

c. 2

FREE RINGS AND THEIR RELATIONS

Second edition

P. M. Cohn, FRS

*Department of Mathematics,
University College London*



1985



Academic Press

(Harcourt Brace Jovanovich, Publishers)

London Orlando San Diego New York
Toronto Montreal Sydney Tokyo

COPYRIGHT © 1985, BY ACADEMIC PRESS INC. (LONDON) LTD.
ALL RIGHTS RESERVED.
NO PART OF THIS PUBLICATION MAY BE REPRODUCED OR
TRANSMITTED IN ANY FORM OR BY ANY MEANS, ELECTRONIC
OR MECHANICAL, INCLUDING PHOTOCOPY, RECORDING, OR
ANY INFORMATION STORAGE AND RETRIEVAL SYSTEM, WITHOUT
PERMISSION IN WRITING FROM THE PUBLISHER.

ACADEMIC PRESS INC. (LONDON) LTD.
24-28 Oval Road
LONDON NW1 7DX

United States Edition published by
ACADEMIC PRESS, INC.
Orlando, Florida 32887

British Library Cataloguing in Publication Data

Cohn, P. M.
Free rings and their relations.—2nd ed.—
(Monographs/London Mathematical Society,
ISSN 0076-0560)
1. Associative rings
I. Title II. Series
512'.4 QA251.5

Library of Congress Cataloging in Publication Data

Cohn, P. M. (Paul Moritz)
Free rings and their relations.

Bibliography: p.
Includes indexes.
1. Associative rings. I. Title.
QA251.5.C62 1985 512'.4 85-1421
ISBN 0-12-179152-1 (alk. paper)

PRINTED IN THE UNITED STATES OF AMERICA

*To the memory
of my parents*

Preface to the second edition

It is not your duty to complete the work,—
But neither are you free to desist from it.

R. Tarphon, Sayings of the Fathers

The present edition has been completely reset, and this has given me an opportunity to make substantial changes. In many respects it is now possible to see more clearly how the subject ought to be developed; the resulting changes, while in themselves small, have affected nearly every page and have, I hope, helped to make the book easier to read. At the same time a number of new developments have been included, particularly: (i) Sylvester domains (5.5–6, 7.10), (ii) localization *à la* Gerasimov and Malcolmson (7.11), (iii) automorphisms of free algebras (6.8–11), (iv) normal forms of matrices over free algebras (5.8) and (v) a closer study of the universal field of fractions of a semifir (7.6–9). Some older topics, which now have more accessible proofs, have been included, such as (vi) the computation of the dependence number, leading to a simple construction of n -firs with prescribed properties (2.11), (vii) the specialization lemma in a new and more general form (5.9) and (viii) the Malcev–Neumann construction (8.8). To give one example, the embeddability of a free algebra in a field, for which there were two proofs in the old edition, has five proofs in the present one. There have also been many smaller innovations, both in the forms of the results and in the proofs. Of the 86 open problems of the first edition, 28 have been solved; most of the solutions are included here, as well as some new problems.

These changes have inevitably caused the volume to grow, but I have been conscious of the need to keep the size within bounds. However, I did not feel justified in leaving out topics from the first edition (except for some better-known background results, for which a reference is readily available), in order to avoid a situation where the first edition continues as a reference for the parts omitted from the second. As in the previous edition, the bibliography is selective but aims to include most of the relevant recent papers; this has meant that some papers previously included have now been omitted, and the total has not quite doubled in size.

It goes without saying that all known errors have been corrected; in this task I have been helped by many correspondents, to whom I would like to

express my appreciation. My thanks also go to L. A. Bokut', who was responsible for the Russian translation and who provided me with a list of errors as well as further references. I have continued to benefit from George M. Bergman's comments on the work, as well as his very substantial help with recent papers of mine which are used in the text. From October 1982 to June 1984 the Bedford College Ring Theory Study Group went through the manuscript on which this edition is based, and I am grateful to the participants for their help and persistence. I owe a special debt to Warren Dicks who played a major role in the Study Group; in particular, he presented his own account of automorphisms and invariants of free algebras, and his notes have (with little change) formed the basis of sections 6.8–11. He also read and commented on substantial parts of the manuscript. This often saved me from error and helped me to see the wood among the trees (*cf.* p. 253). I have been particularly conscious of his help with homological proofs, which are briefer and clearer as a result. The manuscript has also been read by Mark L. Roberts and both he and Mark Hedges have read the proofs and in the process caught a number of errors; I am grateful for their help. Finally I should like to thank the staff of Academic Press for the very efficient way they have coped with the publication of this volume.

University College London
April 1985

P. M. Cohn

From the preface to the first edition

I have had a dream,—
past the wit of man to say what dream it was:
man is but an ass,
if he go about to expound his dream.

Shakespeare, *A Midsummer Night's Dream*

In the development of ring theory there are two strands, each with its own problems and methods, and although with many points of contact, they have never merged completely. One of them is the theory of algebras; this was a non-commutative (indeed sometimes even non-associative) theory from the beginning but there were always heavy finiteness restrictions, which are only gradually being relaxed. Thus while there is a fairly substantial theory of Artinian rings, the theory of Noetherian rings is still in its early stages, and although it is being developed vigorously, it is clear that some type of maximum condition is essential for its development.

A second and quite distinct line originated with the study of arithmetic in algebraic number fields. In the hands of Kummer, Dedekind and E. Noether this led to the abstract notion of a Dedekind ring. Meanwhile, algebraic geometers found the need for affine rings in the study of algebraic varieties; here Dedekind rings appear again, but as a rather special case (essentially the 1-dimensional case). Now in the last few years our way of describing the geometrical notions has changed quite radically, with the result that the correspondence (rings) $\rightarrow$ (varieties) has been extended and made precise: There is a contravariant functor from the category of commutative rings to the category of affine schemes, which is an equivalence. The effect of this connexion has been profound in both directions; we are concerned here particularly with the help afforded by the geometrical notions in studying rings. Although some hypotheses may need to be imposed to obtain particular results, there is only one overall restriction left: the rings must be commutative. Moreover, it is not at all clear that this restriction is essential, but simply that there is so far not even a proper beginning of 'non-commutative algebraic geometry'.

To make such a beginning one would need to study solutions of algebraic

equations in non-commuting indeterminates with coefficients in skew fields. This in itself is a daunting task, raising many problems of its own, and if one is to attack them successfully, one will need some knowledge of non-commutative polynomial rings over skew fields, i.e. free associative algebras. Such knowledge is still rather scanty, in fact since free algebras come under neither of the two main branches, they are not even mentioned in most books on ring theory.

It is the object of this book to present an account of free algebras and related rings, as far as they are known today. Our basic objects of study will be *firs* (free ideal rings) and their variants; this is a class of rings arising rather naturally and including free algebras as well as many other types. In fact it stands in the same relation to free algebras as commutative principal ideal domains do to polynomial rings in one variable over a field. Much of the theory of *firs* is a natural generalization of the commutative theory, although there is a wider range of possibilities here.

Although on the face of it, the topic of this book is rather specialized, our aim has been to bring out the connexion with both of the major branches of ring theory, and to emphasize that all is part of the same subject. Secondly it is hoped that the information on free rings provided here will help to further the study of 'non-commutative algebraic geometry'. Thirdly, the methods used here on *firs* may give some indication of what is needed to study more general rings.

Throughout the book a substantial number of exercises and open problems has been included. The former range from simple questions to test the reader's understanding to brief versions of further developments and counter-examples. The open problems range from open-ended exercises to deep problems whose solution would materially advance the subject. It goes without saying that the author would be glad to hear from any reader who discovers errors in the text or exercises. He would even more like to hear from readers who succeed in solving an open problem.

Finally a word about the sources. Early results in this field, although scanty, are widely scattered; we have tried, in the bibliography at the end, to include most papers of interest. The systematic development of the subject dates from about 1960; a landmark was the appearance of G. M. Bergman's thesis in 1967, in which some outstanding problems were solved, and which simplified and generalized much of what was then known. The material presented here includes roughly the first half of Bergman's thesis; that source has been drawn on freely. It is a pleasure to acknowledge the very great debt I owe to Professor Bergman for allowing me to use his thesis and other unpublished work so extensively for this book, as well as offering much helpful criticism and advice during its preparation.

I have also benefitted greatly from the advice of friends on whom I tried my ideas, and the comments of auditors at my lectures on the topic over the last ten years in three continents. Mr Warren Dicks has read most of the proofsheets and made some valuable comments. My thanks go to all of them, as well as to the staff of the Academic Press for their helpfulness and their efficiency in getting the book published.

Bedford College,
London
November 1971

P. M. Cohn

Contents

	<i>Page</i>
<i>Preface to the second edition</i>	vii
<i>From the preface to the first edition</i>	ix
<i>Note to the reader</i>	xvii
<i>Some terminology, notations and conventions used throughout the book</i>	xviii
<i>List of special notation</i>	xxi
 CHAPTER 0 PRELIMINARIES	 1
0.1 Matrix rings	1
0.2 Rank conditions on free modules	6
0.3 Projective modules	10
0.4 Hermite rings	15
0.5 The monoid of projectives	19
0.6 The matrix of definition of a module	23
0.7 Eigenrings and centralizers	30
0.8 Groups and rings of fractions	34
0.9 Modules over Ore domains	46
0.10 Skew polynomial rings	50
0.11 Free associative algebras and tensor rings	59
Notes and comments	62
 CHAPTER 1 FIRS, SEMIFIRS AND n -FIRS	 64
1.1 Dependence relations	64
1.2 Firs and α -firs	71
1.3 Strong G_n -rings	74
1.4 Homological properties of firs and semifirs	77
1.5 Further properties of n -firs	80
1.6 Inert extensions	83
Notes and comments	85
 CHAPTER 2 THE WEAK ALGORITHM	 87
2.1 The division algorithm	87
2.2 The n -term weak algorithm	94

2.3	The associated graded ring	100
2.4	Weak algebra bases in filtered rings and free algebras	104
2.5	The Hilbert series of a filtered ring.	107
2.6	The construction of rings with a weak algorithm	111
2.7	Generators and relations for $\text{GE}_2(R)$	115
2.8	The 2-term weak algorithm	121
2.9	The inverse weak algorithm	124
2.10	The transfinite weak algorithm	138
2.11	Computation of the dependence number	143
	Notes and comments	150
CHAPTER 3 FACTORIZATION		153
3.1	The commutative case	153
3.2	Similarity in semifirs	158
3.3	Factorization in matrix rings over semifirs	164
3.4	Rigid factorizations	170
3.5	Factorization in semifirs: a closer look	177
3.6	The primary decomposition	184
	Notes and comments	191
CHAPTER 4 RINGS WITH A DISTRIBUTIVE FACTOR LATTICE		193
4.1	Distributive modules	193
4.2	Distributive factor lattices	199
4.3	Conditions for a distributive factor lattice	204
4.4	Finite distributive lattices	211
4.5	The factor lattice	215
4.6	Eigenrings	219
	Notes and comments	226
CHAPTER 5 MODULES OVER FIRS AND SEMIFIRS		227
5.1	Bound and unbound modules	227
5.2	Duality.	232
5.3	Positive and negative modules over semifirs	235
5.4	The ranks of matrices	245
5.5	Sylvester domains.	252
5.6	Stable inner rank and pseudo-Sylvester domains	261
5.7	The factorization of matrices over semifirs	265
5.8	A normal form for matrices over a free algebra	272

5.9	The specialization lemma	281
5.10	Ascending chain conditions	287
5.11	The intersection theorem for firs	294
	Notes and comments	296

CHAPTER 6 CENTRALIZERS, SUBALGEBRAS AND AUTOMORPHISMS

298

6.1	Commutative subrings and central elements in 2-firs	298
6.2	2-Firs with prescribed centre	305
6.3	The centre of a fir	309
6.4	Bounded elements in 2-firs	312
6.5	Free monoids	321
6.6	Subalgebras and ideals of free algebras	329
6.7	Centralizers in power series rings and in free algebras	337
6.8	Automorphisms of polynomial rings	341
6.9	Automorphisms of free algebras	352
6.10	Invariants in free algebras	362
6.11	The Galois correspondence	371
	Notes and comments	378

CHAPTER 7 SKEW FIELDS OF FRACTIONS

381

7.1	The rational closure of a homomorphism	381
7.2	The category of R -fields and specializations	388
7.3	Matrix ideals	395
7.4	Prime matrix ideals	400
7.5	Fields of fractions	412
7.6	Numerators and denominators	421
7.7	The depth	432
7.8	Centralizers in universal fields of fractions of firs	437
7.9	Determinants and valuations	445
7.10	Localization of semifirs and firs	453
7.11	Universal localization: another route	462
	Notes and comments	486

CHAPTER 8 PRINCIPAL IDEAL DOMAINS

489

8.1	The diagonal reduction for matrices over a principal ideal domain	489
8.2	Finitely generated modules over principal ideal domains	492
8.3	Skew polynomial rings	497

8.4	Theory of a single pseudo-linear transformation	503
8.5	Eigenvalues and algebraic matrices	506
8.6	Algebraic skew field extensions	515
8.7	Laurent series and the Malcev–Neumann construction	521
8.8	Iterated skew polynomial rings	532
	Notes and comments	538

APPENDIX

1.	Lattice theory	540
2.	Categories and homological algebra	544
	Bibliography and author index	557
	Subject index	581

Note to the reader

Chapter 0 consists of background material from ring theory that may not be entirely standard, whereas the appendix gives a summary of results from lattice theory and homological algebra, with reference to proofs (or in some cases, proofs). The main subject matter of the book is introduced in Chapter 1 and the reader may wish to start here, referring back to Chapter 0 only when necessary. Chapters 2 and 3 are used throughout (at least the earlier parts), as is Chapter 5, while Chapters 4 and 6 are largely (though not entirely) independent of the rest. The subject matter of Chapter 8 is likely to be more familiar to the reader, and this chapter can be read at any stage, if the reader is willing to turn back for the occasional definition. The first half of Chapter 7 (7.1–5) is quite independent of the rest, except for some applications in 7.5, and it can also be read at any stage.

All theorems, propositions, lemmas and corollaries are numbered consecutively in a single series in each section, thus Th. 2.4 is followed by Cor. 2.5, and this is followed by Lemma 2.6, in Section 4.2 (and outside Chapter 4 they are referred to as Th. 4.2.4, Cor. 4.2.5, Lemma 4.2.6, respectively). The end (or absence) of a proof is indicated by ■. There are exercises at the end of each section; the harder ones are marked * and open-ended (or open) problems are marked °.

References to the bibliography are by author's name and last two digits of year of publication (if after 1900), thus Ore [31], with primes to distinguish publications by the same author in the same year.

Some terminology, notations and conventions used throughout the book

All rings occurring are associative, but not generally commutative (in fact, much of the book reduces to trivialities in the commutative case). Every ring has a unit-element, denoted by 1, which is inherited by subrings, preserved by homomorphisms and acts as the identity operator on modules. The same convention applies to monoids. A ring may consist of 0 alone; this is so precisely if $1 = 0$, and R is the zero ring.

We shall adopt the convention of writing (as far as practicable) homomorphisms of left modules on the right and vice versa. Mappings will be composed accordingly; this applies in particular to multiplication in endomorphism rings and it obviates in many cases the need for passing to the opposite ring. If R is a ring, then for any left R -module M , the *dual* is $M^* = \text{Hom}_R(M, R)$; it is a right R -module. Similarly on the other side. With any property of a ring we associate its left-right dual, which is the corresponding property of the opposite ring. Left-right duals of theorems, etc., will not usually be stated explicitly.

In any ring R the set of non-zero elements is denoted by $R^\times$, but this notation is mostly used for integral domains, where $R^\times$ is closed under multiplication. If $R^\times$ is a group under multiplication, R is called a *field*; occasionally the prefix 'skew' is used, to emphasize that our fields need not be commutative. An element u in a ring or monoid is *invertible* or a *unit* if it has an *inverse* u^{-1} satisfying $uu^{-1} = u^{-1}u = 1$. Such an inverse is unique if it exists at all. The units of a ring (or monoid) R form a group, denoted by $U(R)$. The ring of all $n \times n$ matrices over R is written $\mathfrak{M}_n(R)$ or R_n . The set of all square matrices over R is denoted by $\mathfrak{M}(R)$, and the set of all matrices over R by $\text{Mat}(R)$. Instead of $U(R_n)$ one also writes $\text{GL}_n(R)$, and the union of the $\text{GL}_n(R)$ is written $\text{GL}(R)$ (here 'union' is understood in the sense explained in 1.3).

An element u of a ring is called a *left zero-divisor* if $u \neq 0$ and $uv = 0$ for some $v \neq 0$; if u is neither 0 nor a left zero-divisor, it is called *right regular*. Thus u is right regular whenever $uv = 0$ implies $v = 0$. Corresponding definitions hold with left and right interchanged. A left or right zero-divisor is called a *zero-divisor*, and an element that is neither 0 nor a zero-divisor is

called *regular*. Over a field a square matrix which is a zero-divisor or 0 is also called *singular*; thus 'regular' equals 'non-singular' in this case.

An element of a ring is called an *atom* if it is a regular non-unit and cannot be written as a product of two non-units. A *proper* factorization is one in which all factors are non-units, and a *complete* factorization is one in which all factors are atoms. An integral domain is said to be *atomic* if every element other than zero or a unit has a complete factorization.

Two elements a, b of a ring (or monoid) R are *associated* if $a = ubv$ for some $u, v \in U(R)$. If $u = 1$ ($v = 1$), they are *right* (*left*) *associated*; if $u = v^{-1}$, they are *conjugate* under $U(R)$. A polynomial in one variable (over any ring) is said to be *monic*, if the coefficient of the highest power is 1. Two elements a, b of a ring R are *left coprime* if they have no left factor apart from units; they are *right comaximal* if $aR + bR = R$. Clearly two right comaximal elements are left coprime, but not necessarily conversely. A row $(a_1, \dots, a_n)$ is said to be *unimodular* if the right ideal generated by the a_i is the whole ring; thus a pair is unimodular precisely when it is right comaximal. Two elements a, b are *right commensurable* if there exist a', b' such that $ab' = ba' \neq 0$. Again, corresponding definitions hold on the other side.

Let A be a commutative ring; by an A -*algebra* we understand a ring R which is an A -module such that the multiplication is bilinear. Sometimes we shall want a non-commutative coefficient ring A ; this means that our ring R is an A -bimodule such that $x(yz) = (xy)z$ for any x, y, z from R or A ; this will be called an A -*ring*. To rephrase the definitions, an A -ring is a ring R with a homomorphism $\alpha \mapsto \alpha \cdot 1$ of A into R , while an A -algebra is a ring R with a homomorphism of A into the centre of R . Moreover, the use of the term ' A -algebra' implies that A is commutative. Frequently our coefficient ring will be a skew field, usually written K , or also k when it is assumed to be commutative.

Let R be an A -ring. A family (u_i) of elements of R is *right linearly dependent* over A or *right A -dependent* if there exist $\lambda_i \in A$ almost all (i.e. all but a finite number) but not all zero, such that $\sum u_i \lambda_i = 0$. In the contrary case (u_i) is *right A -independent*. Occasionally we speak of a *set* being linearly dependent; this is to be understood as a family indexed by itself. For example, two elements of an integral domain R are right commensurable if and only if they are right R -dependent and both non-zero.

The maximum condition or ascending chain condition in a ring or module is abbreviated as ACC; if a module satisfies ACC on submodules with at most n generators, we shall say that it satisfies ACC_n . In particular, left (right) ACC_n for a ring R is the ACC on n -generator left (right) ideals of R .

If A, B are matrices, we write $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ as $A \oplus B$ or $\text{diag}(A, B)$. We also

sometimes write columns as rows, with a superscript T to indicate transposition (reflexion in the main diagonal). In such cases the blocks are to be transposed as a whole; thus $(A, B)^T$ means $\begin{pmatrix} A \\ B \end{pmatrix}$, not $\begin{pmatrix} A^T \\ B^T \end{pmatrix}$.

The letters $\mathbf{N}$, $\mathbf{N}_+$, $\mathbf{Z}$, $\mathbf{F}_p$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ stand as usual for the set (respectively ring) of non-negative integers, all positive integers, all integers, all integers mod p , rational, real and complex numbers, respectively, and $\text{card}(I)$ or $|I|$ denotes the cardinal of the set I . If $T \subseteq S$, the complement of T in S is written $S \setminus T$.

List of special notation

(Notation that is either standard or only used locally has not been included.)

$R^\times$	Set of non-zero elements in a ring R
$U(R)$	Group of units in a ring R
${}^m M^n$	Set of all $m \times n$ matrices over a module, p. 3
1_0	0×0 matrix, p. 3
$\text{Lat}_R(M)$	Lattice of all R -submodules of M , p. 3
${}_R \mathcal{M}$	Category of all left R -modules, p. 4
$\sim$	Morita-equivalent rings, p. 4; similar elements, p. 158
$V_{m,n}$	Canonical non-IBN ring, p. 11
$\mathcal{P}(R)$	Monoid of projective R -modules, p. 19
${}_R \mathcal{P}$	Category of finitely generated projective left R -modules, p. 19
$J(R)$	Jacobson radical of R , p. 20
$\chi(M)$	Characteristic of a module, p. 24
$i(A)$	Index of a matrix, p. 28
$I(-)$	Idealizer, p. 30; set of invariant elements, p. 156
$E(-)$	Eigenring, p. 30
$\mathcal{G}(-)$	Universal group of a monoid, p. 35
$l(a)$	Length of a (in a free monoid or UFD), p. 60
$d(f)$	Degree of f , p. 60
$o(f)$	Order of f , p. 60, 527
$L(Rc, R)$	Lattice of principal left ideals containing Rc , p. 80
λ_v	Dependence number, p. 96
$r_h(-)$	Number of v -generators in degree h , p. 98
$H(X), H(R:K)$	Hilbert series, p. 107
μ_v	Inverse dependence number, p. 125
$\mathfrak{B}_n$	Matrix reduction functor, p. 147
$\mathcal{T}, \text{Tor}$	Category of torsion modules (over a semifir), p. 165
$\text{Tr}(-)$	Transpose of a module, p. 233
Pos, Neg	Category of positive, negative modules (over a semifir), p. 236. (Pos is also used in 4.4 for the category of finite partially ordered sets, p. 211.)
M_b	Bounded component of M , p. 237
ρ	Inner rank of a matrix, p. 248

$\mathcal{D}, D$	Singularity support of a matrix (over K^X or $X(R)$), p. 286, 410
X^*	Free monoid on an alphabet X , p. 321
$ f $	Length of a word f , p. 322; leading form, p. 344
$\ f\ $	Top component of f , p. 344
d, D	Degree, bidegree, p. 344
$\deg$	Degree, p. 349
$\mathfrak{M}(R)$	Set of all square matrices over R , p. 391
Ker	Singular kernel, p. 391
∇	Determinantal sum, p. 396
$\sqrt{}$	Radical of a matrix ideal, p. 402
$X(R)$	Field spectrum of R , p. 410
$\Phi(R)$	Set of all full matrices over R , p. 415
λ, ρ	Operations on admissible matrices, p. 421
$D(R)$	Divisor group of R , p. 451
$\parallel$	Total divisor, p. 489

0. Preliminaries

This Chapter collects some facts on rings which form neither part of our subject proper, nor part of the general background (described in the appendix). By its nature the contents is rather mixed, and the reader may well begin with Chapter 1 and only turn back when necessary.

In 0.1 we describe the special form taken by Morita theory in the case of interest to us, the projective free rings. These rings and a generalization are discussed in 0.4, and the effect of these properties on the monoid of projectives $\mathcal{P}(R)$ is considered in 0.5. In 0.2 we discuss the invariant basis number, weak finiteness and an intermediate condition, while 0.3 deals with hereditary rings and also includes Kaplansky's theorem on projective modules, with an application.

0.6 describes the relation between a module and its defining matrix and in particular the condition for two matrices to define isomorphic modules. This and the results on eigenrings and centralizers in 0.7 are mainly used in Chapter 3.

The Ore construction of rings of fractions is behind much of the later development even when this does not appear explicitly. In 0.8 we outline this construction for monoids and rings, as well as some other methods of embedding monoids in groups, used later. This is followed in 0.9 by a discussion of modules over Ore domains; it turns out that the (left or right) Ore condition has some unexpected consequences.

In 0.10 we define skew polynomial rings and power series rings, while 0.11 is devoted to free algebras and tensor rings, the main object of study in this book.

0.1 Matrix rings

Let R be a ring and $n \geq 1$; then the $n \times n$ matrix ring over R is denoted by $\mathfrak{M}_n(R)$ or R_n . It may be defined as the ring generated by R and n^2 elements e_{ij} ($i, j = 1, \dots, n$), called the *matrix basis*, satisfying the conditions $ae_{ij} = e_{ij}a$ ($a \in R$) and

$$e_{ij}e_{kl} = \delta_{jk}e_{il}, \quad \sum e_{ii} = 1, \quad (1)$$

where δ_{ij} is the *Kronecker delta*: $\delta_{ij} = 1$ if $i = j$, and 0 otherwise. The general element of R_n is then uniquely expressible as $\sum a_{ij}e_{ij}$ ($a_{ij} \in R$).

It is noteworthy that matrix rings are characterized by (1), which in effect is a decomposition of 1 in R into n isomorphic idempotents.

PROPOSITION 1.1. *Let S be any ring with n^2 elements e_{ij} satisfying (1). Then $S \cong R_n$, where R is the centralizer of all the e_{ij} .*

Proof. If for each $a \in S$ we put $a_{ij} = \sum_v e_{vi} a e_{jv}$, then it is easily checked that $a_{ij} \in R$ and $a = \sum_{ij} a_{ij} e_{ij}$. Now the correspondence $a \leftrightarrow (a_{ij})$ is seen to be an isomorphism $S \cong R_n$. ■

Matrix rings arise naturally as endomorphism rings of direct powers of modules. Let M be a left R -module and M^n the direct sum of n copies of M . If μ_i is the natural inclusion map from the i th summand into M^n and π_i the natural projection from M^n to the i th factor, then we have the equations (where composition is from left to right, in accordance with our conventions)

$$\mu_i \pi_j = \delta_{ij}, \quad \sum \pi_i \mu_i = 1. \quad (2)$$

It follows that the n^2 R -endomorphisms ε_{ij} of M^n defined by

$$\varepsilon_{ij} = \pi_i \mu_j \quad (3)$$

satisfy the equations (1) for a matrix basis.

Let us write

$$E = \text{End}_R(M). \quad (4)$$

Then each element (α_{ij}) of the $n \times n$ matrix ring E_n defines an endomorphism α of M^n by the equation

$$\alpha = \sum \pi_i \alpha_{ij} \mu_j,$$

and the correspondence $(\alpha_{ij}) \mapsto \alpha$, with inverse $\alpha \mapsto (\mu_i \alpha \pi_j)$, is just another expression of the ring isomorphism $E_n \cong \text{End}_R(M^n)$ already encountered in Prop. 1.1. We record this as

PROPOSITION 1.2. *Let M be any R -module and $E = \text{End}_R(M)$. Then for any $n \geq 1$,*

$$\text{End}_R(M^n) \cong E_n. \quad (5)$$

When we considered M^n as R -module, its endomorphism ring turned out to be E_n . But we can also consider M^n as R_n -module; in that case its endo-

morphism ring, i.e. the centralizer of R_n in $\text{End}(M^n)$, is the centralizer of the matrix basis $\{\varepsilon_{ij}\}$ in E_n , i.e. E itself. Thus we have

$$\text{End}_{R_n}(M^n) \cong E. \quad (6)$$

In the two cases (5) and (6), M^n may be visualized as consisting of column vectors and row vectors, respectively, over M . We shall distinguish these cases by writing the set of column vectors as ${}^nM$ and the set of row vectors as M^n . More generally we denote by ${}^mM^n$ the set of all $m \times n$ matrices with entries in M , and omit reference to m or n if either equals 1. For a ring R , R_n is just ${}^nR$, considered as a ring. We shall also allow m or n to be 0. Thus ${}^0M^n$ is the set of matrices with no rows and n columns; there is one such matrix for each n (including $n = 0$). Similarly for ${}^mM^0$; of course R_0 is the zero ring. The unique 0×0 matrix over R will be written 1_0 , and an $m \times n$ matrix where $mn = 0$ will be called a *null matrix*.

If M is an (R, S) -bimodule, then ${}^mM^n$ is an (R_m, S_n) -bimodule in a natural way. As an example, take R itself, considered as R -bimodule; the set of row vectors R^n has a natural (R, R_n) -bimodule structure and the set of column vectors ${}^nR$ a natural (R_n, R) -bimodule structure. Writing $R_{R, R}R$ for R as right, resp. left R -module and $\rho_a: x \mapsto xa$, $\lambda_a: x \mapsto ax$ for the right resp. left multiplication by a , we have $\text{End}_R({}_R R) \cong R$ via the map $a \mapsto \rho_a$ and $\text{End}_R(R_R) \cong R^o$ via the map $a \mapsto \lambda_a$, where the opposite ring R^o means that the R -endomorphisms of R_R form a ring anti-isomorphic to R (because $\lambda_{ab} = \lambda_b \lambda_a$). In this case, equations (5) and (6) become

$$\text{End}_R(R^n) \cong R_n, \quad \text{End}_{R_n}(R^n) \cong R^o. \quad (7)$$

The row vectors $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots)$, ... and the corresponding columns e_i^T form bases for R^n , ${}^nR$ respectively, as R -modules, called the *standard bases*.

Returning to the case of a general R -module M , we can summarize the relation between M and ${}^nM$ as follows.

THEOREM 1.3. *Let R be a ring and M a left R -module with endomorphism ring E . Then ${}^nM$ may be regarded as R_n -module in a natural way, with endomorphism ring E , and there is a lattice isomorphism*

$$\text{Lat}_R(M) \cong \text{Lat}_{R_n}({}^nM), \quad (8)$$

where $\text{Lat}_R(M)$ denotes the lattice of R -submodules of M . Moreover, (R, E) -bimodules correspond to (R_n, E) -bimodules under (8).

Proof. The first assertion is just a restatement of (6). To establish (8) recall that ${}^nM$ consists of column vectors over M ; any submodule N of M

corresponds to a submodule ${}^nN$ of ${}^nM$ and the correspondence

$$N \mapsto {}^nN \quad (9)$$

is an order-preserving map. Conversely, if P is an R_n -submodule of ${}^nM$, then the n projections $\pi_i: P \rightarrow M$ ($i = 1, \dots, n$) are all equal and associate with P a submodule of M . It is easily seen that the correspondence $P \mapsto P\pi_1$ is an order-preserving map inverse to (9), hence (9) is an order isomorphism between lattices, and so a lattice isomorphism. The rest follows because the E -action on M and on ${}^nM$ is compatible with the R -action. ■

It is possible to go beyond Th. 1.3 and obtain an equivalence between the categories of left R -modules and left R_n -modules. For any ring R we denote the category of all left R -modules and all R -homomorphisms between them by ${}_R\mathcal{M}$, and correspondingly denote the category of right R -modules by $\mathcal{M}_R$. Now the categories ${}_R\mathcal{M}$ and ${}_{R_n}\mathcal{M}$ are seen to be equivalent by the functors

$$M \mapsto {}^nM = {}^nR \otimes_R M \quad (M \in {}_R\mathcal{M}), \quad (10)$$

$$P \mapsto P^{\natural} = \text{Hom}_{R_n}({}^nR, P) = R^n \otimes_{R_n} P \quad (P \in {}_{R_n}\mathcal{M}). \quad (11)$$

It may be verified that $({}^nM)^{\natural} \cong M$, $(P^{\natural}) \cong P$, and this shows that for any ring R and any $n \geq 1$, ${}_R\mathcal{M}$ is equivalent to ${}_{R_n}\mathcal{M}$.

Generally, two rings R, S are said to be *Morita-equivalent*: $R \sim S$, if the categories ${}_R\mathcal{M}$ and ${}_S\mathcal{M}$ are equivalent, so what we have proved is that $R \sim R_n$ for any ring R and any $n \geq 1$. In general, a ring R may be Morita-equivalent to a ring not of the form R_n (e.g. when R is itself a matrix ring), but in a special case of importance to us later, this is the only possibility:

THEOREM 1.4. *Let R be a ring over which every finitely generated projective left R -module is free. Then the rings Morita-equivalent to R are precisely the full matrix rings R_n ($n = 1, 2, \dots$).*

Proof. We have seen that $R \sim R_n$. Conversely, let $S \sim R$; then S , being a finitely generated projective left S -module, corresponds to a finitely generated projective left R -module, since this is a categorical property. If S corresponds to R^n in this way, then these modules have isomorphic endomorphism rings, hence by Prop. 1.2 and (7), $S \cong R_n$. ■

The condition of being Morita-equivalent is invariant under change of sides, i.e. $R \sim S$ if and only if $R^o \sim S^o$. This is not immediately obvious from the definition and we shall not prove it here in general, but in the case of

interest to us it will follow from Cor. 4.7 below (for general Morita theory see Bass [68] or Cohn [76]).

The equivalence between R and R_n may be used to reduce any categorical question concerning a finitely generated module to a cyclic module. For, given M , generated as left R -module by $u_1, \dots, u_n$ say, we apply the functor (10) and pass to the left R_n -module ${}^nM$; this is generated by the single element $(u_1, \dots, u_n)^T$. Thus we have proved

THEOREM 1.5. *Any finitely generated R -module M corresponds to a cyclic R_n -module under the equivalence (10), for a suitable n . Here it is enough to take n equal to the cardinal of some generating set of M . ■*

For example, if R is a principal ideal ring, then as is well known (cf. Prop. 8.2.3), any submodule of an n -generator module over R can be generated by n elements. Applying Th. 1.5 we see that any submodule of a cyclic R_n -module is again cyclic. In particular, R_n is again a principal ideal ring. In the other direction, if R_n is a principal ideal ring, then any submodule of a cyclic R_n -module is cyclic, whence it follows that any submodule of an n -generator R -module can be generated by n elements. This can happen for some $n > 1$ in rings that are not principal ideal rings (cf. Webber [70]).

Exercises 0.1

1. If a ring R satisfies left ACC_n (i.e. ACC on n -generator left ideals), show that every finitely generated free left R -module satisfies ACC_n .
2. If R satisfies left ACC_r , show that R_r satisfies left ACC_s .
3. R is a ring in which all finitely generated left ideals are principal; show that R_n ($n \geq 1$) has the same property.
4. Let R be a ring. If every left ideal can be generated by r elements, show that for any n -generator left R -module M , every submodule of M can be generated by nr elements.
5. If a ring R is injective, as left R -module over itself, show that R_n ($n > 1$) has the same property.
6. For any ring R , show that R and R_n ($n > 1$) have isomorphic centres. Prove this fact by characterizing the centre of R as the set of all natural transformations of the identity functor on ${}_R\mathcal{M}$.
7. Let $R = K_n$ be a full matrix ring and $f: R \rightarrow S$ any ring homomorphism. Show that S is an $n \times n$ full matrix ring $S = L_n$ and that there is a homomorphism $\phi: K \rightarrow L$ which induces f .

0.2 Rank conditions on free modules

Let R be any ring, M an R -module and I a set. The direct power of M with index set I is denoted by M^I , while the direct sum is written $M^{(I)}$.

With every left R -module M we associate the set of linear functionals on M :

$$M^* = \text{Hom}_R(M, {}_R R),$$

which has a natural right R -module structure. Similarly, when M is a right R -module, the linear functionals $M^* = \text{Hom}_R(M, R_R)$ form a left R -module. In particular, $(R^n)^* \cong {}^n R$, $({}^n R)^* \cong R^n$; more generally, if P is a finitely generated projective left R -module, then P^* is a finitely generated projective right R -module and $P^{**} \cong P$.

A free R -module $F = R^{(I)}$ is said to have *rank* $|I|$. When I is infinite, the rank of F is unique (cf. Cohn [77], p. 103), but for finite I this need not be so. We shall say that R^n has *unique rank* if it is not isomorphic to R^m for any $m \neq n$. Using the pairing provided by $*$ we see that R^n has unique rank if and only if ${}^n R$ has unique rank. If F is a free module of unique rank n we write $n = \text{rk}(F)$.

A ring R is said to have the *invariant basis property* or *invariant basis number* (IBN) if every free R -module has unique rank. Most rings commonly encountered have IBN, but we shall meet examples for which this property fails to hold.

Occasionally we shall need stronger properties than IBN. A ring R is said to have *unbounded generating number* (UGN) if for every n there is a finitely generated left R -module which cannot be generated by n elements. Since any n -generator module is a homomorphic image of an n -generator free module, it follows that in a ring with UGN a free module of rank n cannot be generated by less than n elements, and this condition characterizes rings with UGN. It also shows that UGN implies IBN.

A ring R is said to be *weakly n -finite* if every generating set of n elements in R^n is free; if this holds for all n , R is said to be *weakly finite*. Weakly 1-finite rings are also sometimes called 'directly finite', 'v. Neumann finite' or 'inverse symmetric'.

Let R be any ring and suppose that R^n has a generating set of m elements for some m, n , then we have an epimorphism $R^m \rightarrow R^n$, giving rise to an exact sequence

$$0 \rightarrow K \rightarrow R^m \rightarrow R^n \rightarrow 0.$$

Since R^n is free, the sequence splits and so $R^m \cong R^n \oplus K$. This shows that the three properties introduced may be stated as follows:

IBN. For all m, n , $R^m \cong R^n$ implies $m = n$.

UGN. For all m, n , $R^m \cong R^n \oplus K$ implies $m \geq n$.

WF. For all n , $R^n \cong R^n \oplus K$ implies $K = 0$.

In terms of matrices, describing a change of basis, we can express these conditions as follows:

IBN. For any $A \in {}^m R^n$, $B \in {}^n R^m$, if $AB = I_m$, $BA = I_n$, then $m = n$.

UGN. For any $A \in {}^m R^n$, $B \in {}^n R^m$, if $AB = I_m$, then $n \geq m$.

WF. For any $A, B \in R_n$, if $AB = I$, then $BA = I$.

Either set of conditions makes it clear that the zero ring is weakly finite, but has neither IBN nor UGN. For a non-zero ring,

$$\text{WF} \Rightarrow \text{UGN} \Rightarrow \text{IBN},$$

and if R has any of these properties, then so does R° . Moreover, if $R \rightarrow S$ is a homomorphism and S has IBN or UGN, then so does R . Clearly any field (even skew) has all three properties; using determinants we see that every non-zero commutative ring also has all three properties. It is not hard to show that every (left or right) Noetherian ring has the properties (cf. Ex. 4), as does any subring of a skew field. Examples of rings having IBN but not UGN, and rings having UGN but not weakly finite, may be found in Cohn [66].

Let R be a non-zero ring without IBN; then there exist positive integers k, h such that

$$R^h \cong R^{h+k}, \quad h, k \geq 1. \quad (1)$$

The first such pair (h, k) in the lexicographic ordering is called the *type* of the ring R . We observe that for a ring R of type (h, k) , $R^m \cong R^n$ holds if and only if $m = n$ or $m, n \geq h$ and $m \equiv n \pmod{k}$ (cf. e.g. Cohn [81], Th. X.3.2, p. 340).

PROPOSITION 2.1. *Let $f: R \rightarrow S$ be a homomorphism between non-zero rings.† If R does not have IBN and its type is (h, k) , then S does not have IBN and if its type is (h', k') , then $h' \leq h$, $k' \mid k$.*

Proof. By hypothesis (1) holds, hence there exist $A \in {}^m R^n$, $B \in {}^n R^m$ satisfying $AB = I$, $BA = I$, with $m = h$, $n = h + k$. Applying f we get such matrices over S , whence it follows that $S^h \cong S^{h+k}$, so S cannot have IBN and $h' \leq h$, $k' \mid k$. ■

† Here it is important to bear in mind that all our rings have a unit element which is preserved by homomorphisms and inherited by subrings.

The connexion between weak finiteness and UGN is elucidated by the following result:

PROPOSITION 2.2. *A ring R has UGN if and only if some non-zero homomorphic image of R is weakly finite.*

Proof. If some non-zero homomorphic image $\bar{R}$ of R is weakly finite, then $\bar{R}$ has UGN, hence so does R . Conversely, assume that the zero ring is the only weakly finite homomorphic image of R . By adjoining the relations $YX = I$, for all pairs of square matrices X, Y satisfying $XY = I$, we obtain a weakly finite ring $\bar{R}$. For suppose that $I - AB = \sum_1^r U_i(I - Y_i X_i) V_i$, where $X_i Y_i = I$. By taking a diagonal sum of the X 's and Y 's on the right, with $U = (U_1, \dots, U_r)$, $V = (V_1, \dots, V_r)^T$, we can write this as

$$I - AB = U(I - YX)V,$$

and $XY = I$. If A, B are $n \times n$ and X, Y are $m \times m$ ($m \geq n$), then U is $n \times m$ and V is $m \times n$. Let $P \in {}^n R^m$, $Q \in {}^m R^n$ be such that $PQ = I$ and write $Z = APX + U(I - YX)$, $W = YQB + (I - YX)V$, then $ZW = I$ and $I - BA = PX(I - WZ)YQ$. Thus $\bar{R}$ is weakly finite, and it follows that $\bar{R}$ must be the zero ring. Hence R becomes zero by adjoining a finite number of such matrix equations, and by taking diagonal sums we can find a single pair X, Y for which this happens. Thus $XY = I$, while the ideal generated by the entries of $I - YX$ is the whole ring. Replacing each of X, Y by a diagonal sum of an appropriate number of copies, we may assume that there exist $p \in R^n$, $q \in {}^n R$, such that $p(I - YX)q = 1$. Hence

$$I_{n+1} = \begin{pmatrix} X \\ p(I - YX) \end{pmatrix} \begin{pmatrix} Y & (I - YX)q \end{pmatrix}$$

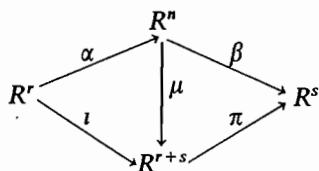
and this equation shows that R does not have UGN. ■

We shall also need the following consequence of weak finiteness (actually equivalent to it):

PROPOSITION 2.3. *Let R be a weakly n -finite ring. Given any sequence of maps between free left R -modules:*

$$R^r \xrightarrow{\alpha} R^n \xrightarrow{\beta} R^s \quad (r+s=n), \quad (2)$$

whose composition is zero, $\alpha\beta = 0$, and such that α has a right inverse and β a left inverse, there exists an automorphism μ of R^n such that the following diagram is commutative:



where i, π are the natural inclusion and projection.

Proof. By hypothesis there exist α', β' such that $\alpha\alpha' = I$, $\beta'\beta = I$, hence $\ker \beta$ is a direct summand of R^n , with complement $(R^n)\beta'\beta' \cong R^s$. Similarly, $(R^r)\alpha \cong R^r$ is a direct summand of $\ker \beta$; denoting the complement by I , we have $R^n \cong (R^r \oplus I) \oplus R^s$. By weak n -finiteness, $I = 0$ and the sequence (2) is split exact. Now a suitable change of basis in R^n achieves the conclusion. ■

The content of this proposition can also be translated into matrix language, and it will be of use to us later in this form. We state it as

PROPOSITION 2.4. *Let R be a weakly n -finite ring and let $A \in {}^rR^n$, $A' \in {}^nR^r$, $B \in {}^nR^s$, $B' \in {}^sR^n$ be such that $AB = 0$, $AA' = I_r$, $B'B = I_s$, where $r+s = n$. Then there exists $P \in \text{GL}_n(R)$ such that*

$$A = (I_r \quad 0)P, \quad B = P^{-1} \begin{pmatrix} 0 \\ I_s \end{pmatrix}.$$

These equations just state that A constitutes the first r rows of P , while B constitutes the last s columns of P^{-1} . To prove the result in this form, we have, by hypothesis,

$$\begin{pmatrix} A \\ B' \end{pmatrix} (A' \quad B) = \begin{pmatrix} I_r & 0 \\ B'A' & I_s \end{pmatrix},$$

where the matrices are $n \times n$. Clearly the right-hand side is invertible, so $(A, B')^T$ has a right inverse and by weak finiteness is invertible. Likewise (A', B) is invertible. ■

Exercises 0.2

1. Show that over a ring of type (h, k) every finitely generated module can be generated by h elements. Find a bound for the least number of elements in a basis of a finitely generated free module.

2. If K is any non-zero ring and I an infinite set, show that $R = \text{End}(K^{(I)})$ does not have IBN and determine its type.

3. If every finitely generated left R -module is cyclic, show that R cannot be an integral domain; in particular, obtain this conclusion for a ring of type $(1, k)$.

4. Show that a ring with left ACC_n for some $n \geq 1$ is weakly n -finite. Deduce that a ring with left (or right) ACC_n for all n is weakly finite. Obtain the same conclusion for DCC_n . (Hint: see Ex. 8.12.)

5. Let R be a ring without IBN and for fixed m, n ($m \neq n$) consider pairs of mutually inverse matrices $A \in {}^m R^n, B \in {}^n R^m$. Show that if A', B' is another such pair, then $P = A'B$ is an invertible matrix such that $PA = A', BP^{-1} = B'$. What is P^{-1} ?

6. Show that the conclusion of Prop. 2.3 is equivalent to weak finiteness.

7. Show that a ring R is weakly n -finite if and only if (F): Every surjective endomorphism of R^n is an automorphism. If R is non-zero and has the property (F), show that every free homomorphic image of R^n has rank at most n . Deduce that a non-zero weakly finite ring has UGN.

8. Show that weak finiteness, UGN, IBN are all Morita invariants.

9°. Characterize the rings all of whose homomorphic images are weakly finite.

10. (Jacobson) Let R be a non-zero ring which is not weakly 1-finite, say $ab = 1 \neq ba$. Writing $e_{ij} = b^{i-1}a^{j-1} - b^i a^j$, show that the e_{ij} are a set of matrix units, and if e is the ideal generated by them, then R/e is the universal weakly 1-finite image of R .

11*. (Montgomery [83]) Let A be an algebra over the real numbers with generators a_0, a_1, b_0, b_1 and defining relations $a_0 b_0 - a_1 b_1 = 1, a_1 b_0 + a_0 b_1 = 0$. Show (by using a normal form for the elements of A) that A is an integral domain, hence weakly 1-finite, but not weakly 2-finite. Show also that $A \otimes_{\mathbb{R}} \mathbb{C}$ is not weakly 1-finite. (cf. also Ex. 2.11.8.)

12°. Is the tensor product of two weakly finite k -algebras again weakly finite?

13°. Is every weakly 1-finite v. Neumann regular ring weakly finite?

14. Let $V_{m,n}$ be a k -algebra with $2mn$ generators, arranged as an $m \times n$ matrix A and an $n \times m$ matrix B and defining relations $AB = I, BA = I$ (the 'canonical non-IBN ring' for $m \neq n$). Show that $V_{1,n}$ is a simple ring for $n > 1$; what is $V_{1,1}$?

15. (M. Kirezci) If $V_{m,n}$ is defined as in Ex. 14 and $m < n$, show that there is a homomorphism $V_{m,n+(n-m)} \rightarrow V_{m,n}$. (Hint: If in $V_{m,n}$ $A = (A_1, A_2), B = (B_1, B_2)^T$, where A_1, B_1 are square, verify that $(A'_1, A'^{-1}_1 A_2, \dots, A_2)$ and $(B'_1, B_2 B'^{-1}_1, \dots, B_2)^T$ are mutually inverse.) Deduce that $V_{1,n}$, for $n > 1$, can be embedded in $V_{1,2}$.

0.3 Projective modules

A well-known result states that if all left ideals of a ring are projective, then all submodules of projective modules are projective. By taking a little care in

the proof it is possible to derive a more precise result, which will be needed later. If α is a cardinal (finite or infinite), then a module is said to be α -generated or an α -generator module if it has a generating set of cardinal not exceeding α . A ring R is called left α -hereditary if every α -generated left ideal of R is projective. Thus to say that R is left $|R|$ -hereditary is to say that every left ideal of R is projective, i.e. R is left-hereditary; a left semihereditary ring is a ring which is left n -hereditary for all n . Corresponding definitions apply on the right.

THEOREM 3.1. *Let α, κ be any cardinals and let R be a left α -hereditary ring. If F is a free left R -module of rank κ , then every α -generated submodule N of F is isomorphic to a direct sum of κ left ideals of R , each α -generated projective.*

Proof. We shall identify cardinals with their least ordinals. It will be enough to prove a direct sum decomposition $N = \bigoplus_{\lambda < \kappa} P_\lambda$, where each P_λ is isomorphic to a left ideal, for then P_λ as homomorphic image of N is α -generated and hence projective.

Let $\{e_i \mid i < \kappa\}$ be a basis of F , put $F_\lambda = \sum_{i < \lambda} Re_i$ and for each $i < \kappa$ write $p_i: F \rightarrow R$ for the i th coordinate projection. If $N \cap F_{i+1}$ is α -generated, then $(N \cap F_{i+1})p_{i+1}$ is an α -generated left ideal of R , hence projective and $N \cap F_{i+1}$ has a submodule $P_{i+1} \cong (N \cap F_{i+1})p_{i+1}$ such that $N \cap F_{i+1} = N_i \oplus P_{i+1}$, where N_i is the kernel of p_{i+1} restricted to $N \cap F_{i+1}$; but this is $N \cap F_i$, so we obtain

$$N \cap F_{i+1} = (N \cap F_i) \oplus P_{i+1}. \quad (1)$$

If this holds for all $i < \kappa$, then it follows that $N = \bigoplus_{i < \kappa} P_{i+1}$ as desired, so it remains to show that for all $i < \kappa$, $N \cap F_{i+1}$ is α -generated. We distinguish two cases.

(i) α is finite. Then N lies in a submodule of F generated by finitely many of the e 's, so we may assume $\kappa = k$ finite. Then $F = F_k$, $N \cap F_k = N$ is α -generated and by (1) for $i+1 = k$, $N \cap F_{k-1}$ is α -generated. Now a downward induction on k shows that each $N \cap F_i$ is α -generated, as we had to prove.

(ii) α is infinite; we shall use induction on α . Let $\{n_\beta \mid \beta < \alpha\}$ be a generating set for N and for each $\beta < \alpha$ denote by N_β the submodule of N generated by all n_γ ($\gamma < \beta$). For each $\lambda < \kappa$ there is a decomposition $F = F_\lambda \oplus F'_\lambda$, where $F'_\lambda = \sum Re_i$ ($\lambda \leq i < \kappa$). Denote by $q_\lambda: F \rightarrow F'_\lambda$ the canonical projection with kernel F_λ ; its restriction to N_β has kernel $N_\beta \cap F_\lambda$. Now $N_\beta q_\lambda$ is a submodule of F'_λ , generated by less than α elements, hence it is projective, by the induction hypothesis. Therefore N_β splits over the kernel $N_\beta \cap F_\lambda$ which is thus a direct summand of N_β and so α -generated. It follows that

$N \cap F_\lambda = \bigcup_{\beta < \alpha} (N_\beta \cap F_\lambda)$ is a union of α submodules, each α -generated and since $\alpha^2 = \alpha$ (cf. e.g. Cohn [81], p. 240), it is α -generated, as claimed. ■

We note the special cases of hereditary and semihereditary rings; when R is hereditary, $\alpha = |R|$ and we can omit the hypothesis on N :

COROLLARY 3.2. *Let R be a left hereditary ring. If F is a free left R -module of rank κ , then every submodule of F is isomorphic to a direct sum of at most κ projective left ideals. ■*

COROLLARY 3.3. *Let R be a left semihereditary ring. If F is a free left R -module of rank κ , then every n -generator submodule of F is isomorphic to a direct sum of at most κ n -generator projective left ideals. ■*

In a principal left ideal domain every left ideal is free of rank at most 1, so by Cor. 3.2 any submodule of a free left module of rank n is again free of rank at most n . Secondly, if R is an integral domain in which every finitely generated left ideal is principal, i.e. R is a *left Bezout domain*, then by Cor. 3.3, every finitely generated submodule of a free left module of rank n is again free of rank at most n . Thus we have

COROLLARY 3.4. *Suppose that either (i) R is a principal left ideal domain and M is a submodule of R^n , or (ii) R is a left Bezout domain and M a finitely generated submodule of R^n , then M is free of rank at most n . ■*

The hypothesis of Cor. 3.3 can still be weakened by using the following remarkable result of Kaplansky. Let R be any ring and M an R -module; we shall say that M is of *type K* if it can be written as a direct sum of (any number of) countably generated modules. Kaplansky's result states that every projective module is of type K. It is convenient to prove a little more:

THEOREM 3.5. *Let R be any ring and M an R -module of type K. Then any direct summand of M is again of type K.*

Proof. Assume that $M = \bigoplus M_i = P \oplus Q$, where the M_i are countably generated. To prove that P is of type K we first show how to express M as a union of a well-ordered ascending sequence, $M = \bigcup S_\alpha$, with the following properties:

- (i) at a limit ordinal α , $S_\alpha = \bigcup \{S_\beta \mid \beta < \alpha\}$,
- (ii) $S_{\alpha+1}/S_\alpha$ is countably generated,

- (iii) each S_α is a direct sum of a subfamily of the M_i and
- (iv) $S_\alpha = P_\alpha \oplus Q_\alpha$, where $P_\alpha = P \cap S_\alpha$, $Q_\alpha = Q \cap S_\alpha$.

At a limit ordinal α the construction of S_α is given by (i); it remains to construct $S_{\alpha+1}$, given S_α . Choose some M_j not contained in S_α and let a countable generating set of M_j be $x_{11}, x_{12}, \dots$. Now split x_{11} into its P - and Q -components; each of these new elements, in the expression of M as the direct sum of the M_i 's, has a non-zero component in only a finite number of the M_i 's. The sum of this finite collection of M_i 's is again countably generated, by $x_{21}, x_{22}, x_{23}, \dots$, say. We now treat x_{12} in the same fashion as x_{11} and obtain another countable sequence $x_{31}, x_{32}, \dots$. We continue in this way to construct an infinite matrix (x_{rs}) , adding enough M_i 's to split each x_{rs} into its P - and Q -components, treating the x 's in the order $x_{11}, x_{12}, x_{21}, x_{13}, x_{22}, x_{31}, \dots$ along successive diagonals. Let $S_{\alpha+1}$ be the submodule generated by S_α and all the x 's, then it is clear that (i)–(iv) hold by construction.

To complete the proof we observe that for each α , P_α is a direct summand of S_α which is a direct summand of M . Hence P_α is a direct summand of M , and it follows that it is a direct summand of $P_{\alpha+1}$. Since $S_{\alpha+1}/S_\alpha \cong P_{\alpha+1}/P_\alpha \oplus Q_{\alpha+1}/Q_\alpha$, we conclude that $P_{\alpha+1}/P_\alpha$ is countably generated (as image of $S_{\alpha+1}/S_\alpha$). At a limit ordinal, $P_\alpha = \bigcup \{P_\beta \mid \beta < \alpha\}$, hence $P \cong \bigoplus P_{\alpha+1}/P_\alpha$ and this shows P to be of type K . ■

Any free module, being a direct sum of copies of R , is of type K , hence we have

COROLLARY 3.6. *Every projective module is a direct sum of countably generated modules.* ■

As an application we shall prove a result of Bergman to be used later, which in effect is a left–right symmetric form of Cor. 3.3. To state it we shall need another definition.

A ring is said to be *weakly semihereditary* if, given two maps between finitely generated projective modules, $\alpha: P_0 \rightarrow P_1$, $\beta: P_1 \rightarrow P_2$ such that $\alpha\beta = 0$, we can write $P_1 = P'_1 \oplus P''_1$ such that $\text{im } \alpha \subseteq P'_1 \subseteq \ker \beta$. In terms of matrices, a ring R is weakly semihereditary if for any $A \in {}^r R^n$, $B \in {}^n R^s$ such that $AB = 0$, there exists an idempotent $n \times n$ matrix E such that $AE = A$, $EB = 0$.

By applying the duality $*$ (or replacing E by $I - E$ in the matrix condition) we see that this condition is left–right symmetric. In a left semihereditary ring, if α, β are as in the above definition, then $\text{im } \beta$ is a projective module (by Cor. 3.3), hence P_1 splits over $\text{im } \beta$, $P_1 \cong \text{im } \beta \oplus \ker \beta$, and $\text{im } \alpha \subseteq \ker \beta$, so the condition of the definition is satisfied. Thus every left semihereditary ring

is weakly semihereditary, and by symmetry the same holds for right semihereditary rings. Now we have

THEOREM 3.7. *Any projective module over a weakly semihereditary ring is a direct sum of finitely generated modules.*

Proof. Let P be a projective module, say $P \oplus Q = F$ is free, and let A be a finite subset of P . We first show that A is contained in a finitely generated direct summand of P .

The elements of A involve only finitely many coordinates in F , hence A lies in a finitely generated free direct summand F_0 of F . Let $|A| = n$ and take $\alpha: R^n \rightarrow F_0$ as the homomorphism mapping a standard basis to A ; denote by β the restriction to F_0 of the projection $F \rightarrow Q$. Then $\alpha\beta = 0$, hence $F_0 = P' \oplus P''$, where $\text{im } \alpha \subseteq P' \subseteq \ker \beta = P \cap F_0$. Thus P' is a direct summand of F_0 and is contained in P , hence it is a direct summand of P and $P' \supseteq \text{im } \alpha \supseteq A$, as we wished to show.

We now complete the proof as follows. By Cor. 3.6, we may take our projective module to be countably generated, by $e_1, e_2, \dots$, say. Suppose that we already have a direct decomposition $P = P_1 \oplus P_2 \oplus \dots \oplus P_n \oplus P'_n$, where each P_i is finitely generated and the sum of the P_i contains $e_1, \dots, e_n$. By the first part of the proof we can find a decomposition $P = Q \oplus Q'$, where Q is finitely generated and contains $e_1, \dots, e_{n+1}$ and a generating set for $P_1, \dots, P_n$. Then $P_1 \oplus \dots \oplus P_n$ is contained in Q and it is a direct summand in P , hence it is a direct summand in Q , say $Q = P_1 \oplus \dots \oplus P_{n+1}$ and $P = P_1 \oplus \dots \oplus P_n \oplus P_{n+1} \oplus Q'$. By induction on n , P has a submodule $P_1 \oplus P_2 \oplus \dots$ containing $e_1, e_2, \dots$; but this set generates P , so P has been expressed as a direct sum of finitely generated modules. ■

Exercises 0.3

1. Let R be a right semihereditary ring and let $A \in {}^m R^n$. Show that the right annihilator of A in R_n has the form ER_n , where E is an idempotent $n \times n$ matrix.

2. (Kaplansky [58]) Let M be a countably generated module over a ring R . Assume that each direct summand N of M is such that any $x \in N$ can be embedded in a finitely generated direct summand of N ; show that M is a direct sum of finitely generated modules.

3. Show that a ring is weakly semihereditary if and only if, given any finitely generated projective left R -module P , a finite subset A of P and a finite subset B of P^* such that $AB = 0$, there exists a direct decomposition $P = P' \oplus P''$ such that $A \subseteq P'$ and $P'B = 0$.

4. Show that any left Noetherian weakly semihereditary ring is left semihereditary.

5. (Bergman [72]) Show that if in the definition of Ex. 3 of weakly semihereditary ring we delete the condition that B resp. A be finite, then we obtain a characterization of left resp. right semihereditary rings.

0.4 Hermite rings

The conditions of IBN, UGN and weak finiteness discussed in 0.2 hold in most rings normally encountered, and counter-examples belong to the pathology of the subject. By contrast, the property defined below forms a significant restriction on the ring.

A module P is said to be *stably free* if $P \oplus R^m \cong R^n$ for some integers m, n . Of course such a P must be finitely generated projective. We remark that if $P \oplus R^m$ is free but not finitely generated, then P is necessarily free (cf. Ex. 6); in any case, we shall mainly be concerned with finitely generated modules.

A ring R is called an *Hermite ring* if it has IBN and any stably free module is free. More specifically, if n -generator modules have unique rank and for any left R -module P , $P \oplus R^r \cong R^m$ ($r \leq m \leq n$) implies that P is free, then R is called an *n -Hermite ring*. Thus, denoting by $\mathcal{H}$ ($\mathcal{H}_n$) the class of all (n)-Hermite rings, we have

$$\mathcal{H}_1 \supseteq \mathcal{H}_2 \supseteq \dots \quad \text{and} \quad \bigcap \mathcal{H}_n = \mathcal{H}.$$

In an n -Hermite ring R , if $P \oplus R^n = R^n$, P is free, say $P = R^r$ and so $R^{n+r} \cong R^n$; by IBN, $r = 0$, so any n -Hermite ring is weakly n -finite.

To describe an Hermite ring in terms of matrices, let us call an $m \times n$ matrix A over a ring R *completable* in R , if either (i) $m = n$ and A is invertible, (ii) $m < n$ and for some $(n-m) \times n$ matrix B , $\begin{pmatrix} A \\ B \end{pmatrix}$ is invertible or (iii) $m > n$ and for some $m \times (m-n)$ matrix C , (A, C) is invertible. We now have

THEOREM 4.1. *Let R be a ring and $n \geq 1$, then R is n -Hermite if and only if*

C_n . *Every $r \times m$ matrix A ($r \leq m \leq n$) with a right inverse is completable.*

Moreover, R is n -Hermite if and only if R^0 is.

Proof. Let R be n -Hermite and suppose that $AB = I_r$. Interpreting A, B as mappings α, β between R^r and R^m , we have a split exact sequence

$$0 \longrightarrow R^r \xrightarrow{\alpha} R^m \xrightarrow{\beta'} P \longrightarrow 0, \quad (1)$$

where $\beta' = \text{coker } \alpha$. Hence $R^m \cong R^r \oplus P$, so $P \cong R^s$, where $r+s=m$. Since

the sequence is split, β' has a left inverse $\alpha': R^s \rightarrow R^m$ and if A', B' are the matrices corresponding to α', β' , then $AB = I$, $A'B' = I$, $AB' = 0$, so by Prop. 2.4, $(A, A')^T$ and (B, B') are both invertible. This shows in effect that an n -Hermite ring satisfies both C_n and the left-right dual C_n^o .

Conversely, if R satisfies C_n and $P \oplus R^r \cong R^m$, where $m-r=s$, then we again have a split exact sequence (1). Let A be the $r \times m$ matrix corresponding to α ; since the sequence is split, there is a map $\beta: R^m \rightarrow R^r$ such that $\alpha\beta = 1$, so if β is represented by B , then $AB = I$. By C_n there exists $A' \in {}^s R^m$ such that $(A, A')^T$ is invertible. Let the inverse (correspondingly partitioned) be (C, C') , then $\beta': R^m \rightarrow R^r$ is represented by C' , hence $P \cong R^s$ and this shows R to be n -Hermite. Now the last part follows because an n -Hermite ring satisfies both C_n and C_n^o , or more directly, by the duality of projective modules. ■

The fact that n -Hermite rings satisfy C_n and C_n^o may be expressed as follows. In an n -Hermite ring R , any matrix equation

$$AB = I_r, \quad \text{where } A \in {}^r R^m, B \in {}^m R^r \quad (r \leq m \leq n), \quad (2)$$

may be completed to an equation

$$\begin{pmatrix} A \\ A' \end{pmatrix} (B \quad B') = I_m = (B \quad B') \begin{pmatrix} A \\ A' \end{pmatrix}. \quad (3)$$

Alternatively we may express the condition by saying that given (2), there exists $P \in \text{GL}_m(R)$ such that

$$A = (I_r \quad 0)P, \quad B = P^{-1} \begin{pmatrix} I_r \\ 0 \end{pmatrix}.$$

We note the following criterion for an integral domain to be a 2-Hermite ring.

PROPOSITION 4.2. *An integral domain R is a 2-Hermite ring if and only if, for any right comaximal pair a, b , $aR \cap bR$ is principal.*

Proof. If a, b are right comaximal, then the mapping $\mu: (x, y)^T \mapsto ax - by$ is a surjective homomorphism of right R -modules ${}^2R \rightarrow R$. Writing $P = \ker \mu$, we have the exact sequence

$$0 \longrightarrow P \longrightarrow {}^2R \xrightarrow{\mu} R \longrightarrow 0. \quad (4)$$

Thus any right comaximal pair a, b leads to a sequence (4); such a sequence necessarily splits to give $R \oplus P \cong {}^2R$. Conversely, if $R \oplus P \cong {}^2R$, we have a split exact sequence (4), and if $(a, -b)$ is the matrix of μ , then a, b are right comaximal. Now $(x, y)^T \in P$ if and only if $ax = by \in aR \cap bR$, thus $P \cong aR \cap bR$, and P is free if and only if $aR \cap bR$ is principal. ■

Over an Hermite ring there is a stability property of matrix factorizations which is sometimes useful.

PROPOSITION 4.3. *Let R be an $(n+r)$ -Hermite ring, $C \in R_n$ and suppose that we have a factorization into square matrices*

$$C \oplus I_r = P_1 \dots P_t.$$

Then there are invertible matrices U_i ($i = 0, \dots, t$), $U_0 = U_t = I$, such that $U_{i-1}^{-1} P_i U_i = P'_i \oplus P''_i$ and $C = P'_1 \dots P'_t$, $I = P''_1 \dots P''_t$.

Proof. By induction it will be enough to treat the case of two factors, thus

$$\begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix} = AB = \begin{pmatrix} A' \\ A'' \end{pmatrix} \begin{pmatrix} B' & B'' \end{pmatrix},$$

with an appropriate block decomposition. We have $A''B'' = I$, and since R is $(n+r)$ -Hermite, there exists an invertible matrix P such that $A'' = (0, I)P$, $B'' = P^{-1} \begin{pmatrix} 0 \\ I \end{pmatrix}$. If we replace A, B by AP^{-1}, PB , we obtain

$$\begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} A_1 & A_2 \\ 0 & I \end{pmatrix} \begin{pmatrix} B_1 & 0 \\ B_2 & I \end{pmatrix}.$$

On multiplying out we find that $A_2 = 0$, $B_2 = 0$, $C = A_1 B_1$ and so the conclusion follows by induction. ■

Let $C \in R_n$; we shall say that C is a *stable atom* if $C \oplus I_r$ is an atom for all $r \geq 1$. From Prop. 4.3 we obtain

COROLLARY 4.4. *Over an Hermite ring every matrix atom is stable.* ■

We now specialize still further. Let $n \geq 1$; a ring R is said to be *n-projective free* if every n -generator projective left R -module is free, of unique rank. Clearly ' n -projective free' implies ' m -projective free' for all $m \leq n$. A ring which is n -projective free for all n is called *projective free*. It is clear from the definition that every n -projective free ring is n -Hermite; thus for non-zero rings the following properties become more restrictive as we go down the list:

1. invariant basis number,
2. unbounded generating number,
3. weak finiteness,
4. Hermite rings,
5. projective free rings.

A finitely generated projective module P may be described as a direct

summand of R^n for some $n: P \oplus P' \cong R^n$, hence P is also determined by the projection of R^n on the direct summand, i.e. an idempotent $n \times n$ matrix. Further, P and P' are free, of ranks adding up to n if and only if there is an automorphism of R^n mapping P to R^r and P' to R^{n-r} . If we interpret this in matrix terms we obtain the following criterion for a ring to be projective free:

PROPOSITION 4.5. *A ring R is n -projective free if and only if R^n has unique rank and each idempotent $n \times n$ matrix C is conjugate (by an invertible matrix) to a matrix of the form $I_r \oplus 0$, for some $0 \leq r \leq n$. ■*

Occasionally a generalization of projective free rings is needed. We shall say that a ring R is *projective trivial* if there exists a projective left R -module P , called the *minimal projective* of R , such that every finitely generated projective left R -module M has the form P^n , for an integer n which is uniquely determined by M . Clearly a projective free ring is projective trivial, with R itself as minimal projective; but note that the property of being projective trivial, unlike that of being projective free, is a Morita invariant. The precise relationship between these two concepts is elucidated in

THEOREM 4.6. *For any ring R , the following properties are equivalent:*

- (a) *R is a full matrix ring over a projective free ring,*
- (b) *R is Morita equivalent to a projective free ring,*
- (c) *R is projective trivial.*

Proof. Clearly (a) $\Rightarrow$ (b) $\Rightarrow$ (c). Now assume (c): R is projective trivial, with minimal projective P , say. Since R is finitely generated projective, we have

$$R \cong P^n \quad (5)$$

for some positive integer n . Write $E = \text{End}_R(P)$, then by Prop. 1.2 we find, on taking endomorphism rings in (5),

$$R \cong E_n.$$

Here E is again projective trivial, and (5) shows that its minimal projective is E itself; hence it is projective free, i.e. (a) holds. ■

From Prop. 4.5 and the left-right symmetry of the condition $R^m \cong R^n$ we deduce that the condition of being n -projective free is left-right symmetric. By Th. 4.6 we have

COROLLARY 4.7. *The condition of being projective trivial is left-right symmetric. ■*

Exercises 0.4

1. Show that a 1-Hermite ring is the same as a weakly 1-finite ring.
2. Let A, B be matrices whose indices have the same sign (cf. 0.6). Show that $A \oplus B$ is completable if and only if A and B are.
3. Describe the rings Morita equivalent to Hermite rings.
4. If in an Hermite ring, $AB = I$ and B is completed to an invertible matrix $\begin{pmatrix} B & B' \end{pmatrix}$, show that for suitably chosen $A', \begin{pmatrix} A & A' \end{pmatrix}$ has the inverse $\begin{pmatrix} B & B' - BAB' \end{pmatrix}$.
5. If R is any commutative ring and $P \oplus R^{n-1} \cong R^n$, show that $P \cong R$. (Hint: In the exterior algebra on P , show that $\Lambda^k P = 0$ for $k > 1$ and deduce the result by calculating $\Lambda^n(P \oplus R^{n-1})$.)
- 6*. (Gabel, cf. Lam [78], p. 23) If P is not finitely generated but $P \oplus R^m = F$, where F is free, show (by writing this as a split exact sequence with middle term F) that we can write $F = F_0 \oplus F_1$, where each F_i is free, F_0 is finitely generated and $F = P + F_0$, $F_0 \cong (P \cap F_0) \oplus R^m$. Deduce that $P/(P \cap F_0) \cong F_1$ and hence show that P is free.
7. (Lam [76]) Let R be any ring and P a projective module which has R as a direct summand. If $P \oplus R^m \cong R^n$, where $n > m$, show that P^{m+1} is free. (Hint: If $P \cong Q \oplus R$, then $P^{m+1} \cong P \oplus R^m \oplus Q^m$; now use R^m to 'liberate' P and the resulting R^n to 'liberate' Q^m .)

0.5 The monoid of projectives

Let R be a ring and denote by ${}_R\mathcal{P}$ ($\mathcal{P}_R$) the category of all finitely generated projective left (right) R -modules and all homomorphisms between them. We have seen that the correspondence $P \mapsto P^*$ defines a duality between ${}_R\mathcal{P}$ and $\mathcal{P}_R$ such that $P^{**} \cong P$.

We shall denote by $\mathcal{S}(R)$ the monoid whose elements are the isomorphism classes of objects of ${}_R\mathcal{P}$; thus each $P \in {}_R\mathcal{P}$ defines an element $[P]$ of $\mathcal{S}(R)$, where $[P] = [P']$ if and only if $P \cong P'$. The operation on $\mathcal{S}(R)$ is given by

$$[P] + [Q] = [P \oplus Q].$$

It is clear that this is well-defined, i.e. the right-hand side depends only on $[P], [Q]$ and not on P, Q . We see that $\mathcal{S}(R)$ is a commutative monoid, in which we may regard $[R]$ as a distinguished element. It is *conical*, i.e. $\alpha + \beta = 0$ implies $\alpha = \beta = 0$. Its universal group (cf. 0.8 below) is the projective module group $\mathbf{K}_0(R)$ (cf. e.g. Milnor [71]). By the duality between ${}_R\mathcal{P}$ and $\mathcal{P}_R$ we have $\mathcal{S}(R^o) \cong \mathcal{S}(R)$, $\mathbf{K}_0(R^o) \cong \mathbf{K}_0(R)$. The element of $\mathbf{K}_0(R)$ corresponding to P is again written $[P]$; now the general element has the

form $[P] - [Q]$ and we have $[P] = [P']$ if and only if $P \oplus S \cong P' \oplus S$ for some $S \in {}_R\mathcal{P}$. Here S may be taken to be free (of finite rank), so we have

$$[P] = [P'] \text{ in } K_0 \text{ if and only if } P \oplus R^n \cong P' \oplus R^n \text{ for some } n. \quad (1)$$

We can equally well define $\mathcal{S}(R)$ in terms of idempotent matrices. Given two idempotent matrices E, F over R , we define E, F to be *isomorphic*, $E \sim F$, if and only if there exist matrices X, Y (generally rectangular) such that $X = EXF$, $Y = FYE$, $XY = E$, $YX = F$. It is easily checked that E, F are isomorphic if and only if the projective left (or right) modules defined by them are isomorphic. Moreover, if E, F correspond to P, Q , respectively, then $E \oplus F$ corresponds to $P \oplus Q$. Thus $\mathcal{S}(R)$ may be defined as the set of isomorphism classes of idempotent matrices with the operation $E \oplus F$.

The structure of $\mathcal{S}(R)$ is closely related to certain properties of the ring R , while $K_0(R)$ reflects the corresponding stable properties. This is illustrated in the next result.

PROPOSITION 5.1. *Let R be any ring and denote by $\lambda: \mathbb{Z} \rightarrow K_0(R)$ the homomorphism defined by $\lambda(1) = [R]$. Then (i) R has IBN if and only if λ is injective and (ii) every finitely generated projective module is stably free if and only if λ is surjective.*

Further, R is projective free if and only if the natural homomorphism $N \rightarrow \mathcal{S}(R)$ is an isomorphism.

Proof. Clearly λ fails to be injective if and only if $n[R] = 0$ in $K_0(R)$ for some $n \neq 0$, say $n > 0$. Then $R^n \oplus P \cong P$, and if $P \oplus P' \cong R^r$, then $R^{n+r} \cong R^r$, which expresses the fact that R does not have IBN. If every finitely generated projective module is stably free, then $P \oplus R^m \cong R^n$, hence $[P] = n[R] - m[R] = (n-m)[R] = \lambda(n-m)$, and conversely, if $[P] = \lambda(r)$, then $P \oplus R^m \cong R^{m+r}$ and P is stably free. The final assertion follows because R is projective free if and only if every P satisfies $[P] = n[R]$ in $\mathcal{S}(R)$, for a unique $n \in \mathbb{N}$, depending on P . ■

It will be useful to relate the monoid $\mathcal{S}(R)$ to $\mathcal{S}(R/J)$, where $J = J(R)$ is the Jacobson radical of R . We recall that $J(R)$ is defined as the set of all elements $a \in R$ satisfying the following equivalent conditions (cf. e.g. Cohn [77], p. 340f. for a proof):

- (a) a annihilates every simple left R -module M , $aM = 0$,
- (b) a belongs to every maximal left ideal of R ,
- (c) $1 - xa$ has a left inverse for every $x \in R$,
- (d) $1 - xay$ has an inverse for every $x, y \in R$.

$(a^0)-(d^0)$: the left-right analogues of (a)–(d). ■

By (b), J is characterized as the intersection of all maximal left ideals and by (b^0) as the intersection of all maximal right ideals of R , hence it is a two-sided ideal of R .

We also recall Nakayama's lemma, in the following form:

LEMMA 5.2. *Let M be any finitely generated R -module and write $J = \mathbf{J}(R)$; if $M \neq 0$, then $M/JM \neq 0$.*

Since $M \neq 0$, it has a maximal proper submodule M' and M/M' is simple. By (a) above, $J \cdot (M/M') = 0$, so $JM \subseteq M' \subset M$ and $M/JM \neq 0$. ■

If we write $\bar{M} = M/JM$, the assertion of the lemma is that $\bar{M} = 0$ implies $M = 0$.

Let $f: R \rightarrow S$ be a homomorphism of rings; this induces a monoid homomorphism $\mathcal{S}(f): \mathcal{S}(R) \rightarrow \mathcal{S}(S)$ which preserves the distinguished element; explicitly we have $\mathcal{S}(f)[P] = [S \otimes_R P]$. Of course, this homomorphism need be neither injective nor surjective, even when f is, but we have

THEOREM 5.3. *Let R be any ring and $J = \mathbf{J}(R)$; then the natural homomorphism $\mathcal{S}(R) \rightarrow \mathcal{S}(R/J)$ is an embedding. Thus if R/J is projective free, then so is R .*

Proof. Write $\bar{R} = R/J$ and for each left R -module M write $\bar{M} = M/JM$ for the corresponding $\bar{R}$ -module. Given any finitely generated projective $\bar{R}$ -modules $\bar{P}, \bar{Q}$ such that $\bar{P} \cong \bar{Q}$, we have the following diagram, where f is an isomorphism:

$$\begin{array}{ccccc} P & \longrightarrow & \bar{P} & \longrightarrow & 0 \\ \downarrow g & & \downarrow f & & \\ Q & \longrightarrow & \bar{Q} & \longrightarrow & 0 \end{array}$$

Since P is projective, there is a map $g: P \rightarrow Q$ to make the diagram commute. Now $\text{coker } g = 0$, hence $\text{coker } g = 0$ by Nakayama's lemma, i.e. g is surjective. Therefore Q splits P , i.e. $P \cong Q \oplus \ker g$ and again from the diagram $\ker g = 0$, hence $\ker g = 0$ and so g is an isomorphism, as claimed. To prove the final remark, when R/J is projective free, then $\mathcal{S}(R/J) \cong \mathbf{N}$ and the natural homomorphism preserves the generator of $\mathbf{N}$, whence $\mathcal{S}(R) \cong \mathbf{N}$. ■

We remark that the result also follows by considering the idempotent matrices E, F , say, corresponding to P, Q (cf. e.g. Cohn [77], 10.5, Prop. 4, p. 350; the proof there is stated for idempotent elements, but still works for idempotent matrices, not necessarily of the same order).

We shall be particularly concerned with rings R such that R/J is a field. Such a ring is said to be *local*; equivalent characterizations are given in

PROPOSITION 5.4. *For any ring R the following conditions are equivalent:*

- (a) R is a local ring, i.e. $R/J(R)$ is a (skew) field,
- (b) the non-units in R form an ideal,
- (c) for any $a \in R$, either a or $1 - a$ has a one-sided inverse (not necessarily on the same side).

Proof. (a) $\Rightarrow$ (b). If R/J is a field, then J is the unique maximal ideal and hence consists of non-units. Given $u \notin J$, we have $uv \equiv vu \equiv 1 \pmod{J}$ for some $v \in R$, hence $uv = 1 + n$ ($n \in J$) is a unit, therefore so is u . Thus J is the set of all non-units.

(b) $\Rightarrow$ (c) is clear. To prove (c) $\Rightarrow$ (a), we first note that an idempotent $e \neq 1$ cannot have a one-sided inverse; for if $eu = 1$, then $e = e^2u = eu = 1$. Thus, when (c) holds, R has no idempotents $\neq 0, 1$. Next, let $ab = 1$, then ba is a non-zero idempotent, hence $ba = 1$, so all one-sided inverses in R are two-sided.

Now let $u \in R$; if u has no inverse, then neither does ux , for any $x \in R$, hence $1 - ux$ always has an inverse, so $u \in J$. Thus any non-unit is in J and (a) follows. ■

Here (a) or (b) is the usual form of the definition of a local ring, while (c) is the easiest to verify. Since every field is projective free we obtain, from Th. 5.3,

COROLLARY 5.5. *Every local ring is projective free.* ■

Sometimes a more general notion of local ring is needed. Let us call a ring R a *matrix local ring* if $R/J(R)$ is simple Artinian. By Wedderburn's theorem this means that $R/J \cong K_n$, where K is a field and $n \geq 1$; n is sometimes called the *capacity* of R . When $n = 1$, we are back in the case of a local ring. By contrast this is sometimes called a *scalar local ring*, but we generally omit the qualifier, so 'local ring' will mean as usual 'scalar local ring'.

Exercises 0.5

1. Verify the equivalence of the two definitions of $\mathcal{S}(R)$, in terms of projective modules and idempotent matrices.

2. Let R be a ring and $J = J(R)$. By considering the kernel of the homomorphism $R_n \rightarrow (R/J)_n$ induced by the natural homomorphism $R \rightarrow R/J$, show that $J(R_n) \cong J_n$.
3. Show that if R is a matrix local ring, then so is R_n .
4. For any ring R , show that $R/J(R)$ is weakly finite if and only if R is weakly finite. Deduce that any matrix local ring is weakly finite.
5. Show that a weakly semihereditary local ring is an integral domain.
6. Let K be a field and R a subring such that for any $x \in K$, either x or x^{-1} lies in R . If a is a non-unit in R , show that $a(1-a)^{-1} \in R$; deduce that R is a local ring.
7. In a local ring show that the additive order of 1 is 0 or a prime power.
8. Let R be a local ring with residue class field $K = R/J$. If M is a finitely generated left R -module such that $K \otimes M = 0$, show that $M = 0$.
9. Show that any Artinian matrix local ring is a full matrix ring over a scalar local ring. (Hint: Recall that idempotents can be lifted from R/J to R in this case, cf. e.g. Cohn [77], p. 350.)
10. Let R be the ring of rational quaternions with denominator prime to p , an odd prime. Show that pR is the Jacobson radical of R and R/pR is the ring of quaternions over $\mathbb{F}_p$; deduce that R is a matrix local ring which is not a matrix ring over a scalar local ring.
11. Show that the following are equivalent, for any ring R :
 - (a) R is local and any finitely generated left ideal is principal,
 - (b) the principal left ideals are totally ordered by inclusion,
 - (c) all left ideals are totally ordered by inclusion.
12. (Beck [72]) Let P be a finitely generated projective R -module. If P/JP is free over R/J , show that P is free over R . (This holds even if P is not finitely generated, cf. Beck [72].)

13*. (Kaplansky [58]) Let P be a projective module over a local ring. Show that any element of P can be embedded in a free direct summand of P ; deduce that every projective module over a local ring is free.

0.6 The matrix of definition of a module

Let R be any ring and M an R -module; then we have a presentation

$$G \xrightarrow{\alpha} F \longrightarrow M \longrightarrow 0,$$

where F, G are free. If F may be taken of finite rank, M is *finitely generated*; if G may be taken of finite rank, M is *finitely related*, and if both F, G can be taken of finite rank, M is *finitely presented*. Thus a finitely presented left R -module M has a presentation

$$R^m \xrightarrow{\alpha} R^n \longrightarrow M \longrightarrow 0. \quad (1)$$

Here M is determined up to isomorphism by the $m \times n$ matrix A describing α , and every $m \times n$ matrix defines a finitely presented left R -module M in this way, as $M = \text{coker } \alpha$. It is clear that α is injective if and only if A is left regular. In that case M has a 'finite free' resolution of length 1:

$$0 \longrightarrow R^m \xrightarrow{\alpha} R^n \longrightarrow M \longrightarrow 0, \quad (2)$$

and we define the *characteristic* of M by $\chi(M) = n - m$. In a ring with IBN this is well-defined and independent of the choice of presentation, by Schanuel's lemma (cf. Th. 6.2 below).

It should be observed that for general rings, modules with a finite free resolution of length at most 1 are very special; however, for the rings considered in later chapters they include all finitely presented modules, which is why we treat them in more detail. In particular, we can show that for such modules, short exact sequences correspond to matrix equations.

PROPOSITION 6.1. *Let R be a ring with IBN. Any short exact sequence*

$$0 \rightarrow M' \rightarrow M \rightarrow M'' \rightarrow 0 \quad (3)$$

of modules, each with a finite free resolution of length at most 1 corresponds to a matrix equation

$$C = AB \quad (4)$$

of left regular matrices A, B, C presenting M', M'', M , respectively, and

$$\chi(M) = \chi(M') + \chi(M''). \quad (5)$$

Conversely, any equation (4) between left regular matrices corresponds to a short exact sequence (3).

Proof. Given (3), by hypothesis there exists a free module F mapping onto M , with free kernel H , both of finite rank. We also have a surjection $F \rightarrow M''$, and if the kernel is denoted by K , then $K \cong H$ and we have the commutative

diagram with exact rows and columns:

$$\begin{array}{ccccccc}
 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \\
 0 & \longrightarrow & H & \xrightarrow{1} & H & \longrightarrow & 0 \\
 & & \downarrow A & & \downarrow C & & \\
 0 & \longrightarrow & K & \xrightarrow{B} & F & \longrightarrow & M'' \longrightarrow 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 0 & \longrightarrow & M' & \longrightarrow & M & \longrightarrow & M'' \longrightarrow 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 & & 0 & & 0 & & 0
 \end{array}$$

Moreover, since M'' has a finite free resolution of length 1, there are free modules $F_1 \cong K_1$ of finite rank such that $M'' \cong F_1/K_1$. By Schanuel's lemma we have $K \oplus F_1 \cong K_1 \oplus F$, therefore on replacing F, K by $F \oplus F_1, K \oplus F_1$, respectively, we can ensure that K is also free. If the matrices defining M', M'', M are A, B, C , respectively, then by commutativity, $C = AB$. Now $\chi(M) = \text{rk } F - \text{rk } H$, $\chi(M') = \text{rk } K - \text{rk } H$, $\chi(M'') = \text{rk } F - \text{rk } K$, and (5) follows by addition. Conversely, given left regular matrices A, B, C satisfying (4), we obtain the first two columns of the above diagram, hence the third follows by the dual of the 3×3 lemma. ■

Let us call two matrices over R *left similar* if the left modules they define are isomorphic; *right similar* matrices are defined correspondingly, and two matrices are called *similar* if they are left and right similar.

The precise relationship between similar matrices was found by Fitting [36]. This relation can be simplified by restricting attention to matrices that are left regular, corresponding to the case where α in (1) is injective. We shall give an explicit description of similarity in this case; in essence this is just a formulation of Schanuel's lemma.

Two maps between R -modules, $\alpha: Q \rightarrow P$, $\alpha': Q' \rightarrow P'$ are said to be *associated* if there is a commutative square

$$\begin{array}{ccc}
 Q & \xrightarrow{\alpha} & P \\
 \downarrow & & \downarrow \\
 Q' & \xrightarrow{\alpha'} & P'
 \end{array}$$

where the vertical maps are isomorphisms. If there are two R -modules S, T such that $\alpha \oplus 1_S$ is associated to $1_T \oplus \alpha'$, then α and α' are said to be *stably associated*. The next result and its corollaries describe similarity of matrices in terms of stable association.

THEOREM 6.2. *Let R be a ring and let $\alpha: Q \rightarrow P$ and $\alpha': Q' \rightarrow P'$ be two homomorphisms of left R -modules. Then the following conditions are equivalent:*

(a) *there is an isomorphism $\mu: Q \oplus P' \rightarrow P \oplus Q'$ of the form*

$$\mu = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \text{ with inverse } \mu^{-1} = \begin{pmatrix} \delta' & \beta' \\ \gamma' & \alpha' \end{pmatrix},$$

(b) *α is stably associated to α' .*

Further, these conditions imply

(c) *$\text{coker } \alpha \cong \text{coker } \alpha'$,*

and if P, P' are projective modules and α, α' are injective, then the converse holds, so (a), (b), (c) are then equivalent.

Proof. (a) $\Rightarrow$ (b). If we take $S = P', T = P$, we obtain the commutative square

$$\begin{array}{ccc}
 Q \oplus P' & \xrightarrow{\alpha \oplus 1_{P'}} & P \oplus P' \\
 \downarrow \mu & & \downarrow \nu \\
 P \oplus Q' & \xrightarrow{1_P \oplus \alpha'} & P \oplus P'
 \end{array}$$

with the vertical isomorphisms

$$\mu = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad \text{and} \quad \nu = \begin{pmatrix} 1 & 0 \\ \gamma & 1 \end{pmatrix} \begin{pmatrix} 1 & -\beta' \\ 0 & 1 \end{pmatrix}.$$

(b) $\Rightarrow$ (a). If α is stably associated to α' , we have a commutative square, which is expressed by an equation

$$\begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} x & y \\ z & t \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \alpha' \end{pmatrix},$$

where

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix}^{-1} = \begin{pmatrix} s' & q' \\ r' & p' \end{pmatrix}, \quad \begin{pmatrix} x & y \\ z & t \end{pmatrix}^{-1} = \begin{pmatrix} t' & y' \\ z' & x' \end{pmatrix}.$$

Now (a) follows with

$$\mu = \begin{pmatrix} \alpha & y \\ -r' & p't \end{pmatrix}, \quad \mu^{-1} = \begin{pmatrix} p't' & -q \\ z' & \alpha' \end{pmatrix},$$

as is easily checked.

(b) $\Rightarrow$ (c) is clear. Now let P, P' be projective, α, α' injective and assume (c). Then there exist maps $\gamma: P' \rightarrow P$ and $\beta': P \rightarrow P'$ making the following diagram commutative, and γ induces $-\gamma': Q' \rightarrow Q$, while β' induces $-\beta: Q \rightarrow Q'$.

$$\begin{array}{ccccccc} 0 & \longrightarrow & Q & \xrightarrow{\alpha} & P & \xrightarrow{\phi} & \text{coker } \alpha \longrightarrow 0 \\ & & \downarrow -\beta & \uparrow -\gamma' & \downarrow \beta' & \uparrow \gamma & \downarrow \cong \\ 0 & \longrightarrow & Q' & \xrightarrow{\alpha'} & P' & \xrightarrow{\phi'} & \text{coker } \alpha' \longrightarrow 0 \end{array}$$

Further, $(1_P - \beta'\gamma)\phi = 0$, whence $1 - \beta'\gamma = \delta'\alpha$ for some $\delta': P \rightarrow Q$, because P is projective. Likewise $(1_{P'} - \gamma\beta')\phi' = 0$, whence $1 - \gamma\beta' = \delta\alpha'$ for some $\delta: P' \rightarrow Q'$. Now it is easily verified that $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}: Q \oplus P' \rightarrow P \oplus Q'$ has inverse $\begin{pmatrix} \delta' & \beta' \\ \gamma' & \alpha' \end{pmatrix}: P \oplus Q' \rightarrow Q \oplus P'$. ■

The implication (c) $\Rightarrow$ (a) (under the given conditions) is just the assertion of Schanuel's lemma. The proof of the equivalence (a) $\Leftrightarrow$ (b) shows that the definition of stable association can be made a little more precise:

COROLLARY 6.3. *If $\alpha: Q \rightarrow P$ is stably associated to $\alpha': Q' \rightarrow P'$, then $\alpha \oplus 1_P$ is associated to $1_P \oplus \alpha'$. Hence two matrices $A \in {}^r R^m$ and $A' \in {}^s R^n$ are stably associated, qua maps, if and only if $A \oplus I_n$ is associated to $I_m \oplus A'$. ■*

In terms of matrices we obtain the following criteria by taking P, P', Q, Q' to be free:

COROLLARY 6.4. *Let $A \in {}^r R^m$, $A' \in {}^s R^n$ be any two matrices. Then of the following, (a) and (b) are equivalent and imply (c):*

- (a) there exists $M = \begin{pmatrix} A & * \\ * & * \end{pmatrix} \in {}^{r+n}R^{s+m}$ with inverse of the form $\begin{pmatrix} * & * \\ * & A' \end{pmatrix}$,
 (b) A and A' are stably associated,
 (c) A and A' are left similar.

If A, A' are left regular, all three conditions are equivalent. Moreover, two regular matrices are left similar if and only if they are right similar.

Here the last part follows by the evident symmetry of (a) or (b). ■

Thus, for regular matrices, 'similar' and 'stably associated' mean the same. We also note that an invertible (square) matrix is stably associated to the unique 0×0 matrix 1_0 .

From (a) we see that if two matrices A, A' over R are stably associated, then their images under any ring homomorphism are again stably associated; in particular, if A maps to a unit under some homomorphism, then so does A' .

We note that if R has IBN, then $A \in {}^rR^m$ and $A' \in {}^sR^n$ can be stably associated only if $r+n = s+m$, i.e. $n-s = m-r$. Let us define the *index* of an $r \times m$ matrix A as $i(A) = m-r$; then we can say that over a ring with IBN, stably associated matrices have the same index. This is also the characteristic of the corresponding left R -module M , assuming A to be left regular. We observe that a regular matrix A of index t defines a left module of characteristic t and a right module of characteristic $-t$; here we shall mainly be concerned with left modules.

Over a weakly finite ring the notion of similarity of matrices can be simplified as follows. Consider a relation

$$AB' = BA' \quad (6)$$

between matrices. This can also be written

$$(A \ B) \begin{pmatrix} -B' \\ A' \end{pmatrix} = 0.$$

We shall call A, B *right comaximal* if the matrix $(A \ B)$ has a right inverse, and A', B' *left comaximal* if $(A' \ B')^T$ has a left inverse. Now (6) is called a *comaximal relation* if A, B are right comaximal and A', B' left comaximal. We shall find that in a weakly finite ring stable association can be described in terms of comaximal relations.

PROPOSITION 6.5. *Let R be any ring and let $A \in {}^rR^m$, $A' \in {}^sR^n$. Then the following relations are equivalent:*

- (a) A, A' satisfy a comaximal relation (6),

(b) *there is an $(r+n) \times (s+m)$ matrix $\begin{pmatrix} A & B \\ * & * \end{pmatrix}$ with a right inverse of the form $\begin{pmatrix} * & -B' \\ * & A' \end{pmatrix}$.*

In particular, (a) and (b) hold whenever

(c) *A and A' are stably associated,*

and in a weakly finite ring (a)–(c) are equivalent for two matrices of the same index.

Proof. If A, A' satisfy a comaximal relation (6), say

$$AD' - BC' = I, \quad DA' - CB' = I, \quad (7)$$

then, on writing

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} D' & -B' \\ -C' & A' \end{pmatrix} \quad (8)$$

we have $MN = \begin{pmatrix} I & 0 \\ P & I \end{pmatrix}$; hence M has the right inverse

$$\begin{pmatrix} D' & -B' \\ -C' & A' \end{pmatrix} \begin{pmatrix} I & 0 \\ -P & I \end{pmatrix} = \begin{pmatrix} * & -B' \\ * & A' \end{pmatrix},$$

and (b) follows. Conversely, if N in (8) is a right inverse of M , then (6) and (7) hold, hence (6) is then a comaximal relation. This shows that (a) $\Leftrightarrow$ (b).

Now (c) $\Rightarrow$ (b) by Cor. 6.4, and (b) $\Rightarrow$ (c) under the given conditions, because when $m-r = n-s$, then $r+n = s+m$, and for a square matrix over a weakly finite ring, any right inverse is an inverse. Now another application of Cor. 6.4 shows that (c) holds. ■

For later use we note the explicit form of the relation of stable association between A and A' :

$$\begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} D' & -B' \\ -C' & A' \end{pmatrix} = \begin{pmatrix} I+BC' & -B \\ -C' & I \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & A' \end{pmatrix}. \quad (9)$$

Let us also restate the criterion for stable association derived in Prop. 6.5.

PROPOSITION 6.6. *In a weakly finite ring R , two matrices A and A' are stably associated if and only if they have the same index and satisfy a comaximal relation $AB' = BA'$. ■*

Here the condition on the index cannot be omitted, as the equation $(1, x)(1, y)^T = (1, x)(1, y)^T$ shows.

Exercises 0.6

1. Show that a matrix is stably associated to I if and only if it is a unit; it is stably associated to 0 only if it is a zero-divisor.
2. Show that the presenting matrix of a module M is invertible if and only if $M = 0$.
3. Let R be a ring and $A \in {}^m R^n$, $B \in {}^n R^m$. Show that $I + AB$ is stably associated to $I + BA$. (Hint: Show that each is stably associated to $\begin{pmatrix} I & A \\ -B & I \end{pmatrix}$.) Deduce that $I + AB$ is a unit if and only if $I + BA$ is; prove this directly by evaluating $I - B(I + AB)^{-1}A$.
4. Under what circumstances is AB stably associated to BA ?
5. Let R be a ring with UGN. If A, A' satisfy a comaximal relation (6), show that $i(A) \geq i(A')$; deduce that if A, A' satisfy a comaximal relation and A', A likewise, then A and A' are stably associated.
6. Give examples of pairs of matrices (over a weakly finite ring, say) which satisfy a comaximal relation but are not stably associated.
- 7*. Let R be a non-zero ring and $S = \text{End}_R(R^{(\mathbb{N})})$. Show that $S_2 \cong S$ and that any two stably associated 1×1 matrices are associated.

0.7 Eigenrings and centralizers

Let R be a ring, P a projective left R -module and N a submodule of P . We define the *idealizer* of N in P over R as the set

$$I(N) = \{\beta \in \text{End}_R(P) \mid N\beta \subseteq N\}.$$

Clearly $I(N)$ is a subring of $\text{End}_R(P)$ and if we put $\alpha = \text{Hom}_R(P, N)$, then α is a left ideal in $\text{End}_R(P)$ and a two-sided ideal in $I(N)$. The quotient ring $E(N) = I(N)/\alpha$ is called the *eigenring* of N in P over R . Writing $M = P/N$, we have a natural ring homomorphism $I(N) \rightarrow \text{End}_R(M)$. Further, for any $\phi \in \text{End}_R(M)$ we have a commutative diagram:

$$\begin{array}{ccccccc} 0 & \longrightarrow & N & \longrightarrow & P & \longrightarrow & M \longrightarrow 0 \\ & & \downarrow \beta' & & \downarrow \beta & & \downarrow \phi \\ 0 & \longrightarrow & N & \longrightarrow & P & \longrightarrow & M \longrightarrow 0 \end{array}$$

Since P is projective, ϕ can be lifted to an endomorphism β of P such that $N\beta \subseteq N$. Thus the map $I(N) \rightarrow \text{End}_R(M)$ is surjective; clearly the kernel is just α , so we have proved

PROPOSITION 7.1. If R is any ring, P a projective left R -module and N a submodule of P , then there is a natural isomorphism

$$E(N) \cong \text{End}_R(P/N). \blacksquare$$

In particular, if α is a left ideal of R , then taking $P = R$ we have $I(\alpha) = \{x \in R \mid \alpha x \subseteq \alpha\}$ and $E(\alpha) = I(\alpha)/\alpha \cong \text{End}_R(R/\alpha)$.

Now let $A \in {}^m R^n$ and let $\alpha: R^m \rightarrow R^n$ be the corresponding map. Taking $P = R^n$, $N = \text{im } \alpha$, we have $\text{End}_R(R^n) \cong R_n$, hence $I(N) = \{\beta \in R_n \mid N\beta \subseteq N\}$ and $\alpha = \text{Hom}_R(R^n, N) \cong {}^n N$, as left R_n -module. We define the *left idealizer* of A over R as the corresponding set of matrices

$$I(A) = \{B \in R_n \mid AB = B'A \quad \text{for some } B' \in R_m\},$$

and the *left eigenring* of A as the quotient ring $E(A) = I(A)/({}^n R^m)A$. By Prop. 7.1, $E(A) \cong \text{End}_R(M)$, where M is the left R -module defined by A . The right eigenring of A is defined similarly, and it is clear that for a regular matrix A the left and right eigenrings are isomorphic, the isomorphism being induced by the mapping

$$B \mapsto B', \quad \text{where } AB = B'A.$$

In the particular case where $m = n = 1$, the matrix becomes an element a of R and we have $E(a) = I(a)/Ra \cong \text{End}_R(R/Ra)$.

Given any matrices A, B over R , if M, N are the left R -modules defined by them, then each R -homomorphism $f: M \rightarrow N$ is completely specified by a matrix P over R such that

$$AP = P'B \tag{1}$$

for some matrix P' . If $I(A, B)$ denotes the set of all such P and b is the left R -module spanned by the rows of B , then as before,

$$I(A, B)/b \cong \text{Hom}_R(M, N) \tag{2}$$

is an isomorphism of $(E(A), E(B))$ -bimodules.

For later use we record the effect of a change of base field on $\text{Hom}_R(M, N)$:

PROPOSITION 7.2. Let R be a k -algebra, where k is a commutative field. Given a field extension E/k , write $R^E = R \otimes_k E$ and for any R -module M denote the extension $M \otimes_k E$ by M^E . If M, N are R -modules such that $\text{Hom}_R(M, N)$ is finite-dimensional over k , then

$$\text{Hom}_R(M, N) \otimes_k E \cong \text{Hom}_{R^E}(M^E, N^E). \tag{3}$$

Proof. There is a natural map from the left- to the right-hand side in (3),

which is clearly injective, so it will be enough to show that both sides have the same dimension.

Let $(e_i), (f_\lambda)$ be bases for M, N as k -spaces (possibly infinite-dimensional), then

$$e_i x = \sum \rho_{ij}(x) e_j, \quad f_\lambda x = \sum \sigma_{\lambda\mu}(x) f_\mu \quad (x \in R),$$

and $\text{Hom}_R(M, N)$ is the space of all solutions α over k of the system

$$\sum_j \rho_{ij}(x) \alpha_{jm} = \sum_\lambda \alpha_{i\lambda} \sigma_{\lambda m}(x). \quad (4)$$

Let $V = \text{Hom}_k(M, N)$ as k -space and C the subspace of solutions $(\alpha_{i\lambda})$ of (4), then $V = C \oplus D$ for some k -space D , and by hypothesis C is finite-dimensional, say $\dim_k C = n$. Let $p_1, \dots, p_n$ be linearly independent functionals on V such that $\cap \ker p_v = D$, then (4) together with

$$p_v(\alpha) = 0 \quad (v = 1, \dots, n)$$

has only the trivial solution over k . Hence it has only the trivial solution over E . It follows that the solution of (4) is again n -dimensional over E , and so (3) holds. ■

The eigenring of an element is closely related to its centralizer, and to some extent both may be treated by the same method, by the device of adjoining an indeterminate. The basic result is:

THEOREM 7.3. *Let R be a ring and $S = R[t]$ the ring obtained by adjoining a central indeterminate t to R . Given $a, b \in R$, write $C = C(a, b) = \{x \in R \mid ax = xb\}$. Then there is a natural isomorphism of abelian groups:*

$$C(a, b) \cong \text{Hom}_S(S/S(t-a), S/S(t-b)).$$

Proof. Let R_a denote R viewed as left S -module with t acting by right multiplication by a . By mapping $t \mapsto a$ we define a left S -module homomorphism $S \rightarrow R_a$ with kernel $S(t-a)$, hence $R_a \cong S/S(t-a)$, and so

$$\text{Hom}_S(S/S(t-a), S/S(t-b)) \cong \text{Hom}_S(R_a, R_b).$$

For any $f: R_a \rightarrow R_b$ we have $a((1)f) = (a)f = (1t)f = t((1)f) = ((1)f)b$, therefore the rule $f \mapsto (1)f$ defines a homomorphism from $\text{Hom}_S(R_a, R_b)$ to $C(a, b)$. Conversely, for any $x \in C(a, b)$, right multiplication by x defines a left S -linear map $R_a \rightarrow R_b$; so we have a homomorphism $C(a, b) \rightarrow \text{Hom}_S(R_a, R_b)$, clearly inverse to the previous map. ■

By putting $b = a$ we can express the centralizer of a as an eigenring:

COROLLARY 7.4. *The centralizer of an element $a \in R$ is isomorphic to the eigenring of $t - a$ in the polynomial ring $R[t]$. ■*

The following result is well known in the special case of matrix rings over a field, where it is used to obtain the canonical form of a matrix.

PROPOSITION 7.5. *Let R be any ring and t a central indeterminate. Then two elements a, b of R are conjugate under $U(R)$ if and only if $t - a$ and $t - b$ satisfy a comaximal relation*

$$f \cdot (t - a) = (t - b) \cdot g \quad (5)$$

in $R[t]$. Moreover, in any such comaximal relation (5), f and g may be taken to lie in $U(R)$.

Proof. If a, b are conjugate, say $ua = bu$, where $u \in U(R)$, then clearly $u(t - a) = (t - b)u$ is a comaximal relation. Conversely, assume a comaximal relation (5). By subtracting an expression $(t - b)h(t - a)$ from both sides, we get the equation

$$u(t - a) = (t - b)v, \quad (6)$$

where $u = f - (t - b)h$, $v = g - h(t - a)$. Here we may choose h so that u has degree 0 in t , i.e. $u \in R$. Then on comparing degrees in (6) we find that $v \in R$, while a comparison of highest terms shows that $v = u$ and so

$$ua = bu. \quad (7)$$

Further, since $u \equiv f \pmod{(t - b)R[t]}$, u and $t - b$ are still right comaximal, say

$$up + (t - b)q = 1, \quad \text{where } p, q \in R[t]. \quad (8)$$

Replacing p by $p - (t - a)k$ for suitable $k \in R[t]$ and using (6), we can reduce (8) to the case where p has degree 0. Then $q = 0$, by comparing degrees, and (8) shows p to be a right inverse to u . By the symmetry of (7), u also has a left inverse and so is a unit. Now (7) shows a and b to be conjugate, as claimed. ■

Exercises 0.7

1. In any ring R , if $ab' = ba'$, show that $a'b$ lies in the idealizer of $Rb'b$ and that of $a'aR$.
2. Let R be a ring and t a central indeterminate. Given $a, b \in R^\times$, show that the

elements $a^n b$ ($n = 0, 1, \dots$) are right linearly dependent over R if and only if $t - a$ and b are right commensurable in $R[t]$.

3. Let A be a matrix with eigenring E . Show that $A \oplus \dots \oplus A$ (r terms A) has eigenring E_r .

4. Let R be a ring and t a central indeterminate. If $t - a$ and $h \in R[t]$ satisfy a comaximal relation $f \cdot (t - a) = h \cdot g$, show that g can be taken to lie in R , but not in general f . (Hint: Use a nilpotent element of R to construct f as an invertible element of degree 2.)

5. Show that a unit has zero eigenring and conversely, an element with zero eigenring over an integral domain is a unit.

6*. (Robson [72]) (a) let $a = m_1 \cap \dots \cap m_k$, where the m_i are maximal left ideals (such an a is called *semimaximal*). If $B = \{b \in R \mid ab \subseteq m_1\}$, and A is the idealizer of a , show that B/m_1 is a simple left A -module.

(b) If a and A are as before, show that any simple left R -module is either simple as left A -module or a homomorphic image of R/a .

(c) With the notation as before, let M be a simple left R -module. Then M is simple as left A -module, unless for some i , $M \cong R/m_i$ and $aR \not\subseteq m_i$. In that case M has a unique composition series

$$R \supset A + m_i \supset m_i.$$

0.8 Groups and rings of fractions

Let S be a monoid and T a subset of S . A homomorphism $f: S \rightarrow S'$ to another monoid S' is said to be *T -inverting* if T is mapped by f into the subgroup of units of S' . When f is an embedding, the submonoid of S' generated by Sf and $(Tf)^{-1}$ is called a *monoid of fractions* of S , or a *group of fractions* when $T = S$.

By the *universal T -inverting monoid* on S we understand a monoid S_T with a T -inverting homomorphism $\lambda: S \rightarrow S_T$ which is universal for all such homomorphisms, i.e. any T -inverting homomorphism $f: S \rightarrow S'$ can be factored uniquely by λ . Like all universal constructions, S_T is unique up to isomorphism; it may be constructed by writing down a presentation of S as a monoid in terms of generators and defining relations, and for each $t \in T$ adjoining an element t' with relations $tt' = t't = 1$.

$$\begin{array}{ccc} S & \xrightarrow{\lambda} & S_T \\ & \searrow & \downarrow \text{dashed} \\ & & S' \end{array}$$

In particular, when $T = S$, S_S is a group, written $\mathcal{G}(S)$ and called the *universal group* of S . A presentation for $\mathcal{G}(S)$ is obtained by taking a presentation for S as a monoid and treating it as a group presentation.

By the uniqueness of S_T we obtain the following result on the extension of homomorphisms:

PROPOSITION 8.1. *Let S, S' be monoids with subsets T, T' . Then any homomorphism $f: S \rightarrow S'$ which maps T into T' can be extended in just one way to a homomorphism $S_T \rightarrow S'_{T'}$.*

For the composition $S \rightarrow S' \rightarrow S'_{T'}$ is T -inverting and hence can be factored by $\lambda: S \rightarrow S_T$. It is unique because it is prescribed on a generating set of S_T , viz. $S \cup T^{-1}$. ■

We shall mainly be concerned with the question: When is $\lambda: S \rightarrow S_T$ injective? A necessary condition for λ to be injective is that S should admit cancellation by T :

$$st = s't \text{ or } ts = ts' \text{ implies } s = s', \text{ for all } s, s' \in S, t \in T. \quad (1)$$

In the commutative case this is also sufficient, as is well known, but in general it is only the first of an infinite sequence of conditions which are necessary and sufficient for S to be embeddable in S_T . These are the *Malcev conditions* (cf. e.g. Cohn [81], p. 268). They are usually stated in the case $T = S$; a monoid S satisfying (1) for $T = S$ is called a *cancellation monoid*.

We shall not discuss the Malcev conditions in more detail, but examine some special cases where the embedding is more easily proved directly. The most important of these is the case where, on identifying S with its image in S_T , every element of S_T can be written as st^{-1} ($s \in S, t \in T$). In particular, $t^{-1}s$ must have this form, say $t^{-1}s = s_1t_1^{-1}$ and on multiplying up we find $st_1 = ts_1$. Let us define a *right Ore set* in a monoid S as a submonoid T satisfying (1) and

$$sT \cap tS \neq \emptyset \quad \text{for all } s \in S, t \in T. \quad (2)$$

Then the basic result may be stated as

THEOREM 8.2. *Let S be a monoid and T a right Ore set in S . Then the natural map $\lambda: S \rightarrow S_T$ into the universal T -inverting monoid is an embedding and every element of S_T has the form st^{-1} ($s \in S, t \in T$). Conversely, when $S_T = ST^{-1}$ and λ is an embedding, then T is a right Ore set.*

We shall sketch the proof (cf. e.g. Cohn [77], Ch. 12). Define an equivalence on $S \times T$ by the rule: $(s_1, t_1) \sim (s_2, t_2)$ whenever there exist $u_1, u_2 \in S$ such that $t_1 u_1 = t_2 u_2 \in T$, $s_1 u_1 = s_2 u_2$. This is easily seen to be an equivalence. We define a multiplication on the equivalence classes as follows: Given $(s_1, t_1), (s_2, t_2)$, we can find $u_1 \in T, u_2 \in S$ such that $t_1 u_1 = s_2 u_2$. Now put $(s_1, t_1) \cdot (s_2, t_2) = (s_1 u_2, t_2 u_1)$; since $t_2 u_1 \in T$, this defines an element of $S \times T$ and it is easily checked that its class depends only on the equivalence classes of the factors, not on the factors themselves. Moreover, the multiplication is associative, and we have a homomorphism which takes s to the class of $(s, 1)$. It is injective, since $(s_1, 1) \sim (s_2, 1)$ means that $s_1 u_1 = s_2 u_1$ for some $u_1 = u_2 \in T$, whence $s_1 = s_2$ by (1). Moreover, $(t, 1)$ has the inverse $(1, t)$ for any $t \in T$, so that we can write $(s, t) = (s, 1)(1, t) = st^{-1}$, on identifying S with its image. Conversely, when S is embedded in S_T and $S_T = ST^{-1}$, then the argument leading to (2) shows that T is a right Ore set. ■

We remark that any finite set of elements of S_T may be brought to a common denominator, which is a right multiple of the given denominators. Taking first the case of two elements, let $c_i = a_i b_i^{-1}$, $a_i \in S$, $b_i \in T$ ($i = 1, 2$), then $b_1 u_1 = b_2 u_2$ for suitable $u_1 \in S$, $u_2 \in T$, hence $b = b_1 u_2 \in T$ and $c_1 = a_1 u_2 (b_1 u_2)^{-1}$, $c_2 = a_2 u_1 (b_2 u_1)^{-1}$. When we have $n \geq 2$ elements, $c_1, \dots, c_n$, we first bring $c_2, \dots, c_n$ to a common denominator b' and then bring c_1, b' to a common denominator b . Then b is the desired common denominator of $c_1, \dots, c_n$, and a right multiple of the original denominators.

This remark can be used to show that for right Ore sets the monoid of fractions is unique up to isomorphism:

COROLLARY 8.3. *Let S be a monoid and T a right Ore set. Then any monoid of fractions for T is isomorphic to S_T .*

Proof. Let S' be a monoid of fractions, with T -inverting homomorphism $f: S \rightarrow S'$. By hypothesis this is injective, and of the form $\lambda f'$, where $f': S_T \rightarrow S'$. Now let $c_1, c_2 \in S_T$ have the same image under f' ; we may write them as $c_i = a_i b^{-1}$ with a common denominator. By hypothesis, $c_1 f' = c_2 f'$, hence $a_1 f = c_1 f' \cdot b f = c_2 f' \cdot b f = a_2 f$ and, since f is injective, $a_1 = a_2$. Thus f' is injective, clearly it is surjective and so is an isomorphism. ■

To find Ore sets in S one looks for its 'large' elements. An element c of a monoid S is said to be *right large* if $cS \cap aS \neq \emptyset$ for all $a \in S$. The set L of all right large elements is always a submonoid. For clearly $1 \in L$ and if $a, b \in L$, take $c \in S$, then there exist $x, y \in S$ such that $ax = cy$ and there exist $u, v \in S$ such that $bu = xv$, hence $abu = axv = cyv$ and this shows that $ab \in L$. If S admits cancellation and $aS \cap bS \neq \emptyset$ implies $ab' = ba'$ with either a' or b' in

L , then L is a right Ore set. For let $a \in L$ and $b \in S$, then $aS \cap bS \neq \emptyset$, say $ab' = ba'$. If $a' \in L$, then $aS \cap bL \neq \emptyset$ as claimed; otherwise $b' \in L$ and then $ab' \in L$, hence $ba' \in L$. Now for any $c \in S$, there exist $x, y \in S$ with $ba'x = bcy$, hence $a'x = cy$ and this shows again that $a' \in L$. Thus, we have proved

PROPOSITION 8.4. *In any monoid the set L of right large elements is a submonoid. In a cancellation monoid S , if $aS \cap bS \neq \emptyset$ implies $ab' = ba'$, where either a' or b' is right large, then the set L of right large elements in S is a right Ore set and the natural map $\lambda: S \rightarrow S_L$ is an embedding. ■*

If in Prop. 8.4, $L = S$, we obtain a sufficient condition for S to have a group of fractions, i.e. to be embeddable in a group. Of course, this follows also directly from Th. 8.2.

COROLLARY 8.5. *Let S be a monoid and $\mathcal{G}(S)$ its universal group. Then the natural map $S \rightarrow \mathcal{G}(S)$ is an embedding and $\mathcal{G}(S) = SS^{-1}$ if and only if S is a cancellation monoid and every element of S is right large.*

The sufficiency of the conditions follows from Th. 8.2 and the necessity from the remarks preceding that result. ■

The definitions and constructions given so far can all be applied to a ring, since the latter is a monoid under multiplication. However, we must be careful not to invert 0 if we want to avoid the zero ring. Thus if R is any ring and T a subset of $R^\times$, it is clear what is meant by a T -inverting homomorphism, and the *universal T -inverting ring* R_T is a ring with a homomorphism $\lambda: R \rightarrow R_T$ which is universal for all T -inverting homomorphisms. This ring R_T is also called the *localization of R at T* , and when λ is injective, the *universal ring of fractions of R by T* .

If R is an integral domain, then $R^\times$ is a monoid and the Malcev conditions tell us (in principle) when $R^\times$ is embeddable in a group. However, even when this is so, R need not be embeddable in a field (cf. Ch. 7). But if R is an integral domain and $R^\times$ a right Ore set—so that we can embed $R^\times$ in its universal group $\mathcal{G}(R^\times)$ —it turns out that we can extend the addition on $R^\times$ to $K = \mathcal{G}(R^\times) \cup \{0\}$ in such a way as to obtain a field containing R as subring. More generally, we can form a ring of fractions with respect to any right Ore set.

Thus let R be a ring and T a right Ore set in R . By Th. 8.2 there is an embedding $R \rightarrow R_T$ of R into the universal T -inverting monoid. To define addition in R_T we recall that any two elements of R_T may be brought to a common denominator: a_1b^{-1} and a_2b^{-1} , say. Now their sum is defined by

the equation

$$a_1 b^{-1} + a_2 b^{-1} = (a_1 + a_2) b^{-1}. \quad (3)$$

It is easily checked that this is independent of the form in which the elements of R_T are written and that R_T forms a ring with respect to this addition, with R as subring. We summarize the result as

THEOREM 8.6. *Let R be any ring and T a right Ore set in R , i.e. T is a submonoid of R such that for any $r \in R$ and $t \in T$, $rT \cap tR \neq \emptyset$ and $rt = 0$ or $tr = 0$ implies $r = 0$. Then the universal T -inverting monoid R_T has a ring structure such that the natural mapping $\lambda: R \rightarrow R_T$ is a ring monomorphism and $R_T = RT^{-1}$. Conversely, if T is a submonoid of R such that $\lambda: R \rightarrow R_T$ is injective and $R_T = RT^{-1}$, then T is a right Ore set. Moreover, when T is a right Ore set, the ring of fractions is unique up to isomorphism.*

Here the last part is a consequence of Cor. 8.3. ■

The general case of localization, by a set which may not be right Ore, will be taken up in Ch. 7.

Clearly, a right Ore set in a non-zero ring R cannot contain 0 and so must be a subset of $R^\times$. The case when the whole of $R^\times$ is a right Ore set is of particular importance. It is easily seen that this is so precisely when R is an integral domain satisfying the following condition, known as the *right Ore condition*:

$$aR \cap bR \neq 0 \quad \text{for all } a, b \in R^\times. \quad (4)$$

An integral domain R satisfying (4) is called a *right Ore domain*. Left Ore domains are defined correspondingly and an *Ore domain* is a domain which is left and right Ore. Now Th. 8.6 yields

COROLLARY 8.7. *Any right Ore domain R can be embedded in a field and the least such field, unique up to isomorphism, is the universal $R^\times$ -inverting ring $R_{R^\times}$. Hence any (left or right) Ore domain is weakly finite. ■*

The field obtained here is also called the *field of right fractions* of R . By symmetry every left Ore domain can be embedded in a field of left fractions, and for a two-sided Ore domain the fields of left and right fractions coincide, by the uniqueness of the latter. We also note that every commutative integral domain is both a left and right Ore domain.

If R is a ring and T a right Ore set in R , then any T -inverting homomorphism $f: R \rightarrow S$ extends in a unique fashion to a homomorphism of R_T

into S , by the universal property of R_T . Sometimes we shall need this result for R -subrings of R_T ; the proof is quite similar to that of Th. 8.6, though it does not actually follow as a special case:

PROPOSITION 8.8. *Let $f: R \rightarrow S$ be an injective homomorphism. If T is a right Ore set in R such that Tf is regular and R' is an R -subring of R_T such that*

$$a \in R'b \quad (a \in R, b \in T) \quad \Rightarrow af \in S.bf,$$

then f extends to a unique homomorphism $f': R' \rightarrow S$ and f' is again injective.

We outline the proof. Given $r \in R'$, we can write $r = ab^{-1}$ ($a \in R, b \in T$), thus $a = rb$, and hence $af = s.bf$ for some $s \in S$. We define $rf' = s$ and note that if instead of a, b we had used au, bu , where $bu \in T$, then $(au)f = s.(bu)f$ with the same s , so any expression $r = au.(bu)^{-1}$ leads to the same value of rf' . Since any two representations of r can be brought to a common denominator, they lead to the same value of rf' and this shows f' to be well-defined. Now the homomorphism property follows as for Th. 8.6 and injectivity as in Cor. 8.3. ■

An integral domain R that is not a right Ore domain must contain two non-zero elements a, b that are right incommensurable: $aR \cap bR = 0$. It follows that the right ideal $aR + bR$ is a free right R -module of rank 2. Moreover, the elements $a^n b$ ($n = 0, 1, 2, \dots$) are right linearly independent; for if $\sum a^i b c_i = 0$, then by cancelling as many factors a as possible on the left we can write this equation as

$$bc_0 + abc_1 + \dots + a^n bc_n = 0 \quad (c_0 \neq 0),$$

hence $bc_0 \in aR \cap bR$, a contradiction. This proves

PROPOSITION 8.9. *An integral domain that is not a right Ore domain contains free right ideals of any finite or countable rank. ■*

Since a free right ideal of countable rank is not finitely generated, we obtain

COROLLARY 8.10. *Any right Noetherian integral domain is a right Ore domain. ■*

Examples of non-Ore domains are free associative algebras of rank at least two (soon to be formally defined, in 0.11).

Let R, A, B be any rings, $\alpha: R \rightarrow A, \beta: R \rightarrow B$ two homomorphisms and M an (A, B) -bimodule. Then an (α, β) -derivation from R to M is a map $\delta: R \rightarrow M$

which is additive and satisfies

$$(xy)^\delta = x^\alpha y^\delta + x^\delta y^\beta. \quad (5)$$

In particular, if $A = R$ and $\alpha = 1$, we speak of a (right) β -derivation. Putting $x = y = 1$ in (5) and observing that $1^\alpha = 1^\beta = 1$, we see that any (α, β) -derivation satisfies

$$1^\delta = 0. \quad (6)$$

It is easily verified that $\ker \delta$ is a subring of R , called the ring of *constants* (with respect to δ). Moreover, any element of $\ker \delta$ which is invertible in R is also invertible in $\ker \delta$, as follows by the formula (itself easily checked):

$$(x^{-1})^\delta = -(x^{-1})^\alpha \cdot x^\delta \cdot (x^{-1})^\beta. \quad (7)$$

We list some examples of derivations.

1. Let $R = A = B = k(t)$ be the field of rational functions in t over some field k , and let f' be the usual derivative of f , then on taking $\alpha = \beta = 1$, we obtain the familiar formula

$$(fg)' = fg' + f'g,$$

as a special case of (5).

2. Let $R = A = B$ and let α be any automorphism of R , then $f^\delta = f^\alpha - f$ is a right α -derivation. In particular, when $R = k[t]$ and $\alpha: f(t) \mapsto f(t+1)$, then δ is the differencing operator $f(t) \mapsto f(t+1) - f(t)$.

3. Take $m \in M$ and define $\delta_m: R \rightarrow M$ by the rule

$$\delta_m: x \mapsto x^\alpha m - mx^\beta. \quad (8)$$

This is easily seen to be an (α, β) -derivation; it is called the *inner* (α, β) -derivation induced by m . Thus the differencing operator in No. 2 is the inner α -derivation induced by 1, where α is the translation operator. A derivation which is not inner is said to be *outer*.

With any (A, B) -bimodule M we can associate the ring $\begin{pmatrix} A & M \\ 0 & B \end{pmatrix}$ consisting of all matrices

$$\begin{pmatrix} a & m \\ 0 & b \end{pmatrix} \quad (a \in A, b \in B, m \in M),$$

with the usual matrix addition and multiplication. The (A, B) -bimodule property just ensures that we get a ring in this way:

$$\begin{pmatrix} a & m \\ 0 & b \end{pmatrix} \begin{pmatrix} a' & m' \\ 0 & b' \end{pmatrix} = \begin{pmatrix} aa' & am' + mb' \\ 0 & bb' \end{pmatrix}.$$

Given maps $\alpha: R \rightarrow A$, $\beta: R \rightarrow B$, $\delta: R \rightarrow M$, we can define a map from R to $\begin{pmatrix} A & M \\ 0 & B \end{pmatrix}$ by the rule

$$x \mapsto \begin{pmatrix} x^\alpha & x^\delta \\ 0 & x^\beta \end{pmatrix}, \quad (9)$$

and it is easily checked that this is a ring homomorphism if and only if α, β are homomorphisms and δ is an (α, β) -derivation. This alternative method of defining derivations is often useful, e.g. in proving

THEOREM 8.11. *Let R, A, B be rings, T a multiplicative subset of $R^\times$ and M an (A, B) -bimodule. Then any T -inverting homomorphism $\alpha: R \rightarrow A$ extends to a unique homomorphism $\alpha': R_T \rightarrow A$, and given T -inverting homomorphisms $\alpha: R \rightarrow A$, $\beta: R \rightarrow B$, any (α, β) -derivation $\delta: R \rightarrow M$ extends to an (α', β') -derivation of R_T into M .*

Proof. The existence and uniqueness of α' follows because R_T is universal T -inverting. Now δ defines a homomorphism (9) from R to $\begin{pmatrix} A & M \\ 0 & B \end{pmatrix}$, which is T -inverting and therefore extends to a homomorphism of R_T : $x \mapsto \begin{pmatrix} x^{\alpha'} & x^{\delta'} \\ 0 & x^{\beta'} \end{pmatrix}$. It follows that δ' is an (α', β') -derivation. ■

We conclude this section by briefly discussing a special class of monoids that are embeddable in groups. Let S be a cancellation monoid; an element c of S is said to be *rigid* if $ab' = ba' = c$ implies that $a = bu$ or $b = au$ for some $u \in S$. Thus c is rigid if the left factors of c form a chain under the ordering by divisibility. When $ab' = ba' = c$ and $a = bu$, then $bub' = ba'$, hence $a' = ub'$; this shows the condition to be left-right symmetric. A monoid is said to be *rigid* if it admits cancellation and all its elements are rigid. Thus S is rigid if it is a cancellation monoid such that

$$aS \cap bS \neq \emptyset \text{ implies } aS \subseteq bS \text{ or } bS \subseteq aS. \quad (10)$$

THEOREM 8.12. *Every rigid monoid is embeddable in a group.*

Proof. Let S be a rigid monoid and for $a, b \in S$ denote by $\{a \cdot b\}$ the set of all elements of the form $a'b''$, where $a = a'a'$, $b = b'b''$. We first establish the following assertions:

- (α) $u \in \{ac \cdot b\}, u \notin \{ad \cdot b\} \Rightarrow u \in aS$,
- (β) $uc \in \{a \cdot bc\} \Rightarrow u \in \{a \cdot b\}$.

To prove (α) , we have by definition, $u = pq, ac = px, b = yq$. By rigidity, $a \in pS$ or $p \in aS$. In the first case $u \in \{a \cdot b\}$, hence $u \in \{ad \cdot b\}$, which is excluded, so $p \in aS$ and hence $u \in aS$. To prove (β) , let $uc = pq, a = px, bc = yq$. Now either $p = uz$, then $a = uzx$, and so $u \in \{a \cdot b\}$, or $u = pz$, then $q = zc$, so $b = yz$ and again $u \in \{a \cdot b\}$. Thus (α) and (β) are established.

We now consider the set of all expressions

$$p = a_0 b_1^{-1} a_1 b_2^{-1} \dots a_{n-1} b_n^{-1} a_n, \quad \text{where } a_i, b_i \in S. \quad (11)$$

The expression (11) is said to be of length n ; it is said to be *reduced* if

$$a_i \notin \{b_i \cdot b_{i+1}\} \quad (i = 1, \dots, n-1), \quad b_i \notin \{a_i \cdot a_{i-1}\} \quad (i = 1, \dots, n). \quad (12)$$

Clearly if (11) is reduced, then $a_i \neq 1$ ($i = 1, \dots, n-1$) and $b_i \neq 1$ ($i = 1, \dots, n$). We define the following elementary transformations on (11), for any $x \in S$:

$$\begin{aligned} R_i: a_{i-1} b_i^{-1} &\mapsto a_{i-1} x (b_i x)^{-1} & (i = 1, \dots, n), \\ L_i: b_i^{-1} a_i &\mapsto (x b_i)^{-1} x a_i & (i = 1, \dots, n). \end{aligned}$$

Two expressions are said to be *equivalent* if they are related by a finite chain of such elementary transformations and their inverses; clearly this is an equivalence. We note in particular that every element of S forms an expression of length 0, which is reduced and admits no elementary transformations.

We claim that a reduced expression stays reduced under elementary transformation. Consider the effect of R_j on (11) and the first condition (12). It is clear that this will not be affected unless $j = i$ or $i + 1$; we take these cases in turn.

R_i . Suppose that $a_i \in \{b_i x \cdot b_{i+1}\}$. Since $a_i \notin \{b_i \cdot b_{i+1}\}$, we have $a_i \in b_i S$ by (α) , but this contradicts the fact that $b_i \notin \{a_i \cdot a_{i-1}\}$.

R_{i+1} . Since $a_i \notin \{b_i \cdot b_{i+1}\}$, we have $a_i x \notin \{b_i \cdot b_{i+1} x\}$ by (β) .

R_i^{-1} . Let $a_{i-1} = a'_{i-1} x$, $b_i = b'_i x$; if $a_i \in \{b'_i \cdot b_{i+1}\}$, then $a_i \in \{b'_i x \cdot b_{i+1}\} = \{b_i \cdot b_{i+1}\}$, a contradiction.

R_{i+1}^{-1} . Let $a_i = a'_i x$, $b_{i+1} = b'_{i+1} x$ and $a'_i \in \{b'_i \cdot b'_{i+1}\}$, then

$$a_i = a'_i x \in \{b'_i \cdot b'_{i+1} x\} = \{b_i \cdot b_{i+1}\},$$

again a contradiction.

By symmetry L_j leaves the first condition (12) unaffected, and we can deal similarly with the second condition (12) by considering the formal inverse of p .

Thus the conditions (12) are unaffected by elementary transformations, so for any reduced expression the length is an invariant of the equivalence class. In particular, two expressions of length 0 are equivalent if and only if they are equal, as elements of S .

We now define $\mathcal{G}(S)$ as a group of permutations on the set of equivalence

classes of reduced expressions and verify that S acts faithfully; this will show that S is embedded in $\mathcal{G}(S)$. Given $c \in S$ and a reduced expression p as in (11), we define

$$pc = \begin{cases} a_0 b_1^{-1} a_1 \dots a_{n-1} b_n^{-1} a_n c & \text{if } b_n \notin \{a_n c, a_{n-1}\}, \\ a_0 b_1^{-1} a_1 \dots b_{n-1}^{-1} a' c'' & \text{if } c = c' c'', a_{n-1} = a' a'', b_n = a_n c' a''. \end{cases} \quad (13)$$

Clearly the first form is reduced; when it does not apply, we have $b_n \in \{a_n c, a_{n-1}\}$, but $b_n \notin \{a_n, a_{n-1}\}$, hence by (α) , $b_n = a_n u$ and so $a_n u \in \{a_n c, a_{n-1}\}$. By the left-right dual of (β) , $u \in \{c, a_{n-1}\}$, so $c = c' c''$, $a_{n-1} = a' a''$, $u = c' a''$ and $b_n = a_n c' a''$, which is the second alternative in (13). It is reduced, for if $b_{n-1} \in \{a' c'', a_{n-2}\}$, then since $b_{n-1} \notin \{a_{n-1}, a_{n-2}\}$, we have $b_{n-1} \notin \{a', a_{n-2}\}$ and so by (α) , $b_{n-1} = a' v$, but then $a_{n-1} \in \{a' v, a_n c' a''\}$, which contradicts the condition $a_{n-1} \notin \{b_{n-1}, b_n\}$.

A routine verification shows that this action is compatible with the elementary transformations R_i , L_i and their inverses, so that (13) defines an action on the equivalence classes. Next we define, for $c \in S$ and p as in (11),

$$pc^{-1} = \begin{cases} a_0 b_1^{-1} a_1 \dots b_n^{-1} a_n c^{-1} & \text{if } c \notin \{1, a_n\} \text{ and } a_n \notin \{b_n, c\}, \\ a_0 b_1^{-1} a_1 \dots b_n^{-1} u & \text{if } a_n = uc, b_n \notin uS, \\ a_0 b_1^{-1} a_1 \dots b_{n-1}^{-1} a_{n-1} (c' b'')^{-1} & \text{if } a_n = b' c'', b_n = b' b'', c = c' c''. \end{cases} \quad (14)$$

The first form is clearly reduced. If it does not hold, suppose that $c \in \{1, a_n\}$, say $a_n = uc$, but $a_n \notin \{b_n, c\}$. Then $b_n \notin uS$ and by hypothesis, $b_n \notin \{uc, a_{n-1}\} = \{a_n, a_{n-1}\}$, hence $b_n \notin \{u, a_{n-1}\}$ and this shows the second form to be reduced. Finally, if $c \notin \{1, a_n\}$ but $a_n \in \{b_n, c\}$, let $c = c' c''$, $b_n = b' b''$, $a_n = b' c''$, then we are in the third case, and it will be reduced if we can show that $a_{n-1} \notin \{b_{n-1}, c' b''\}$. So suppose that $a_{n-1} \in \{b_{n-1}, c' b''\}$; since $a_{n-1} \notin \{b_{n-1}, b_n\}$, we have $a_{n-1} \in Sb''$ by the left-right dual of (α) , say $a_{n-1} = vb''$, but then $b_n \in \{a_n, a_{n-1}\}$, which is a contradiction. Again it is straightforward to show that the action is compatible with L_i , R_i .

It remains to verify that we have a representation; we shall use $\sim$ to indicate equivalence.

(i) $cc^{-1} = 1$. If pc has the first form (13), it is clear that $pcc^{-1} = p$. For the second form we have

$$\begin{aligned} pcc^{-1} &= a_0 b_1^{-1} a_1 \dots b_{n-1}^{-1} a' c'^{-1} \\ &\sim a_0 b_1^{-1} a_1 \dots b_{n-1}^{-1} a_{n-1} (c' a'')^{-1} \sim p. \end{aligned}$$

(ii) $c^{-1}c = 1$. If pc^{-1} has the first form (14), all is clear. If the second applies, we have

$$pc^{-1}c = a_0 b_1^{-1} \dots b_n^{-1} uc \sim p,$$

and for the third,

$$\begin{aligned} pc^{-1}c &= a_0b_1^{-1} \dots b_{n-1}^{-1}a_{n-1}(c'b'')^{-1}c \\ &\sim a_0b_1^{-1} \dots b_{n-1}^{-1}a_{n-1}b''^{-1}c'' \sim p. \end{aligned}$$

(iii) $(pc)d = p(cd)$. If $p(cd)$ has the first form (13), then $b_n \notin \{a_ncd, a_{n-1}\}$, hence $b_n \notin \{a_nc, a_{n-1}\}$, and so

$$(pc)d = a_0b_1^{-1} \dots b_n^{-1}a_ncd = p(cd).$$

If $p(cd)$ has the second form (13) and $b_{n-1} \notin \{a'c'd, a_{n-2}\}$, then

$$(pc)d = a_0b_1^{-1} \dots b_{n-1}^{-1}a'c'd = p(cd).$$

If $b_{n-1} \in \{a'c'd, a_{n-2}\}$, say $a'c'd = e'e'', a_{n-2} = f'f'', b_{n-1} = e'f''$, then

$$(pc)d = a_0b_1^{-1} \dots b_{n-2}^{-1}f'e' = p(cd).$$

It is clear that $p1 = p$, so we have a representation of S by permutations of the classes of reduced expressions (11). Further, for any $x, y \in S$, if x and y have the same action, then $x = 1 \cdot x = 1 \cdot y = y$, and this shows that S acts faithfully, hence S is embedded in $\mathcal{G}(S)$, as claimed. ■

Exercises 0.8

1. In a monoid S , if aba is invertible, show that a and b are both invertible. Show also that it is not enough for ab to be invertible.

2. Let S be a cancellation monoid; show that any central submonoid (i.e. contained in the centre of S) is a right (and left) Ore set.

3. Let R be any ring; show that any $R^\times$ -inverting homomorphism into a non-zero ring must be injective.

4. Let R be an integral domain. Show that any Ore subring of R is contained in a maximal Ore subring.

5. Show that a direct limit of rigid monoids is again rigid. Show that a direct limit of Ore domains is again Ore.

6. Let R be a ring and T a right Ore set in R . If R is (left or right) Ore, Noetherian or Artinian, show that the same is true of R_T .

7. Let R be a ring and S a multiplicative subset such that (i) for each $a \in R, s \in S$, $aS \cap sR \neq \emptyset$, (ii) for each $a \in R, s \in S$, if $sa = 0$, then $at = 0$ for some $t \in S$. Show that the universal S -inverting ring R_S can be constructed as the set of fractions as^{-1} , where $a_1s_1^{-1} = a_2s_2^{-1}$ if and only if $a_1u = a_2v$ and $s_1u = s_2v \in S$ for some $u, v \in R$. Show that the kernel of the natural homomorphism from R to R_S is $\{a \in R \mid at = 0 \text{ for some } t \in S\}$.

8. Describe the analogue of Ex. 7 for monoids and determine the kernel in this case (the *kernel* of a monoid homomorphism f from S is defined as the set of pairs $(s, t) \in S^2$ such that $sf = tf$).

9. In any monoid, show that any left factor of a right large element is again right large. In a cancellation monoid, is the same true of every right factor?

10. Let R be a right Ore domain with right ACC_1 and a an ideal of R which is principal as left ideal. If R/a is an integral domain, show that it is again a right Ore domain.

11. If R is an ordered ring which is an Ore domain, show that the ordering can be extended in a unique way to the field of fractions of R .

12*. Let R be an integral domain which is not right Ore and let $n \geq 1$. Show that ${}^nR$ can be embedded in R as right ideal, and if ${}^nR$ does not have unique rank, show that ${}^nR$ contains a strictly descending chain of direct summands that are free of rank n . Deduce that if an integral domain satisfies right ACC_n then ${}^nR$ has unique rank.

13. Let R be a right Ore domain and K its field of fractions. If $A \in R_n$ is right regular in R_n , show that it is right regular in K_n and hence is invertible, with an inverse of the form Bd^{-1} , $B \in R_n$, $d \in R^*$. Deduce that every right zerodivisor in R_n is a left zerodivisor. Does the reverse implication hold generally?

14. Let $E \supset F$ be a skew field extension of finite right dimension. Show that in the polynomial ring $E[x]$ the monic polynomials with coefficients in F form a right Ore set. (Hint: In the relation $au' = ua'$ ($a, a' \in E[x]$, $u, u' \in F[x]$, u monic) equate coefficients and eliminate the coefficients of a' . This gives a system of linear equations in the coefficients of u' , say $\sum v_i \mu_i = 0$, if $u' = \sum x^i \mu_i$. Show that v_i is independent of the degree of u' and hence the v_i are linearly dependent over F .)

15. (Makar-Limanov) Let S be a cancellation monoid. Given $a, b \in S$, denote by T the submonoid generated by a, b . Show that if $aT \cap bT = \emptyset$, then T is free on a, b . Deduce that a cancellation monoid containing no free submonoid on more than one element can be embedded in a group.

16. If S is a monoid satisfying the hypotheses of Prop. 8.4, show that S_L is rigid.

17. Let R be a ring with IBN and S a right Ore set; show that the localization R_S need not have IBN. (Hint: Take a ring with generators a_{ij}, b_{ji}, λ and relations stating (in matrix form) $AB = \lambda I, BA = \lambda I, \lambda$ central.)

18°. Let R be a right hereditary right Ore domain. Can every right ideal be generated by 2 elements? (This is true in the commutative case but, as we shall see later, false in the non-Ore case.)

19*. (S. Rosset) Let G be a group and A a torsion-free abelian normal subgroup of G . Show that in the group algebra kG , the set $(kA)^*$ is a left and right Ore set consisting of regular elements.

20. Show that the kernel of a derivation acting on a local ring is again a local ring.

21. Prove Leibniz's formula for derivations:

$$(ab)\delta^n = \sum \binom{n}{i} (a\delta^i)(b\delta^{n-i}).$$

More generally, if δ is an α -derivation, show that

$$(ab)\delta^n = \sum \alpha \delta^i \cdot b f_i^n(\alpha, \delta),$$

where $f_i^n(\alpha, \delta)$ is the coefficient of t^i in the formal expansion of $(t\alpha + \delta)^n$.

22. If δ is a derivation on an integral domain of characteristic p , show that δ^p is again a derivation.

23. If δ is a nilpotent derivation of exponent r on an integral domain K (i.e. $\delta^r = 0 \neq \delta^{r-1}$) and $r > 1$, show that K has prime characteristic p and $r = p^t$. (Hint: Apply δ^r to ab , where $b\delta \neq 0 = b\delta^2$ and use Leibniz's formula to show that $p|r$. Now repeat the argument with δ replaced by δ^p .)

24. Let R be an Hermite ring and T a right Ore set; show that the localization R_T need not be Hermite. (Hint: See Ex. 17.)

0.9 Modules over Ore domains

Many results on modules over commutative integral domains hold more generally either for right modules or for left modules over right Ore domains. For convenience we shall deal with left modules over left or right Ore domains in this section and leave the reader to make the necessary modifications.

Let R be an integral domain and M a left R -module. An element $x \in M$ is called a *torsion element* if $ax = 0$ for some $a \in R^\times$. When R is a left Ore domain, the set tM of all torsion elements of M is a submodule, the *torsion submodule* of M . To prove that tM is a submodule, let $x \in tM$, say $ax = 0$ for $a \in R^\times$, and take $b \in R^\times$. Then there exist $a', b' \in R$ such that $a'b = b'a \neq 0$, hence $a'bx = b'ax = 0$ and so $bx \in tM$ for all $b \in R^\times$; the closure under addition is proved similarly. If $tM = 0$, we say that M is *torsion free*, if $tM = M$, we call M a *torsion module*.† It is clear that for any module M , tM is a torsion module and M/tM is torsion free; moreover, these two properties serve to determine tM .

Let R be a ring and T a left Ore set in R ; then the localization R_T may be expressed as a direct limit

$$R_T = \varinjlim \{t^{-1}R \mid t \in T\}.$$

† This is the customary definition, but it is mainly applied to modules over Ore domains; we shall use the term in a slightly different sense later (in 3.3 and 5.3).

For, given $t_1, t_2 \in T$, there exists $t \in Tt_1 \cap Tt_2$ and so $t_1^{-1}R \cup t_2^{-1}R \subseteq t^{-1}R$. This process can be applied to modules as well as rings; for simplicity we state the result only for Ore domains, the case of principal interest.

PROPOSITION 9.1. *Let R be a left Ore domain, K its field of fractions and M a left R -module. Then $K \otimes_R M$ can be described as the set of formal products $b^{-1}x$ ($x \in M, b \in R^\times$) subject to the relations: $b^{-1}x = b'^{-1}x'$ if and only if there exist $u, v \in R^\times$ such that $ux = vx', ub = vb'$. Moreover, the kernel of the canonical map*

$$M \rightarrow K \otimes M \quad (1)$$

is the torsion submodule tM , so (1) is an embedding if and only if M is torsion free.

Proof. Any element of $K \otimes M$ has the form $x = \sum b_i^{-1}a_i \otimes x_i$. If b is a common left multiple for the $b_i: c_i b_i = b$, then

$$x = \sum b^{-1}c_i a_i \otimes x_i = b^{-1}(\sum c_i a_i x_i).$$

Thus every element of $K \otimes M$ has the form $b^{-1}x, x \in M, b \in R^\times$. Given $p = b^{-1}x$ and $p' = b'^{-1}x'$, there exist $u, v \in R^\times$ such that $ub = vb' = c$, and we have $cp = ux, cp' = vx'$. Clearly $p = p'$ if and only if $cp = cp'$, i.e. $tux = tvx'$ in M for some $t \in R^\times$. Now it follows that $b^{-1}x = 0$ if and only if $ux = 0$ for some $u \in R^\times$, i.e. precisely if $x \in tM$. Hence the kernel of (1) is tM and the rest is clear. ■

For a right R -module there is no such convenient description, but in that case there are two ways of describing the linear functionals on M , using the dual $M^* = \text{Hom}_R(M, R)$.

PROPOSITION 9.2. *Let R be a right Ore domain with field of fractions K , and let M be a finitely generated left R -module. Then there is a natural isomorphism of right K -modules*

$$M^* \otimes_R K \cong \text{Hom}_K(K \otimes M, K). \quad (2)$$

Proof. By adjoint associativity, applied to $({}_K K_R, {}_R M, {}_K K)$, we have

$$\begin{aligned} \text{Hom}_K(K \otimes_R M, K) &\cong \text{Hom}_R(M, K) \\ &\cong \text{Hom}_R(M, \varinjlim Rb^{-1}). \end{aligned}$$

Since M is finitely generated, we can find a common denominator for the images of M , so we can replace $\varinjlim Rb^{-1}$ by Rb^{-1} for any given homomor-

phism; thus we have

$$\begin{aligned}\operatorname{Hom}_R(M, \varinjlim Rb^{-1}) &\cong \varinjlim \operatorname{Hom}_R(M, R)b^{-1} \\ &\cong M^* \otimes_R K. \blacksquare\end{aligned}$$

We have seen that the field of fractions K of a left Ore domain R has the form $K = \varinjlim b^{-1}R$; here each $b^{-1}R$ is a free right R -module. Let us call a module *semifree*[†] if every finitely generated submodule is contained in a finitely generated free module. Then we can say that K is semifree as right R -module, hence flat, and so by the independence property of the tensor product, if a family of elements in a left R -module M is linearly independent, so is its image in $K \otimes M$. Therefore the dimension of $K \otimes M$ as vector space over K equals the cardinal of a maximal linearly independent subset of M . This number is an invariant of M which we shall call the *rank* of M and denote by $\operatorname{rk} M$. In particular, $\operatorname{rk} M = 0$ if and only if M is a torsion module. On free modules the rank clearly agrees with our previous definition of rank, and since tensoring preserves exactness, we have

PROPOSITION 9.3. *Let R be a left Ore domain. If $0 \rightarrow M' \rightarrow M \rightarrow M'' \rightarrow 0$ is an exact sequence of left R -modules, then*

$$\operatorname{rk} M = \operatorname{rk} M' + \operatorname{rk} M''.$$

In particular, if N is a submodule or a homomorphic image of M , then $\operatorname{rk} N \leq \operatorname{rk} M$. ■

The last assertion, relating to homomorphic images, holds (under an appropriate definition of rank) for a large class of rings, including all that can be embedded in fields, and hence most of the rings considered later. However, apart from this, none of the other assertions holds with 'right' in place of 'left' Ore domain. Thus let R be any right Ore domain (or, indeed, any integral domain) that is not left Ore and let x, y be two left linearly independent elements of R , then R contains the left ideal $Rx + Ry$ which is isomorphic to R^2 ; this shows that the first part of Prop. 9.3 cannot be extended to such rings. For an example showing that K need not be semifree as left R -module see Ex. 5.1.7. This also shows that Prop. 9.1 does not extend; however, it remains true for right Bezout domains (cf. Prop. 1.1.9).

The following property of right Ore domains is not in general shared by left Ore domains (cf. Ex. 3).

[†] This is sometimes called 'locally free', but we shall avoid that term, as it has quite a different meaning in commutative algebra.

PROPOSITION 9.4. *Let R be a right Ore domain and K its field of fractions. Then any left K -module, considered as left R -module, is semifree.*

Proof. Let M be a finitely generated R -submodule of a left K -module, which may without loss of generality be taken to be K^n , for some n . We can choose a common right denominator $c \in R^\times$ for the components of the finite generating set of M ; then $M \subseteq R^n \cdot c^{-1}$ and the latter is a free R -module. ■

Combining this with Prop. 9.1, we obtain

COROLLARY 9.5. *If R is a left and right Ore domain, then every finitely generated torsion free R -module is embeddable in a free R -module.* ■

Finally we note that the flatness of the ring of fractions, well known in the commutative case, continues to hold in the Ore case.

PROPOSITION 9.6. *Let R be a ring and T a right Ore set. Then R_T is left R -flat. If R is any integral domain, then $R_{R^\times}$ is non-zero and left R -flat if and only if R is a right Ore domain.*

Proof. We check that the criterion given in Appendix 2 (Prop. A.4, p. 553) holds. Suppose that $\sum c_i u_i = 0$, where $c_i \in R$, $u_i \in R_T$ and write $u_i = a_i b^{-1}$ with a common denominator b , then $\sum c_i a_i = 0$ and $u_i = a_i b^{-1}$. Hence R_T is left R -flat, in particular $K = R_{R^\times}$ is so when R is right Ore. Conversely, if $R_{R^\times}$ is left R -flat and non-zero, take $a, b \in R^\times$, then $a \cdot a^{-1} - b \cdot b^{-1} = 0$, hence there exist $u_i \in K$, $p_i, q_i \in R$ such that $a^{-1} = \sum p_i u_i$, $b^{-1} = \sum q_i u_i$, $ap_i - bq_i = 0$. Not all the p_i, q_i can vanish, say $p_1, q_1 \neq 0$, then $ap_1 = bq_1$ is the desired right multiple. ■

The last part of this proposition shows in effect that if R is a right Ore domain with field of fractions K , then K is left R -flat but not right R -flat unless R is also left Ore.

Exercises 0.9

1. Let K be a field, F a subfield of its centre and E a commutative field containing F . Show that $K \otimes_F E$ is an Ore domain, provided that it is an integral domain.

2. Let F be a commutative field, E an algebraic commutative field extension and A an F -algebra which is a right Ore domain with field of fractions K . If $A \otimes_F E$ is an integral domain, show that it is a right Ore domain with field of fractions $K \otimes_F E$.

3. Let R be a left but not right Ore domain and K its field of fractions. Show that K , as left R -module, has rank 1 but is not semifree (see Ex. 5.1.7).

4. (Gentile [60]) Let R be a subring of a field. If every torsion free left R -module can be embedded in a free left R -module, show that R is right Ore. Note that this is a converse to Prop. 9.4. Investigate the truth of other possible converses.

5. (Chalabi and Khuri [80]) Show that an integral domain R is left Ore provided that there exists a non-zero torsion free divisible left R -module (M is *divisible* if every equation $ax = m$, $m \in M$, $a \in R^\times$, has a solution in M).

6*. Show that a projective left ideal α of a left Ore domain is finitely generated. (Hint: Use a projective coordinate system to show that α is invertible or 0, cf. Prop. 4 of 4.7, p. 126 of Cohn [77].) Deduce that every projective left R -module which is uniform (i.e. any two non-zero submodules have a non-zero intersection) is finitely generated.

7*. Let R be a right Ore domain, K its field of fractions and α any non-zero right ideal of R . Show that $\alpha \otimes_R K \cong K$ (as right R -modules). Deduce that for any left or right Ore domain R with field of fractions K , $K \otimes_R K \cong K$. (This is equivalent to the assertion that the embedding $R \rightarrow K$ is an epimorphism in the category of rings; this equivalence actually holds for any ring R with a homomorphism to a field K , cf. Ex. 7.2.6.)

8. (Bergman [67]) Let R be a right Ore domain and K its field of fractions. Prove that the following conditions on a finitely generated left K -module M are equivalent:

- (a) the canonical map $M \rightarrow K \otimes M$ is an embedding,
- (b) M is embeddable in a K -module (*qua* left R -module),
- (c) M is embeddable in a free left R -module,
- (d) $\text{Hom}_R(M, R)$ distinguishes elements of M ,
- (e) $\text{Hom}_R(M, K)$ distinguishes elements of M .

9°. Find the relations between (a)–(e) of Ex. 8 when (i) K is a field and R a subring generating K as a field, (ii) K is any ring and R a subring. Find conditions on the finitely generated R -module M for (a)–(e) to be equivalent.

10. Show that for a finitely generated left R -module M over a (left and right) Ore domain with field of fractions K , $K \otimes M \cong K \otimes M^{**}$.

0.10 Skew polynomial rings

There is one type of ring that is especially useful in providing examples and counter-examples at the simplest level. It generalizes the usual polynomial ring $k[x]$ in one indeterminate x over a field k , but differs from it in that k need not be commutative, nor commute with x .

The degree of a polynomial is an example of a degree-function which we

shall define quite generally as follows. In any ring R a *degree-function* is a function which assigns to each $a \in R$ a degree $d(a)$ such that

- D.1. for $a \in R^\times$, $d(a) \in \mathbb{N}$ while $d(0) = -\infty$,
- D.2. $d(a-b) \leq \max\{d(a), d(b)\}$,
- D.3. $d(ab) = d(a) + d(b)$.

By D.3, $d(1) = 0$. By D.2 we have, as for valuations,

$$d(a) = d(-a), \quad (1)$$

and

$$d(a+b) \leq \max\{d(a), d(b)\}, \quad (2)$$

with equality holding whenever $d(a) \neq d(b)$ ('every triangle is isosceles'). By D.1 and D.3, $R^\times$ is closed under multiplication and contains 1, thus every ring with a degree-function is an integral domain. For example, this applies to any polynomial ring over a field or, more generally, over an integral domain. We remark that any integral domain has the *trivial* degree-function which assigns degree zero to each non-zero element.

Now let R be any ring and S a ring containing R as subring, as well as an element x such that every element of the ring A generated by R and x is uniquely expressible in the form

$$f = a_0 + xa_1 + \dots + x^n a_n \quad (a_i \in R). \quad (3)$$

Further, the function

$$d(f) = \max\{i \mid a_i \neq 0\} \quad (4)$$

is assumed to be a degree-function on A , in the sense just explained. Of the properties D.1–3 only D.3 represents a restriction on the base ring R . It entails that R is an integral domain and moreover, for any $a \in R$, there exist $a^\alpha, a^\delta \in R$ such that

$$ax = xa^\alpha + a^\delta \quad (a \in R). \quad (5)$$

In the first place we note that a^α, a^δ are uniquely determined by a , and $a^\alpha = 0$ if and only if $a = 0$. Comparing the expressions for $(a+b)x$ and $ax + bx$, we find that

$$(a+b)^\alpha = a^\alpha + b^\alpha, \quad (a+b)^\delta = a^\delta + b^\delta, \quad (6)$$

thus α and δ are additive maps of R . Similarly a comparison of $a(bx)$ and $(ab)x$ shows that

$$(ab)^\alpha = a^\alpha b^\alpha, \quad (ab)^\delta = a^\delta b^\alpha + ab^\delta. \quad (7)$$

By putting $a = b = 1$ and observing that R is an integral domain, we see that

$$1^\alpha = 1, \quad 1^\delta = 0. \quad (8)$$

Hence α is an endomorphism of R , injective by the remark following (5). Now (6) and (7) show that δ is just an α -derivation of R . The law (5), with the uniqueness of (3), suffices to determine the multiplication on A in terms of R , α and δ . For by the distributive law we need only know $x^m a \cdot x^n b$, and from (5) we see that

$$x^m a \cdot x^n b = x^{m+1} a^\alpha x^{n-1} b + x^m a^\delta x^{n-1} b.$$

By induction on n this defines $x^m a x^n b$ in all cases. Thus A is completely fixed when R, α, δ are given. We shall write

$$A = R[x; \alpha, \delta] \quad (9)$$

and call A the *skew polynomial ring* in x over R determined by α, δ . When $\delta = 0$, we shall write $R[x; \alpha]$ instead of $R[x; \alpha, 0]$, and when $\alpha = 1, \delta = 0$, we write $R[x]$ as usual; this is just the polynomial ring in a central indeterminate x .

It remains to show that for any integral domain R with an injective endomorphism α and an α -derivation δ , there always exists a skew polynomial ring $R[x; \alpha, \delta]$. Consider the direct power $M = R^{\mathbb{N}}$ as right R -module and define an additive group endomorphism of M by

$$x: (a_i) \mapsto (a_i^\delta + a_{i-1}^\alpha) \quad (a_{-1} = 0). \quad (10)$$

If we identify R with its image in $\text{End}(M)$ by right multiplication then the action of the endomorphism x defined by (10) satisfies the rule

$$\begin{aligned} (c_i)ax &= (c_i a)x = ((c_i a)^\delta + (c_{i-1} a)^\alpha) \\ &= (c_i^\delta a^\alpha + c_i a^\delta + c_{i-1}^\alpha a^\alpha) \\ &= (c_i^\delta + c_{i-1}^\alpha) a^\alpha + (c_i) a^\delta \\ &= (c_i) (x \cdot a^\alpha + a^\delta). \end{aligned}$$

Hence (5) holds and, moreover, every element of the subring of $\text{End}(M)$ generated by R and x can be brought to the form (3). This form is unique, because when $a_0 + xa_1 + \dots + x^n a_n$ is applied to $(1, 0, 0, \dots)$, it produces $(a_0, a_1, \dots, a_n, 0, \dots)$. Moreover, the function $d(f)$ defined by (4) is easily seen to be a degree-function, using the fact that R is an integral domain and α is injective. Hence the polynomial ring is also an integral domain. Thus we have proved

THEOREM 10.1. *Let R be an integral domain with an injective endomorphism α and an α -derivation δ . Then there is a skew polynomial ring $R[x; \alpha, \delta]$ which is an integral domain. Every skew polynomial ring arises in this way. ■*

The result is not left-right symmetric because in (5) the coefficients were written on the right. One therefore sometimes introduces the *left* skew polynomial ring, in which the commutation rule, instead of (5) is

$$xa = a^\alpha x + a^\delta. \quad (11)$$

In general $R[x; \alpha, \delta]$ will not be a left skew polynomial ring; but when α is an automorphism of R , with inverse β say, then on replacing a by a^β in (5) and rearranging the terms, we obtain

$$xa = a^\beta x - a^{\beta\delta}. \quad (12)$$

Thus we have

PROPOSITION 10.2. *The ring $R[x; \alpha, \delta]$ is a left skew polynomial ring provided that α is an automorphism. ■*

We have an analogue of the Hilbert basis theorem for skew polynomial rings, with a similar proof.

PROPOSITION 10.3. *Let R be a right Noetherian domain, α an automorphism and δ an α -derivation of R . Then the skew polynomial ring $A = R[x; \alpha, \delta]$ is again right Noetherian.*

The proof is essentially as in the commutative case: Let us consider A as left skew polynomial ring, i.e. with coefficients on the left. This is permissible because α is an automorphism. If A is not right Noetherian, let $\mathfrak{a}$ be a right ideal which is not finitely generated. Let $f_1 \in \mathfrak{a}$ be a non-zero polynomial of least degree; given $f_1, \dots, f_k \in \mathfrak{a}$, we take f_{k+1} in $\mathfrak{a} \setminus \sum_1^k f_i A$ of least degree. Since $\mathfrak{a}$ is not finitely generated, we thus obtain an infinite sequence $f_1, f_2, \dots$; let f_i have degree n_i and leading coefficient a_i , then $n_1 \leq n_2 \leq \dots$. We claim that $a_1 R \subset a_1 R + a_2 R \subset \dots$ is an infinite ascending chain, which will contradict the fact that R is right Noetherian. For an equation $a_{k+1} = \sum_1^k a_i b_i$ ($b_i \in R$) would imply that $g = f_{k+1} - \sum f_i b_i^{\alpha^{-n_i}} x^{n_{k+1}-n_i} \in \mathfrak{a} \setminus \sum_1^k f_i A$, but g has lower degree than f_{k+1} , which is a contradiction. ■

From Prop. 10.3 and Cor. 8.10 we see that any skew polynomial ring $K[x; \alpha, \delta]$ over a field K is a right Ore domain and hence has a field of fractions, which we shall call a *skew rational function field* and denote by

$K(x; \alpha, \delta)$. More generally, let R be a right Ore domain with field of fractions K . If α is an injective endomorphism of R and δ an α -derivation, they can be extended to K , by Th. 8.11, and we have the inclusions

$$R[x; \alpha, \delta] \subseteq K[x; \alpha, \delta] \subseteq K(x; \alpha, \delta).$$

Any element u of $K(x; \alpha, \delta)$ has the form fg^{-1} , where $f, g \in K[x; \alpha, \delta]$. On bringing the coefficients of f and g to a common right denominator we can write $f = f_1 c^{-1}$, $g = g_1 c^{-1}$, where $f_1, g_1 \in R[x; \alpha, \delta]$, $c \in R^\times$. Hence $u = fg^{-1} = f_1 g_1^{-1}$ and we have proved

PROPOSITION 10.4. *A skew polynomial ring over a right Ore domain is again a right Ore domain. ■*

Since we localized here at the set of non-zero polynomials, we obtain by Th. 8.6,

COROLLARY 10.5. *If R is a right Ore domain with an injective endomorphism α and α -derivation δ , then the non-zero polynomials form a right Ore set. ■*

We now give some examples of skew polynomial rings, both as illustration and for later use.

1. The *complex-skew* polynomial ring $C[x; -]$ consists of polynomials with complex coefficients and commutation rule

$$ax = x\bar{a}, \quad \text{where } \bar{a} \text{ is the complex conjugate of } a.$$

We observe that the centre of this ring is $R[x^2]$, the ring of real polynomials in x^2 . The residue class ring $\text{mod } x^2 + 1$ is the field of real quaternions.

2. Let R be an integral domain and α an automorphism of R , then we can localize the skew polynomial ring $R[x; \alpha]$ at the set of powers of x ; the resulting ring is denoted by $R[x, x^{-1}; \alpha]$ and is called the *skew Laurent polynomial ring*.

3. Let $k = F(t)$ be the rational function field in an indeterminate t over a commutative field F . The usual derivative $f \mapsto f'$ defines a derivation on k , and this gives rise to a skew polynomial ring $R = k[x; 1, ']$, the ring of differential operators.

4. Let F be a commutative field of characteristic $p \neq 0$ and E/F a separable extension of degree p , say $E = F(\xi)$, where $\xi^p - \xi \in F$. The map $\alpha: f(\xi) \mapsto f(\xi + 1)$ defines an automorphism of order p and we have the skew polynomial ring $E[\eta; \alpha]$.

5. Let K be any field and denote by $A_1(K)$ the K -ring generated by x, y

centralizing K , with the defining relation $xy - yx = 1$. This ring $A_1(K)$, called the *Weyl algebra* on x, y over K , may also be defined as skew polynomial ring $A[y; 1, \cdot]$, where $A = K[x]$ and $\cdot$ is the derivation with respect to x (as in No. 3). The example in No. 3, above, is obtained by localizing at the set of all monic polynomials in x over k , and the example in No. 4 by putting $\xi = xy$, $\eta = y$ and then localizing at the set of all monic polynomials in ξ over F .

The Weyl algebra is useful as an example of a finitely generated infinite-dimensional algebra which in characteristic 0 is simple. For in any non-zero ideal we can pick an element $f(x, y)$ of least possible x -degree. The ideal still contains $df/dx = fy - yf$; this is of lower degree and so must be 0. Therefore $f = f(y)$ is a polynomial in y alone. If its y -degree is minimal, then $df/dy = xf - fx = 0$, hence $f = c \in K^\times$, so the ideal contains a non-zero element of the ground field and so must be the whole ring. This shows $A_1(K)$ to be simple. Further, $A_1(K)$ is Noetherian, by Prop. 10.3.

6. Let q be a prime power, F_q the field of q elements and T the endomorphism $f \mapsto f^q$ of $F_q[x]$. If the operation of multiplying by $a \in F_q$ is simply denoted by a , then each polynomial $\sum a_i T^i$ defines an endomorphism of $F_q[x]$, and it is easily verified that (applying endomorphisms on the right) we have $aT = Ta^q$; hence, the endomorphisms form a skew polynomial ring $F_q[T; \sigma]$, where $\sigma: a \mapsto a^q$. (This has an application to the theory of finite fields, cf. Ore [33].)

7. The Laurent polynomial ring may be thought of as the group algebra of the infinite cyclic group and many of the concepts introduced apply to general group algebras. Particularly this is true of derivations. Let G be any group, kG its group algebra over a field k and $\varepsilon: kG \rightarrow k$ the augmentation map, defined as a k -linear map such that $g\varepsilon = 1$ for all $g \in G$. The $(\varepsilon, 1)$ -derivations of kF , where F is a free group, are called the *Fox derivatives* (Fox [53]). If F is free on $X = \{x_i\}$, each map $d: X \rightarrow kF$ defines a unique Fox derivative; each such d can be written as a linear combination of the d_i , where d_i maps x_j to δ_{ij} . Moreover, any $u \in kF$ has the form $u = u\varepsilon + \sum (x_i - 1)d_i u$, while any Fox derivative d can be written $df = \sum dx_i \cdot d_i f$.

Besides polynomial rings we shall also need formal power series rings. Taking the case of a zero derivation, we can describe the formal power series ring $A = R[[x; \alpha]]$ over a ring R as the set of all infinite series

$$f = a_0 + \alpha a_1 + \alpha^2 a_2 + \dots \quad (13)$$

with componentwise addition and multiplication based on the commutation rule $ax = \alpha x a$. With each power series f we associate its *order* $o(f)$, defined as the suffix of the first non-zero coefficient in (13). When R is an integral domain and α a monomorphism, this satisfies the conditions for an order-function analogous to D.1-3:

- O.1. for $a \in R^\times$, $o(a) \in \mathbb{N}$, while $o(0) = \infty$,
 O.2. $o(a-b) \geq \min\{o(a), o(b)\}$,
 O.3. $o(ab) = o(a) + o(b)$.

From 0.3 we see that we shall not have an order function unless $\delta = 0$. However, we may now allow higher powers of x ; we shall return to this point in 8.7.

In the formal power series ring $R[[x; \alpha]]$ the powers of x form a (left and right) Ore set and on localizing at this set we obtain the ring $R((x; \alpha))$ of all Laurent series $\sum_{-\infty}^{\infty} x^i a_i$. This ring will be discussed in greater detail in 8.7.

Skew polynomial rings are often useful in constructing counter-examples; to end this section we examine the conditions for such a ring to be non-Hermite.

PROPOSITION 10.6. *Let A be a right Ore domain, α an automorphism and δ an α -derivation of A , and write (i) $R = A[x; \alpha, \delta]$ or (ii) $R = A[x, x^{-1}; \alpha]$. If $a, c \in A$ are such that a is a non-unit and $a, x+c$ are right comaximal in R , and in case (ii) $ca^\alpha \notin aA$, then $aR \cap (x+c)R$ is non-principal, hence R is not 2-Hermite.*

Proof. By Prop. 4.2 we have only to show that $aR \cap (x+c)R$ is not principal. In case (i) every element of R is a polynomial in x , while in case (ii) every element is associated to a polynomial; moreover, the monic polynomials form a right Ore set, so there exists a monic polynomial $f \in R$ such that $af = (x+c)u$.

Secondly, we have

$$ax = xa^\alpha + a^\delta = (x+c)a^\alpha + (a^\delta - ca^\alpha), \quad (14)$$

where δ is to be taken to be 0 in case (ii). By hypothesis there exist $a_1, a_2 \in A$, $a_1 \neq 0$, such that $(a^\delta - ca^\alpha)a_1 = aa_2$, hence on multiplying (14) by a_1 and simplifying, we obtain

$$a(xa_1 - a_2) = (x+c)a^\alpha a_1.$$

We put $g = xa_1 - a_2$. If $aR \cap (x+c)R$ were principal, equal to ahR , say, then $g \in hR$, so $\deg h \leq 1$. Thus h (after multiplication by a suitable unit in case (ii)) has the form $h = xb_1 + b_2$, where $b_i \in A$. Since $ah \in (x+c)R$, $\deg ah \geq 1$, so $\deg h \geq 1$ and $b_1 \neq 0$. Now $f \in hR$, say $f = hd$, where $f = x^r + x^{r-1}f_1 + \dots + f_r$, $d = x^{r-1}d_1 + \dots + d_r$. A comparison of highest terms gives $b_1^{\alpha^{r-1}}d_1 = 1$, and this shows b_1 to be a unit. Multiplying by b_1^{-1} we may take h in the form $h = x+b$ ($b \in A$). By definition of h we have $a(x+b) = (x+c)k$, i.e.

$$a(x+b) = xa^\alpha + a^\delta + ab = (x+c)k \quad \text{for some } k \in R. \quad (15)$$

It follows that $k = a^\alpha$, $ck = ab + a^\delta$, and so

$$ca^\alpha = ab + a^\delta. \quad (16)$$

In case (ii) $a^\delta = 0$ and (16) is excluded by hypothesis. When (i) holds, we have

$$(x+c)p + aq = 1, \quad (17)$$

for some $p, q \in R$. Write $q = (x+b)q_1 + r$, where $q_1 \in R$, $r \in A$, then by (15), $aq = a(x+b)q_1 + ar = (x+c)kq_1 + ar$, so

$$1 = (x+c)(p+kq_1) + ar.$$

A comparison of degrees shows that $p+kq_1 = 0$, $ar = 1$, which contradicts the fact that a is a non-unit. Thus $aR \cap (x+c)R$ is not principal, even though it is stably free. ■

As an example consider the Weyl algebra; this may be written $A[x; 1, \cdot]$, where $A = k[y]$ and $\cdot$ is d/dy . Here y is a non-unit and x, y are comaximal, hence we have

COROLLARY 10.7. $A_1(k)$ is not a 2-Hermite ring. ■

Explicitly this means that $R = A_1(k)$ contains a non-principal right ideal a such that $a \oplus R \cong R^2$.

Exercises 0.10

1. Let $R = K[x; \alpha, \delta]$ be a skew polynomial ring over a field K . Show that K may be regarded as a right R -module by letting each $a \in K$ correspond to right multiplication by a and letting x correspond to the action by δ . When is this representation faithful?
2. Let $R = K[x; \alpha, \delta]$ be a skew polynomial ring. If $\alpha\delta = \delta\alpha$, show that α may be extended to an endomorphism of R by taking $x^\alpha = x$; what is the value of x^δ in this case?
3. Let $R = K[x; \alpha, \delta]$ be a skew polynomial ring, where K is an integral domain and α an endomorphism such that $K^\alpha a \cap K^\alpha = 0$ for some $a \in K^\times$. Show that R is not left Ore. If K is a field, show that R is left Ore if and only if α is an automorphism.
4. Let K be a ring and α an injective endomorphism of K . Show that K can be embedded in a ring $K^{[\alpha]}$ such that α extends to an automorphism of $K^{[\alpha]}$ and that the least such subring is unique up to K -ring isomorphism. If, moreover, K is a field and $R = K(x; \alpha)$, then $K^{[\alpha]} \cong \bigcup x^m K x^{-m}$ (cf. 8.7).

5. Show that in a skew polynomial ring over a right Ore domain, the set of monic polynomials is a right Ore set.

6. Let R be an integral domain with an injective endomorphism α . If S is a right Ore set admitting α , show that S is also a right Ore set for $R[x; \alpha]$ and that $R_S[x; \alpha] \cong R[x; \alpha]_S$.

7*. (D. A. Jordan) Let k be a field, $K = k(x_i | i \in \mathbb{Z})$, $E = k(x_i | i > 0)$ and α the automorphism $x_i \mapsto x_{i+1}$ of K . Let S be the set of all monic polynomials in t over E , put $A = K[t]_S$ and extend α to A by the rule $t^\alpha = t$. Show that $A[t; \alpha]$ is right Noetherian and left Ore but not left Noetherian. (Hint: Use Ex. 0.8.6 for the last part.)

8. (Ore [32]) Let R be a skew polynomial ring over a field K and let f, g be polynomials of degrees m, n , respectively. Denote by K_0 the centralizer in K of f , and assume that the dimension of K as right K_0 -space is r . Show that $\text{Hom}_R(R/fR, R/gR)$ is a right K_0 -space of dimension at most rm . (Hint: Use (2) in 0.7.)

9*. (Lesieur) Let R be a right Noetherian domain and α an endomorphism of R . Show that $R[x; \alpha]$ is right Noetherian if and only if, for any sequence of right ideals a_i such that $a_i^\alpha \subseteq a_{i+1}$, there exists n_0 such that $a_{n+1} = a_n^\alpha R$ for all $n \geq n_0$.

10. (E. L. Ince) In the ring of differential polynomials $R = k(x)[D; 1, \cdot]$ show that $P = D^2 - 2x^{-2}$ and $Q = D^3 - D \cdot 3x^{-2} + 3x^{-3}$ commute but cannot be written as polynomials in the same element of R . (Hint: Verify that $P^3 = Q^2$ and note that P, Q are obtained by conjugating D^2, D^3 , respectively, by $D + x^{-1}$ in the field of fractions of R .)

11. Let $A = k[t]$ and $\alpha: f(t) \mapsto f(t+1)$. Show that for suitable elements $a, c \in R = A[x, x^{-1}; \alpha]$, a and c are right comaximal but $ca^\alpha \notin aR$; deduce that R is not 2-Hermite.

12. Let K be a non-commutative field and $R = K[x, y]$ the polynomial ring in two central indeterminates. If $[a, b] = ab - ba = c \neq 0$, verify that $[x+a, y+b] = c$; deduce that $(x+a)R \cap (y+b)R$ is isomorphic to a stably free but non-free right ideal of R .

13. Show that if R is weakly finite, then so is $R[t]$ and $R[[t]]$.

14. (Bergman and Dicks [78]) Let $\phi: R \rightarrow S$ be a k -algebra homomorphism. The multiplication map $\mu: S \otimes_R S \rightarrow S$ is given by $x \otimes y \mapsto xy$ ($x, y \in S$), and its kernel is denoted by $\Omega_{S/R}$, while the universal derivation of S relative to R , $d: S \rightarrow \Omega_{S/R}$ is defined by $s \mapsto s \otimes 1 - 1 \otimes s$ (Eilenberg). Show that there exists an exact sequence:

$$0 \longrightarrow \text{Tor}_1^R(S, S) \longrightarrow S \otimes_R \Omega_{S/R} \xrightarrow{d\phi} \Omega_{S/k} \longrightarrow \Omega_{S/R} \longrightarrow 0$$

the map $d\phi$ is called the *derivative* of ϕ , relative to the category of k -algebras). Show further that the following are equivalent:

(a) $d\phi$ is surjective, (b) $\Omega_{S/R} = 0$, (c) $S \otimes_R S \rightarrow S$ is surjective and (d) ϕ is an isomorphism (in the category of k -algebras).

0.11 Free associative algebras and tensor rings

Much of the subsequent discussion will centre on free algebras and related rings; in this section we shall describe them briefly and list some of their simpler properties.

Free algebras are most easily defined by the universal mapping property: given a commutative ring k and a set X , there exists a k -algebra $F = k\langle X \rangle$ and a map $i: X \rightarrow F$ such that any map $\theta: X \rightarrow B$ into another k -algebra B can be factored uniquely by i . Thus there exists a unique homomorphism

$$\begin{array}{ccc} X & \xrightarrow{i} & k\langle X \rangle = F \\ & \searrow \theta & \downarrow \theta' \\ & & B \end{array}$$

$\theta': F \rightarrow B$ such that the accompanying diagram commutes. This algebra F , clearly unique up to isomorphism, is easily proved to exist by general considerations (cf. Cohn [81], p. 170). However, one of its characteristic features is the very simple normal form for its elements, which is easily described and can also be used to prove its existence. To avoid trivialities we shall usually assume that k is not the zero ring and that X is non-empty.

Consider first the free monoid on a set X , X^* say. Its elements can be expressed as products

$$x_1 x_2 \dots x_n \quad (x \in X, n \geq 0),$$

including the empty product to represent 1. Sometimes it is more convenient to take X indexed: $X = \{x_i\}$, then the general element of X^* takes the form

$$x_I = x_{i_1} x_{i_2} \dots x_{i_n},$$

where $I = (i_1, \dots, i_n)$ runs over all finite sequences of suffixes (including the empty sequence). Now the free associative algebra $k\langle X \rangle$ on X over k may be described as the monoid algebra of X^* . Its elements are uniquely expressible in the form

$$f = \sum x_I a_I \quad (a_I \in k, \text{ almost all } 0), \quad (1)$$

By the *support* of f we understand the set of all x_I such that $a_I \neq 0$. We remark that since the elements of X do not commute, $k\langle X \rangle$ is non-commutative as soon as $|X| > 1$. When $|X| = 1$, say $X = \{x\}$, then $k\langle x \rangle$ becomes just the polynomial ring $k[x]$ in a single variable over k . One of our main objectives, in later chapters, will be to give a detailed description of $k\langle X \rangle$ in which each property of $k[x]$ has its counterpart.

We first check that $k\langle X \rangle$ indeed satisfies the characteristic property of free algebras. Given any k -algebra B and any map $\theta: x_i \rightarrow b_i$ of X into B , consider the map

$$\theta': \sum x_i a_i \mapsto \sum b_i a_i \quad (b_i = b_{i_1} \dots b_{i_n}, a_i \in k).$$

This is well-defined by the uniqueness of the normal form (1). It is easily seen to be a homomorphism and it extends the given map θ . Moreover, θ' is determined by its values on a generating set, hence for each θ there is just one θ' , as claimed.

In general we cannot easily recover k from $k\langle X \rangle$, but when k is a field (the case that will mainly occupy us later), k is determined by $k\langle X \rangle$ as the set of all units, together with 0. The generating set X is also not generally determined, but at least its cardinal is unique, and is called the *rank* of $k\langle X \rangle$. To prove that the rank is well-defined, take a fixed k -algebra homomorphism $\alpha: k\langle X \rangle \rightarrow k$. The (α, α) -derivations from $k\langle X \rangle$ to k form a k -module D say, and by the universal mapping property for derivations we have an isomorphism of k -modules

$$D \cong k^{(X)}. \quad (2)$$

Since k is commutative non-zero, it has IBN, and on comparing ranks in (2) we find that $|X| = \text{rk}_k D$. This shows the rank of $k\langle X \rangle$ to be independent of the choice of X . The set X is called a *free generating set* of $k\langle X \rangle$; more generally a subset S of a k -algebra is called *free* if the subalgebra generated by S is free, on S as free generating set.

The product $x_1 x_2 \dots x_n$ will be called a *monomial* of $k\langle X \rangle$ and $n = l(x_1 \dots x_n)$ its *length*; for a general element (1) of $k\langle X \rangle$ we define its *degree* as

$$d(f) = \max\{l(x_i) \mid a_i \neq 0\}, \quad (3)$$

and its *order* as

$$o(f) = \min\{l(x_i) \mid a_i \neq 0\}. \quad (4)$$

The degree- (and order-) functions so defined depend, of course, on the choice of the free generating set X ; more precisely, if Y is another free generating set, then we have a bijection

$$\alpha: X \rightarrow Y, \quad (5)$$

which extends to a unique automorphism of $k\langle X \rangle$. Thus the different (ordered) free generating sets of $k\langle X \rangle$ correspond to the automorphisms. The automorphism α given by (5) preserves the degree if and only if the expression of each element of Y in terms of X is linear, and it preserves the order if and only if the expression of each element of Y in terms of X has zero

constant term. The ideal of $k\langle X \rangle$ generated by X is called the *augmentation ideal*; X and Y generate the same augmentation ideal if and only if both define the same order-function. In that case the corresponding automorphism α is said to be *augmentation preserving*.

Instead of defining $k\langle X \rangle$ as a set of sums of products we may also begin with sums. It is convenient to do this in a more general context, where the coefficient ring is no longer assumed to be commutative.

Let D be any non-zero ring. Given a D -bimodule U , we define

$$U^{(0)} = D, \quad U^{(n)} = U \otimes_D U \otimes_D \dots \otimes_D U \quad (n \text{ factors}, n \geq 1),$$

as a D -bimodule. The direct sum

$$D\langle U \rangle = U^{(0)} \oplus U^{(1)} \oplus \dots \quad (6)$$

can be defined as a D -ring by means of the D -bilinear maps $U^{(r)} \times U^{(s)} \rightarrow U^{(r+s)}$. It is called the *tensor D -ring* on U ; since it is usually clear whether U is a D -bimodule or merely a set, no confusion with the earlier notation $k\langle X \rangle$ is likely. The form of the right-hand side of (6) shows that we have a graded ring and we can define a degree-function and an order-function on $D\langle U \rangle$ by assigning to each element f the degree of its highest resp. lowest terms, as in (3) and (4).

As an important example we have $U = (D \otimes_K D)^{(X)}$, where K is a subring of D and X is a set. The general element of U may then be written $\sum a_{xi} x b_{xi}$, where we have put x in place of $(1 \otimes 1)x$. U may be described as the free D -bimodule on X centralizing K , and one often writes $D_K\langle X \rangle$ in place of $D\langle U \rangle$. We remark that $D_K\langle X \rangle$ may also be defined as the D -ring generated by X with the defining relations

$$\alpha x = x \alpha \quad \text{for all } x \in X, \alpha \in K. \quad (7)$$

In terms of the order-function we can form the completion of $D\langle U \rangle$, denoted by $D\ll U \gg = \prod_{n=0}^{\infty} U^{(n)}$ and called the *power series D -ring* over U . Its elements are infinite series $a_0 + a_1 + \dots$ ($a_n \in U^{(n)}$). For a set X we shall denote the completion of $D_K\langle X \rangle$ by $D_K\ll X \gg$. In the case where X is infinite, it is sometimes advantageous to assign degrees to the elements of X in such a way that only finitely many are of degree $< n$, for any n .

Exercises 0.11

1. Let $\mathfrak{a}$ be the augmentation ideal in the free k -algebra $k\langle X \rangle$. Show that $\bigcap \mathfrak{a}^n = 0$, and find an ideal not contained in $\mathfrak{a}$ with the same property (cf. also 5.11).
2. Let kF be the group algebra (over a commutative field k) of the free group F of rank t . Show how to compute t in terms of kF .

3. Given a commutative field k and two disjoint sets X and Y , the *mixed free k -algebra* on X, Y, Y^{-1} is defined as the k -algebra $k\langle X, Y, Y^{-1} \rangle$ generated by X, Y and the inverses of elements of Y , which is universal for Y -inverting maps of $X \cup Y$ into k -algebras. Prove that the numbers of invertible and of non-invertible free generators in a mixed free algebra are independent of the choice of free generating set.

4. Show that every matrix over $k\langle X \rangle$ is stably associated to one of the form $A_0 + \sum A_i x_i$, where A_0, A_i are matrices over k and $x_i \in X$. Show the same for matrices over $k\langle X, X^{-1} \rangle$. (Hint: Note that $c + ab$ is stably associated to $\begin{pmatrix} c & a \\ -b & 1 \end{pmatrix}$, cf. also 5.8.)

5. Let $R = k\langle x_1, \dots, x_r \rangle$ be the free k -algebra, d_i the derivation mapping x_j to δ_{ij} and $D = \sum x_i d_i$ the derivation which is the identity on the x_i . Show that for any element $a = \sum a_n$ of R , where a_n is homogeneous of degree n , $a^D = \sum n a_n$. (This generalizes Euler's theorem on homogeneous functions.)

6. (Jategaonkar [69'], Koshevoi [70]) Let R be an integral domain which is not right Ore. If $xR \cap yR = 0$ ($x, y \neq 0$), show that x and y generate a free algebra (over $\mathbb{Z}$ or $\mathbb{F}_p$); deduce that an integral domain contains a free algebra on two free generators unless it is a left and right Ore domain.

Given a left but not right Ore domain, obtain an embedding of the free algebra of rank 2 in a field.

7. In $R = k\langle x, y \rangle$ show that the elements xy^r ($r = 0, 1, \dots$) form a free generating set of the subalgebra $k + xR$, and deduce that the free algebra of countable rank can be embedded in the free algebra of rank 2.

8. Find the centralizer of X in $\mathbb{Q}_Z\langle X \rangle$. What happens in the general case $D_K\langle X \rangle$? (cf. Dicks [77], p. 575.)

9. Let $R = k\langle x, x^{-1}, y \rangle$ be a mixed free algebra; show that the elements $y_i = x^{-i} y x^i$ ($i \in \mathbb{Z}$) form a free set. Denoting by A the subalgebra generated by these elements and by α the automorphism defined on A by $a \mapsto x^{-1} a x$, show that $R \cong A[x, x^{-1}; \alpha]$.

10. Let $R = k\langle x, y \rangle$ and write $[ab^{(1)}] = ab - ba$, $[ab^{(r+1)}] = [[ab^{(r)}]b]$. Show that the subalgebra S generated by the elements $[yx^{(r)}]$ ($r = 0, 1, \dots$) is freely generated by these elements. If δ is the derivation on S defined by $a \mapsto [ax]$, show that $R \cong S[x; 1, \delta]$.

Notes and comments on Chapter 0

Much of this material is part of the folklore and the citations given below are probably far from exhaustive. The formulation of Morita theory in 0.1 can easily be extracted from the general account (cf. Bass [68] or Cohn [76] and, for the form in which it appears here, Cohn [66']). It is all that is needed for projective free rings; nearly all the rings occurring in the sequel fall into this class.

The first thorough discussion of IBN occurs in Leavitt [57]; Shepherdson [51]

gives an example of a ring which is weakly 1- but not 2-finite. For a connected account of IBN, UGN and weak finiteness (showing that these classes are distinct) see Cohn [66]. Prop. 2.2 is proved for regular rings by Goodearl [79], Th. 18.3 and generally by Malcolmson [80]. Cor. 3.2, 3.3 (without reference to κ) occur in Cartan and Eilenberg [56], p. 13, cf. also Guazzone [62]; the proof of Th. 3.1 is modelled on these cases. Th. 3.5 is due to Kaplansky [58], who uses it to prove (*inter alia*) a commutative form of Th. 3.7: over a commutative semihereditary ring every projective module is a direct sum of finitely generated ideals. This result was generalized to left modules over a left semihereditary ring by Albrecht [61] and to left modules over a right semihereditary ring by Bass [64]. Both results are included in Th. 3.7, which is taken from Bergman [72].

The term Hermite ring (or H-ring) was introduced by Lissner [65]. The condition in Th. 4.1 occurs repeatedly, e.g. in Drogomizhs'ka [70] and Kazimirskii and Lunik [72], see also Lam [78] for the commutative case. Prop. 4.2 is due to Ojanguren and Sridharan [71]. The notion 'projective free' is defined (as ' p -free') in Cohn [66'], where Th. 4.6 is also proved.

The projective module group $K_0(R)$ occurs in K -theory; it contains much of the 'stabilized' information about the category ${}_R\mathcal{P}$, though for best results one needs to take the affine structure into account; see an instructive study by Goodearl and Warfield [81]. The (unstabilized) facts about $\mathcal{P}(R)$ are well known, e.g. Th. 5.3 occurs in Bass [68], p. 90.

The notion of stable association and its connexion with the matrix of definition of a module is implicit in Fitting's work [36]; in the form given here it is developed from the factorization theory in Cohn [63], see also Cohn [82]. The concepts of idealizer and eigenring seem to have been first used by Ore [32] (though already implicit in Levitzki's work) in his papers on the formal theory of differential equations. Prop. 7.1 (for the case $P = R$) is due to Fitting [35], and special cases of Th. 7.3 are well known (see e.g. Amitsur [58]), they are stated in this form by Cohn [70].

Since Ore's original construction (Cor. 8.7, cf. Ore [31]) there have been innumerable papers dealing with extensions, analogues for monoids, etc. We have not discussed the most general case, in which the Ore set may contain zero-divisors (cf. Ex. 0.8.7), but have concentrated on the cases used later in the book. For a comprehensive survey see Elizarov [69]. Cor. 8.10 is due to Goldie [58] (the customary proof actually gives Prop. 8.9). Th. 8.12 is due to Doss [48], who proves it by applying the Malcev conditions (itself a non-trivial verification). The proof in the 1st edition was defective, as L. A. Bokut has pointed out. The present proof is new.

The discussion of 0.9 is taken from Bergman [67]; Prop. 9.2 occurs in Cohn and Schofield [82] and Cor. 9.5 in Gentile [60]. Skew polynomial rings were first systematically studied by Ore [33]. The proof of Prop. 10.3 is modelled on the brief proof in Sarges [76] of the Hilbert basis theorem, and Prop. 10.4 is due to Curtis [52]; the analogue for power series does not hold, see Kerr [82]. Prop. 10.6 is taken from Stafford [85].

Most of 0.11 is folklore, cf. Cohn [69], although the tensor D -rings have only been considered recently, cf. Cohn [74].

1. Firs, semifirs and n -firs

Nearly all the rings discussed in this book have the property that all their (one-sided) ideals are free as left resp. right modules, of unique rank. This leads to the notion of a *free ideal ring*, or *fir*, which forms the subject of this chapter. Firs may be regarded as generalizing the notion of a principal ideal domain.

If one only demands that all finitely generated left ideals, or all n -generated left ideals, shall be free of unique rank, one obtains the wider class of *semifirs* and *n -firs*, respectively. These notions are automatically left-right symmetric and arise naturally if one postulates that certain dependence relations should be trivializable. A number of equivalent conditions, and the relation to the commutative case, is discussed in 1.1, while firs and their basic properties occupy 1.2. In 1.3 we discuss the special case that arises when we insist on 'trivializing' our relations by a subgroup of the general linear group. This will be of importance later when we come to meet firs (such as free associative algebras) in which every invertible matrix is a product of elementary and diagonal matrices.

From the homological point of view, firs are rather trivial. They are hereditary rings in which every projective is free (of unique rank). We can describe not merely firs, but also matrix rings over firs in this way, and in fact give a categorically invariant description of these classes. This is done in 1.4, with a description of flat modules over semifirs. In 1.5 we establish the link with weak Bezout rings (they are just the 2-firs) and consider what happens when fractions are introduced. Finally, 1.6 is devoted to inertia, useful in studying extensions of firs.

1.1 Dependence relations

We begin with a study of conditions under which every relation of the form

$$x \cdot y = x_1 y_1 + \dots + x_n y_n = 0 \quad (1)$$

is in some sense trivial. To be precise, let us call the relation (1) in a ring R *trivial* if for each $i = 1, \dots, n$, either $x_i = 0$ or $y_i = 0$. Of course, every non-zero ring R has non-trivial relations (1), e.g. $x = (a, ab)$, $y = (bc, -c)^T$, for any $a, b, c \in R$, is a relation, non-trivial unless $ab = bc = 0$. Let us say that the

n -term relation (1) is *trivialized* by the invertible matrix P if the relation $xP^{-1}.Py = 0$ is trivial; it is called *trivializable* if an invertible matrix trivializing it exists. We shall be concerned with rings in which every relation (1) is trivializable. Some equivalent conditions for this to happen are given in our first result.

More generally, we may take $x_1, \dots, x_n$ in (1) to lie in ${}^rR$ and $y_1, \dots, y_n$ in R^s , then the x 's may be taken to form the columns of an $r \times n$ matrix X and the y 's form the rows of an $n \times s$ matrix Y . Then (1) is an expression of the matrix equation $XY = 0$, and trivial, trivializable, etc., may be defined as before.

THEOREM 1.1. *Let R be a non-zero ring and n a positive integer. Then the following conditions are equivalent:*

- (a) every m -term relation $\sum x_i y_i = 0$ ($m \leq n$) holding in R can be trivialized by an invertible $m \times m$ matrix,
- (b) given elements $y_1, \dots, y_m \in R$ ($m \leq n$) which are left linearly dependent, there exist $m \times m$ matrices P, Q over R such that $PQ = I$ and the vector $Q(y_1, \dots, y_m)^T$ has zero as its first component,
- (c) any left ideal of R generated by $m \leq n$ left linearly dependent elements has a family of fewer than m generators,
- (d) any left ideal of R on at most n generators is free, of unique rank,
- (e) if F is a free left R -module, then any n -generated submodule of F is free, and R^n has unique rank,
- (f) if ϕ is a map of R^m ($m \leq n$) into a free left R -module, then there exist an integer $r \leq m$ and isomorphisms μ, ν such that the following diagram is commutative:

$$\begin{array}{ccccc}
 \ker \phi & \longrightarrow & R^m & \longrightarrow & \text{im } \phi \\
 \downarrow \mu|_{\ker \phi} & & \downarrow \mu & & \downarrow \nu \\
 R^r & \longrightarrow & R^r \oplus R^s & \longrightarrow & R^s
 \end{array} \tag{2}$$

where $r + s = m$ and the maps in the bottom line are the natural inclusion and projection,

- (g) for any matrix $Y \in {}^mR^s$ ($m \leq n$) there exists an invertible matrix P such that the non-zero rows of PY are left linearly independent,
- (h) any matrix relation $XY = 0$, where $X \in {}^rR^m$, $Y \in {}^mR^s$ ($m \leq n$) can be trivialized by an invertible $m \times m$ matrix.

(a⁰)–(h⁰) The left–right duals of (a)–(h).

The existence of a commutative diagram (2) with isomorphisms μ, ν is

expressed by saying that the two rows of the diagram are *isomorphic*; in that case $\mu|_{\ker \phi}$ is clearly also an isomorphism.

Proof. We shall prove $(a) \Rightarrow (b) \Rightarrow \dots \Rightarrow (h) \Rightarrow (a)$. Since (a) is left-right symmetric, the conditions must then be equivalent to their left-right duals. The implication $(a) \Rightarrow (b)$ is evident.

$(b) \Rightarrow (c)$. Let $\mathfrak{a}$ be a left ideal, generated by $y_1, \dots, y_m$ and suppose that the y_i are left linearly dependent. By (b) there exist $P, Q \in R_m$ such that $PQ = I$ and if $y' = Qy$, where $y = (y_1, \dots, y_m)^T$, then $y'_1 = 0$. Clearly $y'_1, \dots, y'_m \in \mathfrak{a}$ and since $y = PQy = Py'$, $\mathfrak{a}$ is generated by $y'_2, \dots, y'_m$.

$(c) \Rightarrow (d)$. If $\mathfrak{a}$ is an n -generated left ideal, (c) allows us to reduce the number of generators until we get a linearly independent generating set; so $\mathfrak{a}$ will be free on at most n generators. Now let $m \leq n$ be the least integer such that $\mathfrak{a} \cong R^m$ and assume that $\mathfrak{a}$ does not have unique rank, say $\mathfrak{a} \cong R^m \cong R^{m+k}$ ($k > 0$). Then $\mathfrak{a}$ has a surjective endomorphism ϕ with kernel R^k . If we take a set of m generators of $\mathfrak{a}$, their images under ϕ will again generate $\mathfrak{a}$ but will be linearly dependent, because $\ker \phi \neq 0$. Hence by the previous argument $\mathfrak{a}$ will be free of rank $m' < m$, contradicting our choice of m .

$(d) \Rightarrow (e)$. The first assertion follows from Th. 0.3.1. Next we observe that R is an integral domain. For if $a \in R$, then $Ra \cong R/n$, where n is the left annihilator of a in R . Since $n \geq 1$, Ra is free and we have $R = n \oplus \mathfrak{a}$, where $\mathfrak{a} \cong Ra$. This shows n to be a homomorphic image of R , hence principal and so also free. By the uniqueness of the rank of R , either $n = 0$ or $\mathfrak{a} = 0$, so a is either left regular or zero, and this shows R to be an integral domain. If R is a left Ore domain, it clearly has invariant basis number; otherwise, by Prop. 0.8.9, it contains free left ideals of any finite rank, so by (d), R^n has unique rank.

$(e) \Rightarrow (f)$. Given ϕ as in (f), (e) tells us that $\text{im } \phi$ is free, say $\text{im } \phi \cong R^s$. Then the exact sequence

$$0 \rightarrow \ker \phi \rightarrow R^m \rightarrow \text{im } \phi \rightarrow 0 \quad (3)$$

splits and so $\ker \phi$ is a direct summand of R^m , hence also free, say $\ker \phi \cong R^r$. Thus (3) is isomorphic to the bottom line of (2), and since R^m (as direct summand of R^n) has unique rank, it follows that $r + s = m$.

$(f) \Rightarrow (g)$. Let $\phi: R^m \rightarrow R^s$ be right multiplication by Y . By (f) we have a commutative diagram

$$\begin{array}{ccc} R^m & \longrightarrow & R^s \\ \downarrow \mu & & \downarrow \nu \\ R^m & \longrightarrow & R^t \end{array}$$

where ν is injective, μ is bijective and the lower map is the projection onto the last t components. If we represent μ, ν by matrices P, Q , then Q is left regular, P is invertible and $PY = \begin{pmatrix} Q \\ 0 \end{pmatrix}$, so (g) holds.

Now (g) $\Rightarrow$ (h) $\Rightarrow$ (a) is clear. ■

A ring satisfying the equivalent conditions of Th. 1.1 is called an n -fir. The reason for the name will become clear when we come to discuss the case of a fir (= free ideal ring, cf. 1.2). Of the above conditions, we shall find (a) and (f) the most useful tools, while (b) is the easiest to verify. Conditions (c) and (d) are weaker in appearance than the others since they do not explicitly assume anything about the modules R^n .

Clearly an n -fir is also an n' -fir, for all $n' < n$. Consequently a theorem of the form 'If R is an n -fir, then R has the property $P(m)$ for all $m < n$ ' is equivalent to the statement 'Any n -fir has the property $P(n)$ '. We shall often state results in the latter simpler form, with the understanding that $P(m)$ holds for $m < n$ because R is also an m -fir.

We note that a 1-fir is simply an integral domain; by convention a 0-fir is understood to be a non-zero ring. At the other extreme we have a ring which is an n -fir for all n ; such a ring is called a *semifir*. From Th. 1.1(d) we obtain the following characterization of semifirs:

COROLLARY 1.2. *A ring R is a semifir if and only if R has invariant basis number and every finitely generated left ideal is free (equivalently, every finitely generated right ideal is free). ■*

It is clear that an n -fir either has invariant basis number or is of type (h, k) , where $h > n$ (cf. 0.2). Hence by (f) we obtain

COROLLARY 1.3. *An n -fir is weakly n -finite and a semifir is weakly finite. ■*

The trivialization procedure of Th. 1.1 may also be applied to matrix products in which there is merely a block of zeros:

LEMMA 1.4 (Partition lemma). *Let R be an n -fir and $A \in {}^r R^n, B \in {}^n R^s$ such that AB has an $r' \times s''$ block of zeros as shown:*

$$AB = \begin{pmatrix} C' & 0 \\ C & C'' \end{pmatrix} \begin{matrix} r' \\ r'' \\ s' & s'' \end{matrix}$$

where r', r'', s', s'' indicates the number of rows and columns, respectively. Then there exists $T \in \mathbf{GL}_n(R)$ such that

$$AT^{-1} = \begin{pmatrix} A_1 & 0 \\ A_3 & A_4 \end{pmatrix} \begin{matrix} r' \\ r'' \end{matrix}, \quad TB = \begin{pmatrix} B_1 & 0 \\ B_3 & B_4 \end{pmatrix} \begin{matrix} t' \\ t'' \end{matrix}. \quad (4)$$

Proof. Partition A, B as $A = (A', A'')^T$, $B = (B', B'')$, where A', A'' have r', r'' rows, respectively, and B', B'' have s', s'' columns, respectively. Then $A'B'' = 0$, hence by Th. 1.1, there exists $T \in \mathbf{GL}_n(R)$ such that the first t' rows in TB'' are 0 and all columns after the first t' in $A'T^{-1}$ are 0, so AT^{-1} , TB have the form shown in (4). ■

We shall often need the following rank formula:

PROPOSITION 1.5. *Let R be an n -fir and A, B any submodules of a free R -module such that $A \oplus B$ is n -generated. Then the exact sequence*

$$0 \rightarrow A \cap B \rightarrow A \oplus B \rightarrow A + B \rightarrow 0 \quad (5)$$

splits, all modules occurring are free and

$$\mathrm{rk}(A+B) + \mathrm{rk}(A \cap B) = \mathrm{rk} A + \mathrm{rk} B. \quad (6)$$

Proof. We map $A \oplus B$ onto $A+B$ by the rule $(a, b) \mapsto a-b$; it is clear that the kernel is $A \cap B$. Now apply Th. 1.1(f). ■

In a semifir the exact sequence (5) also shows the truth of

COROLLARY 1.6. *Let R be a semifir and F a free R -module. Then the intersection of two finitely generated submodules of F is finitely generated.* ■

This also follows from the fact that a semifir is evidently coherent (cf. Appendix 2, p. 554).

Let R be a semifir and M a left R -module, with a presentation

$$0 \rightarrow K \rightarrow F \rightarrow M \rightarrow 0,$$

where F is free and K is a submodule of F . If M is finitely presented, then F may be taken to be finitely generated and K is then finitely generated, hence free, so the characteristic of M is given by

$$\chi(M) = \mathrm{rk} F - \mathrm{rk} K.$$

If M is finitely generated but not finitely related, we put $\chi(M) = -\infty$, while for M not finitely generated we put $\chi(M) = \infty$. This defines the characteristic of M in all cases, although care is needed in adding characteristics, since both ∞ and $-\infty$ can occur.

Let us see what becomes of our definitions in the commutative case. More generally, let us take a left Ore domain R ; in R any two elements are left linearly dependent, therefore any free left ideal has rank at most 1. Hence if a left Ore domain is a 2-fir, Th. 1.1(c) shows that any two elements generate a principal left ideal. An integral domain with this property is called a *left Bezout domain*; *right Bezout domains* are defined similarly, and by a *Bezout domain* we understand a domain which is left and right Bezout. It is not hard to see the truth of

PROPOSITION 1.7. *For any ring R , the following conditions are equivalent:*

- (a) R is a semifir and a left Ore domain,
- (b) R is a 2-fir and a left Ore domain,
- (c) R is a left Bezout domain,
- (d) R is an integral domain in which every finitely generated left ideal is principal.

Proof. (a) $\Rightarrow$ (b) is clear and (b) $\Rightarrow$ (c) follows from the earlier remarks. To prove (c) $\Rightarrow$ (d), let a be generated by $x_1, \dots, x_n$, then by induction on n , $Rx_2 + \dots + Rx_n$ is principal, say Ry , hence $a = Rx_1 + \dots + Rx_n = Rx_1 + Ry = Rz$, as we had to show. Finally, to show that (d) $\Rightarrow$ (a), if (d) holds, R is certainly a semifir, and if $a, b \neq 0$ but $Ra \cap Rb = 0$, then $Ra + Rb$ is free and not principal, which contradicts (d). ■

The commutative case yields

COROLLARY 1.8. *Every commutative 2-fir is a semifir.* ■

Thus for left (or right) Ore domains our chain of conditions from 2-fir to semifir collapses to a single condition. By contrast there are, in the general case, for each n , n -firs that are not $(n+1)$ -firs, as examples in 2.11 will show.

We have seen in Cor. 0.3.4 that for a left Bezout domain R , every finitely generated submodule of R^n is free (of rank at most n). A corresponding result holds for finitely generated torsion free left R -modules over a *right Bezout domain*:

PROPOSITION 1.9. *Let R be a right Bezout domain. Then any finitely generated torsion free left R -module is free.*

Proof. Let M be torsion free and generated by $u_1, \dots, u_n$, where n is minimal. Suppose that $\sum a_i u_i = 0$, where not all the a_i vanish; then $\sum a_i R = dR \neq 0$ for some $d \in R$, say $a_i = da_{1i}$, and

$$\sum a_{1i} R = R. \quad (7)$$

Since a Bezout domain is Hermite, the row $(a_{11}, \dots, a_{1n})$ forms the first row of an invertible $n \times n$ matrix $A = (a_{ij})$ say. It follows that M is also generated by $v_1, \dots, v_n$, where $v_i = \sum a_{ij} u_j$. But $dv_1 = \sum da_{1i} u_i = \sum a_i u_i = 0$, and since M is torsion free, $v_1 = 0$, so M is generated by $v_2, \dots, v_n$ and this contradicts the minimality of n . ■

It follows that in the situation of Prop. 1.9, torsion free is the same as semi-free (cf. also Cor. 4.7, below).

Exercises 1.1

1. For each $n \geq 1$, determine which of the following are n -firs: $k, k[x], k[x, y]$ (k a commutative field), $\mathbb{Z}, \mathbb{Z}[x]$.

2. For which n is it true that every subring of an n -fir is an n -fir? Give an example of an integral domain which cannot be embedded in a semifir (cf. Ex. 2.11.9).

3. For any $n \geq 1$, show that a direct limit of n -firs is an n -fir. Deduce that a direct limit of semifirs is a semifir.

4. Is the inverse limit of a system of semifirs necessarily a semifir? (Hint: Note that any intersection may be written as an inverse limit.)

5. Let R be an n -fir and S a subring which is also a homomorphic image of R , under a homomorphism fixing S (i.e. S is a retract of R). Show that S is again an n -fir.

6*. (Bergman) Let R be an n -fir and G a finite group of R -linear automorphisms of R^n . Show that the set of fixed elements is a free submodule.

7. Let R be an n -fir and $x_1 y_1 + \dots + x_n y_n = 0$ ($x_i, y_i \in R$). Show that $\text{rk}(\sum x_i R) + \text{rk}(\sum R y_i) \leq n$. Does this remain true if $x_i \in {}^e R, y_i \in R^s$?

8. Let $R = \mathbb{Z} + x\mathbb{Q}[x]$ be the ring of all polynomials in an indeterminate x with rational coefficients and integral constant term. Show that R is a Bezout domain but not principal.

9. Show that for any n -fir R and any left R -module P , $P^m \cong R^m$ (where $1 \leq m \leq n$) implies $P \cong R$.

10⁰. A ring R is called a *metasemifir* if every finitely generated left ideal is free, but R does not have IBN. Is the notion so defined left-right symmetric? If a metasemifir

R is of type (h, k) , show that (i) R is a $(h-1)$ -fir, (ii) every finitely generated R -module is h -generated and (iii) every finitely generated free module can be freely generated by fewer than $h+k$ elements (cf. Ex. 0.2.1).

11*. (Bergman) Let R be a ring such that R^n has unique rank. If every n -term relation in R can be trivialized by a member of $\text{GL}_n(R)$, show that R is an n -fir. (Without the uniqueness of the rank of R^n it can be shown that R is weakly semihereditary, all finitely generated projective modules are free and if R has type (h, k) then $n = i + jk$, where $2h - 1 \leq i < h + k, j \geq 1$.)

12°. Let R be a semifir, K a subfield of R and α an automorphism of R such that for each $n \neq 0$, the fixed ring of α^n is K . Is $R[x, x^{-1}; \alpha]$ necessarily a semifir?

1.2 Firs and α -firs

A *free left ideal ring*, or *left fir* for short, is a ring in which all left ideals are free of unique rank, as left R -modules. By Cor. 1.2, such a ring is certainly a semifir and so has invariant basis number. In particular, it is always an integral domain. *Right firs* are defined correspondingly and a *fir* is a left and right fir. More generally, let α be an infinite cardinal, then by a *left α -fir* we understand a ring in which all α -generated left ideals are free, of unique rank, and similarly for right α -firs. Clearly this generalizes the definition of n -fir in 1.1, but unlike the latter, the above notion turns out not to be left-right symmetric (cf. the examples in 2.10, 3.4 and 8.8).

Our first observation is that in a fir, submodules of free modules are free:

THEOREM 2.1. (i) *In a left fir every submodule of a free left module is free.* (ii) *In a left α -fir every α -generated submodule of a free left module is free.*

This is an immediate consequence of Cor. 0.3.2 and Th. 0.3.1, respectively. ■

We remark that the result holds even for meta-firs (defined as in Ex. 2.3).

We next investigate the presence of chain conditions in firs. It is easy to see that a fir is not Noetherian except in the rather special case of a principal ideal domain. Nevertheless there is a chain condition which is satisfied by all firs, namely the ACC on n -generator left ideals, i.e. left ACC_n . However, we begin with the Ore case.

PROPOSITION 2.2. *For any ring R the following are equivalent:*

- R is a left fir and a left Ore domain,*
- R is a principal left ideal domain,*
- R is a left Noetherian left fir,*
- R is a left Bezout domain with left ACC_1 .*

Proof. (a) $\Rightarrow$ (b) follows because a left Ore domain cannot have a left ideal which is free of rank > 1 . (b) $\Rightarrow$ (c) is clear; (c) $\Rightarrow$ (a) follows by Cor. 0.8.10. Now it is clear that (b) $\Rightarrow$ (d) and (d) $\Rightarrow$ (b), because (d) implies ACC on finitely generated left ideals, so the ring is left Noetherian. ■

This result shows in particular that a commutative fir is just a principal ideal domain.

THEOREM 2.3. *Let R be a left $\aleph_0$ -fir. Then any finitely related left R -module satisfies ACC_n for all n , hence R has left ACC_n for all n .*

Proof. Suppose first that F is a free left R -module and suppose that there is an infinite strictly ascending chain of n -generator submodules

$$N_1 \subset N_2 \subset \dots, \quad (1)$$

then $\bigcup N_i$ is a submodule of F which is countably but not finitely generated, hence by Th. 2.1(ii) it is free of countable rank. Let $u_1, u_2, \dots$ be a basis and denote by P the submodule generated by $u_1, \dots, u_{n+1}$, then $P \subseteq N_{i_0}$ for some i_0 and P is a direct summand of the union, hence of N_{i_0} . But N_{i_0} is free of rank at most n , and so cannot have a direct summand of rank $n+1$, which is a contradiction. This proves the assertion for free left R -modules.

Now let $M = F/L$ be finitely related, where F is free and L is k -generated, say. Then for any ascending chain (1) in M we can write $N_i = F_i/L$. If in (1), each N_i is n -generated, then F_i is $(n+k)$ -generated; by the first part of the proof the sequence (F_i) becomes stationary, and hence so does (1). ■

COROLLARY 2.4. *A semifir with left ACC_n satisfies left ACC_n on free left modules.*

This is obvious for a left Ore domain R , because R will then be left Bezout, hence left principal. Otherwise any ascending chain of n -generated submodules of a free module is contained in a free module of countable rank, and this is isomorphic to a left ideal of R , by Prop. 0.8.9, and here ACC_n holds by hypothesis. ■

Let R be an integral domain with left and right ACC_1 and take any $c \in R^\times$. Then $cR \neq R$ if and only if c is a non-unit, and in that case by right ACC_1 there is a maximal principal right ideal p_1R such that $cR \subseteq p_1R \subset R$. This means that $c = p_1c_1$ and p_1 is an atom. Repeating the procedure on c_1 we see that unless it is a unit, we can write $c_1 = p_2c_2$, where p_2 is an atom. Continuing in this fashion, we get a strictly ascending sequence of principal

left ideals

$$Rc \subset Rc_1 \subset Rc_2 \subset \dots$$

which must terminate by left ACC_1 . It follows that every non-zero element of R is either a unit or a product of atoms. A ring with this property is said to be *atomic*, and what we have proved can be stated as

PROPOSITION 2.5. *Any integral domain with left and right ACC_1 is atomic. ■*

In particular, combining this result with Th. 2.3, we have

COROLLARY 2.6. *Any left and right $\aleph_0$ -fir is atomic. ■*

Th. 2.3 can also be used to factorize matrices over $\aleph_0$ -firs provided they have no zero-divisors as factors. This will be done in Ch. 3 in a more general context, where this rather cumbersome condition on the factors is expressed in a different form.

Exercises 1.2

1. Give a direct proof that every principal left ideal domain is a left fir.
2. Show that an $\aleph_0$ -fir which is also a left Ore domain is a principal left ideal domain.
3. A non-zero ring without IBN in which all left ideals are free is called a left *metafir*. Which results of this section carry over to left metafirs?
4. In a (2-sided) fir, let a be a 2-sided ideal and b a left ideal. Examine the possible relations between $\text{rk}(a+b)$ and $\text{rk } b$.
5. Show that a semifir is a left fir if and only if it is left hereditary. (Hint: Use Th. 0.3.7, cf. 1.4.)
6. If R is weakly semihereditary and $M_1 \subseteq M_2 \subseteq \dots$ is an ascending chain of finitely generated modules whose union M is projective, show that there is a cofinal chain $P_1 \subseteq P_2 \subseteq \dots$ such that each P_i is a direct summand of M . (The P_i and M_j are *cofinal* in M if each M_j is contained in some P_j and each P_i is contained in some M_j . (Hint: Use Th. 0.3.7.)
- 7*. (Schofield) Let R be a left $\aleph_0$ -hereditary ring such that in any infinite sequence $P_1, P_2, \dots$ of non-zero finitely generated projective left R -modules the number of generators of $P_1 \oplus \dots \oplus P_n$ is unbounded as $n \rightarrow \infty$. Show that any projective left R -module satisfies ACC_n for all $n \geq 1$.

8°. Investigate join-irreducible (or meet-irreducible) left ideals in *firs*.

9°. If all maximal left ideals in a semifir R are finitely generated, is R necessarily a left fir?

1.3 Strong G_n -rings

In 1.1 we achieved the triviality of dependence relations by using invertible matrices. Such matrices are most tractable when they can be expressed as products of elementary and diagonal matrices, as is the case over a field, or over the ring $\mathbb{Z}$ of integers. However, over a general ring not every invertible matrix is necessarily a product of elementary and diagonal matrices, and it is reasonable to consider relations that are trivialized by products of elementary matrices or by matrices in some subgroup of $\text{GL}_n(R)$.

Let R be a ring. For each integer n there are $n+1$ natural embeddings of $\text{GL}_n(R)$ in $\text{GL}_{n+1}(R)$, obtained by inserting a row and column intersecting in 1 on the diagonal, with zeros elsewhere. By a *coherent family of matrix groups over R* , $\{G_n\}_{n \geq 1}$, we understand a family of subgroups G_n of $\text{GL}_n(R)$ such that G_{n+1} contains all the $n+1$ natural images of G_n .

Here are some examples of coherent families of matrix groups over R .

1. The general linear group $\{\text{GL}_n(R)\}$,
2. $\{\text{E}_n(R)\}$, where $\text{E}_n(R)$ is generated by all $n \times n$ elementary matrices (differing from the unit matrix in only one off-diagonal place), also called *transvections*,
3. $\{\text{D}_n(R)\}$, the invertible diagonal matrices,
4. $\{\text{GE}_n(R)\}$, where $\text{GE}_n(R)$ is generated by $\text{D}_n(R)$ and $\text{E}_n(R)$,
5. $\{\text{P}_n(R)\}$, the permutation matrices,
6. $\{\text{Tr}_n(R)\}$, the upper unitriangular matrices (having 1's on the main diagonal and 0's below it),
7. $\{\text{O}_n(R)\}$, the orthogonal matrices over a commutative ring R .

We note that in each of these examples G_n is a subfunctor of GL_n , for $n = 1, 2, \dots$; moreover, $\{G_n\}$ is a functor from rings to families of groups.

In order to study rings in which every relation can be trivialized by some member of a coherent family of groups $\{G_n\}$, we prove a theorem analogous to Th. 1.1.

THEOREM 3.1. *Let R be a non-zero ring, $\{G_m\}$ a coherent family of matrix groups over R and n a positive integer. Then the following conditions on R are equivalent:*

- (a) every m -term relation $\sum_1^m x_i y_i = 0$ ($m \leq n$) holding in R can be trivialized by a member of G_m ,

(b) if $y_1, \dots, y_m \in R$ are left linearly dependent ($m \leq n$), then there exists $Q \in G_m$ such that $Q(y_1, \dots, y_m)^T$ has at least one zero component,

(c) given $y_1, \dots, y_m \in R$ ($m \leq n$), there exists $Q \in G_m$ such that the non-zero components of $Q(y_1, \dots, y_m)^T$ are left linearly independent,

(d) R is an n -fir and $\mathbf{GL}_m(R) = G_m \mathbf{P}_m(R) \mathbf{Tr}_m(R) \mathbf{D}_m(R)$ ($1 \leq m \leq n$),

(e) if $y_1, \dots, y_m$ ($m \leq n$) are elements of a free left R -module F , then there exists $Q \in G_m$ such that the non-zero terms of $Q(y_1, \dots, y_m)^T$ are left linearly independent,

(f) if ϕ is a map of R^m ($m \leq n$) into a free left R -module, then there is an integer $r \leq m$ and $Q \in G_m$ corresponding to a map $\mu: R^r \rightarrow R^m$ such that the following diagram is commutative:

$$\begin{array}{ccccc} \ker \phi & \longrightarrow & R^m & \longrightarrow & \operatorname{im} \phi \\ \uparrow \mu' & & \uparrow \mu & & \uparrow \mu'' \\ R^r & \longrightarrow & R^m & \longrightarrow & R^s \end{array}$$

where $s = m - r$, the first of the lower maps is obtained by mapping the r components of R^r to certain r of the components of R^m , and the second lower map is obtained by deleting these r components, while μ', μ'' are induced by μ ,

(g) any relation $XY = 0$, where $X \in {}^r R^m$, $Y \in {}^m R^s$ ($m \leq n$) can be trivialized by some member of G_m .

(a⁰)–(g⁰) The left–right duals of (a)–(g).

Proof. Again (a) is left–right symmetric, so we need only prove (a) $\Rightarrow$ (b) $\Rightarrow$... $\Rightarrow$ (g) $\Rightarrow$ (a). Of these (a) $\Rightarrow$ (b) is clear and (b) $\Rightarrow$ (c) follows by an obvious induction, using the coherence.

(c) $\Rightarrow$ (d). By Th. 1.1(c), R is an n -fir. Now consider $A \in \mathbf{GL}_m(R)$, where $m \leq n$; we note that $A \in \mathbf{P}_m(R) \mathbf{Tr}_m(R) \mathbf{D}_m(R)$ if and only if there is a permutation matrix P such that PA is upper triangular, and we have to bring A to this form by left multiplication by a member of G_m . By (c) there exists $Q \in G_m$ such that the non-zero components of the first column of QA are left linearly independent. Since QA is invertible, its first column is unimodular, so it can only have a single non-zero entry which must be a unit, occurring in the i th row, say. Hence the matrix A' obtained from QA by deleting the first column and the i th row lies in $\mathbf{GL}_{m-1}(R)$. By induction on m there exist $Q' \in G_{m-1}$, $P' \in \mathbf{P}_{m-1}(R)$ such that $P'Q'A'$ is upper triangular. If Q_1 is the i th image of Q' in G_m and P_1 the i th image of P' , then after permuting the first and i th row in $P_1 Q_1 QA$ we obtain an upper triangular matrix, hence $A \in G_m \mathbf{P}_m(R) \mathbf{Tr}_m(R) \mathbf{D}_m(R)$.

(d) $\Rightarrow$ (e). Since $y_1, \dots, y_m$ can involve only finitely many components of

F , we can take F to be of finite rank, say $F = R^k$. Then we can form $Y = (y_1, \dots, y_m)^T$, an $m \times k$ matrix. By Th. 1.1(g) there exists $Q \in \text{GL}_m(R)$ such that the non-zero rows of QY are left linearly independent. Hence for suitable $P \in \text{P}_m(R)$ we have $PQY = (y'_1, \dots, y'_s, 0, \dots, 0)^T$. Now $(PQ)^{-1} = ABCD$, where $A \in G_m$, $B \in \text{P}_m(R)$, $C \in \text{Tr}_m(R)$, $D \in \text{D}_m(R)$, hence $A^{-1}Y = BCD \cdot PQY = BCD(y'_1, \dots, y'_s, 0, \dots, 0)^T$ and it is clear that the non-zero rows are left linearly independent.

(e) $\Rightarrow$ (f). Let $e_1, \dots, e_m$ be the standard basis of R^m (viewed as row vectors). By (e) there exists $Q \in G_m$ such that the non-zero rows of $Q(e_1\phi, \dots, e_m\phi)^T = Q\phi$ are left linearly independent. If there are s such rows, these s rows form a basis of $\text{im } \phi$, giving a natural isomorphism $\mu'' : R^s \rightarrow \text{im } \phi$. The remaining $m - s$ rows form a basis of $\ker \phi$, and so we obtain the desired commutative diagram.

Now (f) $\Rightarrow$ (g) follows as in the proof of Th. 1.1 and (g) $\Rightarrow$ (a) is clear. ■

A ring satisfying the above equivalent conditions will be called a *strong G_n -ring*; if R is a strong G_n -ring for all n , it will be called a *strong G -ring*. When $G_n = \text{GL}_n(R)$ or $\text{E}_n(R)$, etc., we shall omit (R) ; thus, for example, a strong GL_n -ring is just an n -fir, and a strong GL -ring is a semifir. On the other hand, the Euclidean algorithm shows that $\mathbb{Z}$, and more generally any Euclidean domain, is a strong E -ring. We shall meet many other examples of strong E -rings in the next chapter. Similar terminology applies for the other coherent families. Finally, by a GE_n -ring we understand a ring R for which $\text{GL}_m(R) = \text{GE}_m(R) (= \text{D}_m(R)\text{E}_m(R))$ for all $m \leq n$. A GE_n -ring for all n is called a GE -ring. We note

COROLLARY 3.2. *For any non-zero ring R and any integer $n \geq 1$ the following are equivalent:*

- (a) R is a strong E_n -ring,
- (b) R is a strong GE_n -ring,
- (c) R is an n -fir and a GE_n -ring.

Proof. (a) $\Rightarrow$ (b) is clear; if (b) holds, then by Th. 3.1(d), R is an n -fir and $\text{GL}_m(R) = \text{GE}_m(R)\text{P}_m(R)\text{Tr}_m(R)\text{D}_m(R)$, $1 \leq m \leq n$. It is clear that $\text{D}_m(R) \subseteq \text{GE}_m(R)$ and $\text{Tr}_m(R) \subseteq \text{GE}_m(R)$; further, $\text{P}_m(R) \subseteq \text{GE}_m(R)$, as follows from the equation

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Finally, when (c) holds, R is a strong E_n -ring by Th. 3.1(d). ■

Exercises 1.3

1. Show that for any ring R , $\mathbf{GE}_n(R)$ is the subgroup of $\mathbf{GL}_n(R)$ generated by $\mathbf{P}_n(R)$, $\mathbf{Tr}_n(R)$ and $\mathbf{D}_n(R)$. Verify that $\mathbf{D}_n$ normalizes $\mathbf{E}_n$ and $\mathbf{Tr}_n$.
2. Show that a strong $\mathbf{Tr}_n$ -ring is the same as an integral domain for $n = 1$ and the same as a field for $n > 1$.
3. If $\{G_n\}$ is a coherent family such that $G_n(R) \supseteq \mathbf{GE}_n(R)$, show that R is a strong G_n -ring if and only if it is an n -fir and $\mathbf{GL}_m(R) = G_m(R)$ for all $m \leq n$.
4. Let R be a commutative discrete valuation ring and $G_n(R)$ the group of invertible matrices whose entries below the main diagonal lie in the maximal ideal. Show that R is a strong G -ring. If R is a local ring, what can be said about the form of R when R is a strong G -ring?
- 5*. If $\mathbf{O}_n(R)$ denotes the orthogonal group (when R is commutative), show that the field of real numbers is a strong $\mathbf{O}$ -ring. For which other fields is this true?
6. (Lifting property) Show that an n -fir R is a strong $\mathbf{E}_n$ -ring whenever $R/\mathbf{J}(R)$ is a strong $\mathbf{E}_n$ -ring.

1.4 Homological properties of firs and semifirs

From the point of view of category theory, firs and semifirs are rather simple objects. On the one hand, these rings are hereditary and semihereditary, respectively, while on the other hand they are projective free. It turns out that these properties are enough to characterize them.

THEOREM 4.1. *A ring R is a semifir if and only if it is weakly semihereditary and projective free. In this case every projective (left or right) R -module is free, in particular, every projective left (or right) ideal is free.*

Proof. If R is a semifir, then it is left semihereditary, hence weakly semihereditary, and it is projective free, since every finitely generated projective module is a submodule of a free module, hence is itself free.

Conversely, assume that R is weakly semihereditary and projective free, and consider a relation $x_1 y_1 + \dots + x_n y_n = 0$ in R . Define the map $f: R^n \rightarrow R$ by $(a_1, \dots, a_n) \mapsto \sum a_i y_i$, then $x \in \ker f$, and since R is weakly semihereditary, there is a decomposition $R^n = P \oplus Q$ such that $Py = 0$ and $x \in P$. Since R is projective free, P and Q are free, with bases $u_1, \dots, u_r$ for P and $u_{r+1}, \dots, u_n$ for Q . The matrix $U = (u_1, \dots, u_n)^T$ is invertible and since $x \in P$, $x = x'_1 u_1 + \dots + x'_r u_r = (x'_1, \dots, x'_r, 0, \dots, 0)U$, while $Uy = (0, \dots, 0, u_{r+1}y, \dots, u_ny)^T$. Hence U trivializes the relation $xy = 0$, and this shows R indeed to be a

semifir. Now, by Th. 0.3.7 every projective R -module is a direct sum of finitely generated projective modules, which are free, because R is projective free. ■

For example, any local ring which is weakly semihereditary is a semifir, by Cor. 0.5.5. Moreover, by Th. 0.3.1 and its corollaries we have

COROLLARY 4.2. *A ring is a left fir if and only if it is left hereditary and projective free. ■*

COROLLARY 4.3. *Let α be an infinite cardinal. A ring is a left α -fir if and only if it is left α -hereditary and projective free. ■*

Since being weakly semihereditary is a categorical property, Th. 0.4.6 applied to Th. 4.1 yields the following Morita-invariant description of semifirs:

THEOREM 4.4. *For any ring R the following conditions are equivalent:*

- (a) R is a full matrix ring over a semifir,
- (b) R is Morita-equivalent to a semifir,
- (c) R is weakly semihereditary and projective trivial. ■

We can also give an explicit description of flat modules over semifirs.

PROPOSITION 4.5. *Let R be a semifir and U a left R -module. Then U is flat if and only if every finitely generated submodule of U is free.*

Proof. Clearly U is flat whenever all its finitely generated submodules are flat; this is so when the latter are free, so the condition is sufficient. To show that it is necessary, suppose that U is flat and let V be a finitely generated submodule. We have to show that V is free, so let us take a generating set $v_1, \dots, v_n$ of V , where n is minimal. The conclusion will follow if the v_i are linearly independent, so assume that there is a non-trivial dependence relation $xv = 0$, where $v = (v_1, \dots, v_n)^T \in {}^nV$ and $0 \neq x \in R^n$. By Th. 1.1(g^0) there exists $P \in GL_n(R)$ such that the non-zero entries of xP are right linearly independent. Replacing v by $P^{-1}v$ and x by xP , and renumbering the components if necessary, we may assume that for some r , $1 \leq r \leq n$, $x_1, \dots, x_r$ are right linearly independent while $x_j = 0$ for $j > r$. It follows that the map $f: {}^rR \rightarrow R$ given by $(y_1, \dots, y_r)^T \mapsto \sum_1^r x_i y_i$ is injective. Since U is flat, the induced map ${}^rR \otimes U \rightarrow R \otimes U \cong U$ is again injective, i.e. $(u_1, \dots, u_r)^T \mapsto \sum_1^r x_i u_i$ from ${}^rU$ to U . But $(v_1, \dots, v_r)^T$ is in the kernel, so $v_i = 0$ for $i = 1, \dots, r$, but

this contradicts the minimality of n , and it shows that V is free on $v_1, \dots, v_n$ as claimed. ■

It is clear that a semifir R is coherent, hence by Th. A.6, R^N is flat and by Prop. 4.5 we obtain a strengthening of Th. 1.1(e) (cf. the remarks after Th. 5.1.5 for the corresponding statement for n -firs):

COROLLARY 4.6. *Let R be a semifir, then every finitely generated submodule of R^N (qua left R -module) is free. ■*

We note that if in the proof of Prop. 4.5, R is a right Bezout domain, then $r = 1$ and we can weaken the hypothesis by assuming U to be torsion free instead of flat:

COROLLARY 4.7. *Over a right Bezout domain a left module is flat if and only if it is torsion free. ■*

Exercises 1.4

1. Let R be a semifir and S a ring such that $R_n \cong S_n$. Show that $R \cong S$. Does this still hold if R is merely assumed to be an n -fir?

2. Show that Cor. 4.7 fails for semifirs that are not right Bezout.

3. Prove an analogue of Th. 4.4 for firs; in particular show that every ring Morita-equivalent to a left fir is a full matrix ring over a left fir.

4. By an appropriate modification of Cor. 4.6 show that for any n -fir R , every n -generator submodule of R^N is free (cf. 5.1).

5. Given any ring R and modules $U, {}_R V$, $\text{Tor}_1(U, V)$ may be defined as the abelian group generated by the symbols (u, A, v) , where u is a row in U , v a column in V and A a matrix over R such that $uA = Av = 0$, subject to the relations

$$(u, A, v) + (u', A, v) = (u + u', A, v),$$

$$(u, A, v) + (u, A, v') = (u, A, v + v'),$$

$$(uA, B, v) = (u, AB, v), \quad (u, A, Bv) = (u, AB, v)$$

and

$$(u, A, v) + (u, A', v) = (u, A + A', v),$$

whenever the left-hand sides are defined (cf. Mac Lane [63], p. 150). What simplifications are possible if R is (i) a principal ideal domain and (ii) a fir?

6°. Investigate conditions on a ring R which imply that all finitely generated flat R -modules are (i) free and (ii) submodules of free modules.

7. Let R be a principal ideal domain. Show that a left R -module M is injective if and only if for any $c \in R^\times$ and $u \in M$ there exists $v \in M$ such that $u = cv$. Deduce that the field of fractions of R is injective as R -module.

8. Let R be a fir. Show that a left R -module M is injective if and only if for any left linearly independent family (e_i) in R and any family (u_i) in M (on the same suffix set) there exists $v \in M$ such that $u_i = e_i v$. (Hint: Apply Baer's criterion for injectivity.)

1.5 Further properties of n -firs

Of the chain of n -fir conditions, the case $n = 1$ is really too general to be of interest; thus a 1-fir is just an integral domain. On the other hand, 2-firs form an important class, e.g. in the commutative case they already comprise all semifirs.

By Th. 1.1(c), a 2-fir can be characterized as an integral domain in which any two left linearly dependent elements generate a principal left ideal: the case $m = 1$ of (c) asserts that R is an integral domain and $m = 2$ is the above ideal condition. We reformulate this condition as

THEOREM 5.1. *For any integral domain R the following are equivalent:*

- (a) R is a 2-fir,
 - (b) for any $a, b \in R^\times$ we have $Ra \cap Rb = Rm$ for some $m \in R$, while $Ra + Rb$ is principal if and only if $m \neq 0$,
 - (c) in the lattice $\text{Lat}_R(RR)$ of all left ideals of R , the principal left ideals of R containing a given $c \in R^\times$ form a sublattice, denoted by $\mathbf{L}(Rc, R)$,
 - (d) any two principal left ideals with non-zero intersection have a sum which is principal.
- (a⁰)–(d⁰) The right-left duals of (a)–(d).

Proof. (a) $\Rightarrow$ (b). Let $a, b \in R^\times$; by Prop. 1.5, $Ra + Rb$ and $Ra \cap Rb$ are free and have ranks adding up to $\text{rk}(Ra) + \text{rk}(Rb) = 2$; clearly $\text{rk}(Ra + Rb) \geq 1$, so $\text{rk}(Ra \cap Rb) \leq 1$, i.e. $Ra \cap Rb = Rm$ for some $m \in R$. Moreover, $m \neq 0$ if and only if $\text{rk}(Ra + Rb) = 1$, i.e. when $Ra + Rb$ is principal.

(b) $\Rightarrow$ (c) $\Rightarrow$ (d) is clear. To prove (d) $\Rightarrow$ (a), assume (d); then any two elements which are left linearly dependent generate a principal left ideal, hence R is a 2-fir by Th. 1.1(c). Finally the symmetry is clear from (a). ■

In a commutative 2-fir, i.e. a commutative Bezout domain, the lattice $\mathbf{L}(Rc, R)$ is a sublattice of the lattice of all fractional principal ideals, a lattice-ordered group and hence distributive (cf. Birkhoff [67], p. 293). In Ch.

4 we shall investigate 2-firs in which $L(Rc, R)$ is distributive; of course, $L(Rc, R)$ is always modular, as sublattice of $\text{Lat}_R(RR)$.

Let R be an integral domain. Two elements $a, b \in R^\times$ are said to have a *least common left multiple* (LCLM) m if $Ra \cap Rb = Rm \neq 0$; clearly m (if it exists) is unique up to a unit left factor. Suppose now that R is a 2-fir; from Th. 5.1(b) it is clear that any two left commensurable elements in R have an LCLM. In particular, this holds for left comaximal elements:

COROLLARY 5.2. *In a 2-fir any two non-zero left comaximal elements are left commensurable and hence have an LCLM.*

For if $Ra + Rb = R$, then $Ra \cap Rb = Rm \neq 0$ by Th. 5.1(b). ■

These results may be used to derive a normal form for fractions over a 2-fir.

PROPOSITION 5.3. *Let a, b be left commensurable elements of a 2-fir R and assume that R is a subring of a ring S in which every factor of b is invertible. Then the element $s = ab^{-1}$ of S can be written in the form $a'b^{-1}$, where a', b' are left comaximal in R , and if $s = a_1b_1^{-1}$, where $a_1, b_1 \in R$, then $(a_1, b_1)^T \in (a', b')^T R$. In particular, if a_1, b_1 are left comaximal, then $(a_1, b_1)^T$ is right associated to $(a', b')^T$.*

Proof. By Th. 5.1(d), $Ra + Rb = Rd$ for some $d \in R^\times$, say $a = a'd$, $b = b'd$. Then a', b' are left comaximal in R and b', d are invertible in S , hence $s = ab^{-1} = a'dd^{-1}b'^{-1} = a'b'^{-1}$. Further, $Ra' \cap Rb' = Rm$ for some non-zero $m = b_0a' = a_0b'$. Therefore $a_0 = b_0s$ and it is clear that $b_0R \cap a_0R = mR$. Thus if $s = a_1b_1^{-1}$, where $a_1, b_1 \in R$, then $a_0b_1 = b_0sb_1 = b_0a_1 \in a_0R \cap b_0R = mR$, hence for some $r \in R$, $a_0b_1 = b_0a_1 = mr = b_0a'r = a_0b'r$, so $(a_1, b_1)^T = (a', b')^T r$ as claimed. ■

We next show that the property of being a fir is preserved by localization.

PROPOSITION 5.4. *Let R be a right fir and T a right Ore set in R . Then the localization R_T is again a right fir.*

Proof. Let $\mathfrak{a}$ be a right ideal of R_T ; then $\mathfrak{a} \cap R$ is a right ideal of R , and hence is free, with a basis $(v_\beta)_{\beta < \alpha}$. It is clear that the v_β generate $\mathfrak{a}$ as right ideal of R_T ; for $v_\beta \in \mathfrak{a}$ and if $x \in \mathfrak{a}$, then $xd \in \mathfrak{a} \cap R$ for some $d \in T$, so $xd = \sum v_\beta a_\beta$ and hence $x = \sum v_\beta a_\beta d^{-1}$. We claim that the v_β are right linearly independent over R_T ; for if not, let $\sum v_\beta b_\beta = 0$ ($b_\beta \in R_T$) be a non-trivial relation, then only

finitely many of the b_β are non-zero. So we can bring the b 's to a common denominator, and multiplying up, we obtain a relation $\sum v_\beta c_\beta = 0$ ($c_\beta \in R$), and $c_\beta = 0$ if and only if $b_\beta = 0$. But this contradicts the linear independence of the v_β over R , hence they form a free generating set of $\mathfrak{a}$. Moreover, if $\mathfrak{a}$ is finitely generated, then it has a finite basis, and since every basis arises from one of $\mathfrak{a} \cap R$ over R , all have the same number of elements, so R_T is indeed a right fir. ■

It is easily seen that a corresponding result holds for α -firs.

When R is a left or right Bezout domain, we can embed R in a field of fractions. Later (in Ch. 7) we shall see how to embed any semifir in a field, but for the moment we observe that if R is a semifir, then the monoid $R^\times$ can be embedded in a group. This depends on the following theorem from Klein [69]. The proof consists in taking the Malcev conditions for the embeddability of a monoid (cf. e.g. Cohn [81], p. 268), expressing them in matrix form and so showing that they are satisfied; it will not be given here.

KLEIN'S THEOREM. *Let R be an integral domain such that for $n = 1, 2, \dots$ every nilpotent $n \times n$ matrix A over R satisfies $A^n = 0$. Then $R^\times$ is embeddable in a group.* ■

To apply this result to semifirs, we shall need

PROPOSITION 5.5. *If A is a nilpotent $n \times n$ matrix over an n -fir R , then a conjugate of A is strictly upper triangular, with zeros on the main diagonal, and in particular $A^n = 0$.*

Proof. Let m be the least integer such that $A^{m+1} = 0$ and write $B = A^m$, then $B \neq 0$ and $AB = 0$, thus the columns of A are right linearly dependent. Hence (by Th. 1.1(h)) there exists $U \in \text{GL}_n(R)$ such that the first column of AU is 0, and so the first column of $U^{-1}AU$ is also 0. Deleting the first row and column from $U^{-1}AU$ we obtain an $(n-1) \times (n-1)$ matrix which is again nilpotent; by induction it is conjugate to a strictly upper triangular matrix, hence so is A . ■

Applying Klein's theorem, we obtain

COROLLARY 5.6. *If R is a semifir, then $R^\times$ is embeddable in a group.* ■

In 7.11 a corresponding result will be obtained for 2-firs.

Exercises 1.5

1. Show that in a 2-fir, any finite set of pairwise left commensurable elements generates a principal left ideal.
2. Let R be a ring; if the polynomial ring $R[x]$ is a 2-fir, show that R must be a field.
3. Show that for any right invariant element c in an integral domain R (i.e. $Rc \subseteq cR$ and $c \neq 0$) the family of powers (c^n) is a right Ore set.
4. (Klein [69]) Show that any ring satisfying Klein's nilpotence condition is weakly finite. (Hint: If $AB = I$, show that $A(I - B'A')$ is nilpotent.)
5. (Johnson [67]) Let R be a 2-fir and L the set of all right large elements in R . Show that L is a right Ore set and that the localization R_L is again a 2-fir.
6. Prove the analogue of Prop. 5.4 for α -firs.
- 7^o. Prove the analogue of Prop. 5.4 for left firs; what about principal left ideal domains?

1.6 Inert extensions

When we come to discuss extensions of rings, we shall be particularly interested in those extensions that preserve factorizations. This leads to the notion of an inert extension, well known from number theory; we shall need several variants of this notion.

Definition 1. Let S be a ring and R a subring. An $m \times n$ matrix A over R is said to be *inert* in S , if for any factorization

$$A = PQ, \quad P \in {}^m S^r, \quad Q \in {}^r S^n, \quad r = \min(m, n),$$

there exists $U \in \text{GL}_r(S)$ such that $PU, U^{-1}Q$ have entries in R . If every non-zero element of R is inert in S , then R is said to be *1-inert* in S .

For a second, stronger, version we shall need to deal with row and column vectors from a ring. We shall generally denote the components of such a vector by a Latin suffix, thus a has the components a_i and a_λ the components $a_{\lambda i}$. The precise range will be indicated in brackets when it is not clear from the context. We also continue to use the notation ab for the product of a row a and a column b , thus $ab = \sum a_i b_i$.

Let $R \subseteq S$ be a pair of rings. Given $a \in S^n, b \in {}^n S$, the product $ab = \sum a_i b_i$ is said to lie *trivially* in R , if for each $i = 1, \dots, n$, either a_i and b_i lie in R or $a_i = 0$ or $b_i = 0$.

Definition 2. The subring R of S is said to be *totally n -inert* in S if for all $m \leq n$ and any families (a_λ) of rows in S^m and (b_μ) of columns in ${}^mS$ such that $a_\lambda b_\mu \in R$ for all λ, μ , there exists $U \in \text{GL}_m(S)$ such that on writing $a'_\lambda = a_\lambda U$, $b'_\mu = U^{-1}b_\mu$, each product $a'_\lambda b'_\mu$ lies trivially in R . If R is totally n -inert in S for all $n \geq 1$, we say that R is *totally inert* in S .

It is clear that 'totally 1-inert' implies '1-inert'; note that the term ' n -inert' has not been defined for $n > 1$.

To give an example, if we allow for a moment subrings not containing the 1 of the larger ring, an n -fir may be described as a ring in which $\{0\}$ is totally n -inert. More generally, we have the following relation with semifirs.

PROPOSITION 6.1. *Let R be a ring and X a totally inert subring of R . Then every finitely generated right ideal of R is the direct sum of a free right ideal and a right ideal generated by a subset of X .*

Proof. Let α be a finitely generated right ideal of R , generated by $e_1, \dots, e_n$, say, where n is minimal. Let $b_\lambda \in {}^nR$ be the family of all columns satisfying $(e_1, \dots, e_n)b_\lambda \in X$. By total inertia we can apply an invertible matrix so that all these products lie trivially in X . Keeping the same notation, we find that the e 's still generate α , and by the minimality of n , no e_i can vanish, so we may assume that $e_1, \dots, e_r \in X$, while $b_{\lambda j} = 0$ for $j > r$. Let α_1 be the right ideal generated by $e_1, \dots, e_r$ and α_2 generated by $e_{r+1}, \dots, e_n$, then

$$\alpha = \alpha_1 + \alpha_2, \quad (1)$$

and for any $b_1, \dots, b_n \in R$ such that $\sum e_i b_i = 0$ we must have $b_{r+1} = \dots = b_n = 0$, hence the sum in (1) is direct and α_2 is free on $e_{r+1}, \dots, e_n$, which proves the assertion. ■

Let us examine more closely the situation which arises in a totally inert subring. Suppose we have two families $a_\lambda \in R^n$ and $b_\mu \in {}^nR$, such that each product $a_\lambda b_\mu$ lies trivially in X . We can divide the range $\{1, 2, \dots, n\}$ into three parts N' , N'' and N_0 : $i \in N'$ if for some a_λ the i th coordinate is not in X , $i \in N''$ if for some b_μ the i th coordinate is not in X and $i \in N_0$ in all other cases. Then it is clear that $b_{\mu i} = 0$ for $i \in N'$, $a_{\lambda i} = 0$ for $i \in N''$ and $a_{\lambda i}, b_{\mu i} \in X$ for $i \in N_0$. Thus taking the families finite for simplicity, if the a_λ form the rows of a matrix A and the b_μ form the columns of a matrix B , then for a suitable ordering of the columns of A and the rows of B ,

$$A = (A' \quad 0 \quad A_0) \quad B = \begin{pmatrix} 0 \\ B'' \\ B_0 \end{pmatrix} \quad (2)$$

where A_0, B_0 have entries in X . Now the definition of total n -inertia shows the truth of

PROPOSITION 6.2. *Let R be a ring and X a totally n -inert subring. Given $A \in {}^nR^n$, $B \in {}^nR^s$ such that all the entries of AB lie in X , there exists $U \in \text{GL}_n(R)$ such that $AU, U^{-1}B$ have the respective forms shown in (2). ■*

To give an example, any field K (even skew) is totally inert in the polynomial ring $K[x]$. For, given any families of row and column vectors a_λ, b_μ over $K[x]$ such that $a_\lambda b_\mu \in K$, we may assume that $a_\lambda \neq 0$ for some λ . Fix such a λ , say $\lambda = 0$, and apply an invertible matrix which brings a_0 to the form $(d, 0, \dots, 0)$, where d is the highest common factor of the components of a_0 in $K[x]$. If the transformed columns are $b'_\mu = (b'_{\mu 1}, \dots, b'_{\mu n})^T$, then $db'_{\mu 1} \in K$ and $d \neq 0$, hence either $b'_{\mu 1} = 0$ for all μ , or $d, b'_{\mu 1} \in K$ and $a'_{\lambda 1} \in K$ for all λ . In either case $\sum_2^n a'_\lambda b'_{\mu i} \in K$ and an induction on n shows that K is totally inert in $K[x]$. Other examples will be encountered in Ch. 2 and Ch. 5.

Exercises 1.6

1. Prove the transitivity of inertia: if R is totally inert in S and S is totally inert in T , then R is totally inert in T .
2. Given an integral domain R and a central indeterminate t , show that R is totally 1-inert in $R[t]$.
3. Let K be a commutative field and R a subring which is a unique factorization domain. Show that R is totally 1-inert in K .
4. Let R be a commutative unique factorization domain and k its field of fractions. Show that $R[t]$ is 1-inert in $k[t]$. (This is essentially Gauss's lemma.)
5. If R is 1-inert in S , show that any atom in R remains one in S .
6. Let R be a semifir. Show that the $n \times n$ triangular matrices with non-zero diagonal elements are inert in R_n .
- 7°. Let R be an n -fir and S a totally n -inert subring. Is S again an n -fir?

Notes and comments on Chapter 1

The basic references for firs and semifirs are Cohn [64'] and, for the presentation used here, Bergman [67]. The partition Lemma 1.4 is taken from Cohn [82]; it is used in Ch. 7. n -firs arose out of the construction in Cohn [66], which leads to a general method of constructing n -firs that are not $(n+1)$ -firs, cf. Cohn [69'], Bergman [74'] and 2.11, below. α -firs were introduced by Bergman [67], who proved Cor. 2.4,

generalizing the fact (proved by Cohn [67]) that any left fir has left ACC_n . The fact that free modules over firs satisfy ACC_n can be viewed as a result in universal algebra (cf. B. Baumslag and G. Baumslag [71]).

The treatment of strong G_n -rings is taken from Bergman [67] and generalizes earlier results on strong GE_n -rings in Cohn [66', 69]. Th. 4.4 first occurs in Cohn [66']. We remark that the hypotheses of Th. 4.4 are symmetric and show incidentally that for a projective trivial ring, 'left semihereditary' is equivalent to 'right semihereditary'. For general rings this is no longer so, as an example in Chase [61] shows. The remainder of 1.4 was new in the first edition; for an application of Prop. 4.5 see Jensen [69]. Recently Prest [83] has shown that for any existentially complete prime ring R , every non-zero finitely generated projective module is isomorphic to R , thus such a ring is a metasemifir (of type (1, 1), cf. Ex. 1.10).

2-Firs were defined (under the name 'weak Bezout rings') and studied in Cohn [63]; the present (weaker) form of their definition is due (independently) to Bergman [67], Bowtell [67] (cf. Ex. 3.2.3), Williams [68] and, for right Bezout rings (i.e. rings in which every 2-generator ideal is principal), Lunik [70]. This observation is essentially Th. 5.1(a) $\Rightarrow$ first part of (b). Commutative Bezout domains are studied (under that name) by Jaffard [60] and Bourbaki [72]; the name is intended to convey that any two coprime elements a, b satisfy the 'Bezout identity' $au - bv = 1$.

Klein's theorem (as well as Cor. 5.6) occurs in Klein [69]. For generalizations of Cor. 5.6 see 7.11. The notion of inertia described in 1.6 is the natural generalization of the commutative concept introduced in Cohn [68]. Of course, the underlying idea already occurs in Gauss's lemma and the notion of 'inertial' primes in algebraic number theory.

A 'ringoid', i.e. a small preadditive category, may be regarded as a ring with several objects, and many of the results on general rings, appropriately modified, still hold for ringoids (cf. Mitchell [72]). By analogy, left free ideal ringoids ('left firoids') may be defined, and many of the results of Ch. 1 proved in this context. Thus Faizov [81] obtains analogues of Th. 1.1 and constructs firoids by a form of the weak algorithm (cf. Ch. 2). Wong [78] defines, for any small category $\mathcal{C}$ and any ring R , the category ring $R\mathcal{C}$ (analogous to the group ring). By a bridge category Wong understands the free category freely generated by an oriented graph with arrows A and inverses for a certain subset B of A . For the moment let us call a category a *delta* if all its morphisms are isomorphisms and its only endomorphisms are the identities. Now Wong [78] proves that if R is any ring and $\mathcal{C}$ a small category, then $R\mathcal{C}$ is a firoid if and only if either (i) $\mathcal{C}$ is a delta and R is a fir or (ii) $\mathcal{C}$ is a bridge category and R is a field. Thus for any free monoid $M \neq 1$, the monoid ring RM is a fir if and only if R is a field. In a similar vein, Dicks and Menal [79] have shown that for any non-trivial group G and any ring R , RG is a semifir if and only if R is a field and G is locally free. More generally, Menal [81] shows that for any monoid $M \neq 1$, finitely generated over its unit group, and any ring R , RM is a semifir if and only if R is a field and M is the free product of a free monoid and a locally free group. It will follow from later results (Ch. 2 and 7) that the group algebra of a free group is a fir.

2. The weak algorithm

Just as the Euclidean algorithm singles out some important classes of principal ideal domains, so there are some special classes of firs that may be described by a 'weak' algorithm. This generalizes the Euclidean algorithm, to which it reduces in the commutative case.

We begin in 2.1 by recalling the division algorithm, and showing how it can actually be used to characterize skew polynomial rings over fields. Corresponding to n -firs we have the n -term weak algorithm defined (for any filtered rings) in 2.2. To enable us to study this algorithm efficiently we look at graded rings in 2.3; it turns out that the n -term weak algorithm on a filtered ring can be described entirely in terms of the associated graded ring. The weak algorithm is used in 2.4 to characterize free algebras, and the Hilbert series, defined for any filtered ring in 2.5, provides extra information when the weak algorithm holds. In 2.6 we give Bergman's description of rings with a weak algorithm, while 2.7 and 8 are devoted to a closer study of the 2-term weak algorithm, using non-commutative continuants, and include a presentation of the GL_2 for such rings.

The non-commuting power series rings can be characterized in analogous fashion by an 'inverse weak algorithm'. This is in many ways simpler than the weak algorithm; in fact the commutative analogue is so simple that it has never been thought worthy of study. However, in the general case it provides some very explicit information on the relations in free power series rings, which are explored in 2.9.

Algorithms with a transfinite range, described in 2.10, are useful for constructing one-sided counter-examples. Finally, 2.11 gives a method allowing one in many cases to read off from the presentation of a ring whether the n -term weak algorithm holds. This enables one to construct quite naturally n -firs that are not $(n+1)$ -firs.

2.1 The division algorithm

We recall briefly the division algorithm used in deriving the familiar Euclidean algorithm. A ring R is said to satisfy the *division algorithm* relative to a function θ on R taking values in a well-ordered set, if

DA For any $a, b \in R, b \neq 0$, there exist $q, r \in R$ such that

$$a = bq + r, \quad \theta(r) < \theta(b). \quad (1)$$

We note that if b is chosen in R so as to minimize $\theta(b)$, we must have $r = 0$ in (1), so that $a = bq$. Since this holds for all $a \in R$, b must then have a right inverse; this also shows that $\theta(0)$ has the least value, usually taken to be 0.

Strictly speaking DA should be called the *right* division algorithm, since it is not left-right symmetric, but we shall usually omit the qualifying adjective.

It is often convenient to replace DA by the following equivalent condition which demands less:

A For any $a, b \in R$ such that $b \neq 0$ and $\theta(a) \geq \theta(b)$, there exists $c \in R$ such that

$$\theta(a - bc) < \theta(a). \quad (2)$$

If DA holds, then A follows by taking $c = q$. Conversely, if A holds and $a, b \in R, b \neq 0$, choose $q \in R$ such that $\theta(a - bq)$ takes its least value. If $\theta(a - bq) \geq \theta(b)$, then by A there exists $c \in R$ such that $\theta(a - bq - bc) < \theta(a - bq)$, which contradicts the choice of q , hence $\theta(a - bq) < \theta(b)$ and we obtain DA on writing $r = a - bq$.

We note that any condition such as DA or A is relative to a function θ , but to investigate the existence of an algorithm we need not presuppose that θ is given. For any ring R and subsets S, T of R let us put $S + T = \{s + t \mid s \in S, t \in T\}$ and define the *derived set* of S as

$$S' = \{x \in R \mid S + xR = R\}. \quad (3)$$

Thus S' is the set of quotients for which we can always perform the division with a remainder in S . Now define a sequence of subsets $\{S_n\}$ of R recursively by putting

$$S_0 = \{0\}, \quad S_{n+1} = S_n \cup S'_n. \quad (4)$$

For example, S_1 consists of 0 and all right invertible elements of R . These sets form an ascending chain

$$\{0\} = S_0 \subseteq S_1 \subseteq \dots \quad (5)$$

If their union is the whole ring,

$$\bigcup S_n = R, \quad (6)$$

we shall say that R is *Euclidean* and define a function $\phi: R \rightarrow \mathbb{N}$ by

$$\phi(x) = \min\{n \mid x \in S_n\}. \quad (7)$$

Thus S_n consists of the $x \in R$ such that $\phi(x) \leq n$. The concepts of Euclidean ring and division algorithm are related in the following way:

THEOREM 1.1. *If R is a Euclidean ring, then R satisfies the division algorithm relative to the $\mathbf{N}$ -valued function ϕ on R defined by (7). Here*

$$\phi(0) = 0, \quad \phi(1) = 1, \quad (8)$$

and R is an integral domain if and only if

$$\phi(ab) \geq \phi(a) \quad \text{for all } a, b \in R^\times. \quad (9)$$

Conversely, if R satisfies the division algorithm relative to a function $\theta: R \rightarrow \mathbf{N}$, then R is Euclidean and

$$\phi(x) \leq \theta(x) \quad \text{for all } x \in R.$$

Proof. Suppose that R is Euclidean; then it is clear that (8) holds. If $a \in R$, $b \in R^\times$ are given, say $\phi(b) = n > 0$, then $S_{n-1} + bR = R$, so there exist $r \in S_{n-1}$, $c \in R$ such that $r + bc = a$, and hence $\phi(a - bc) = \phi(r) \leq n - 1 < \phi(b)$. This proves DA. If R is an integral domain, then for $a, b \in R^\times$, we have $ab \neq 0$, so $\phi(ab) = n > 0$, hence $R = S_{n-1} + abR \subseteq S_{n-1} + aR$, therefore $a \in S_n$ and $\phi(a) \leq n = \phi(ab)$, i.e. (9). Of course, when (9) holds, then $\phi(ab) > 0$ for $a, b \in R^\times$, and by (8), $1 \neq 0$, so R is then an integral domain.

Conversely, assume that R satisfies the division algorithm relative to some function θ . For $n \in \mathbf{N}$, put $T_n = \{x \in R \mid \theta(x) \leq n\}$; we shall show by induction on n that $T_n \subseteq S_n$ for all n . This is clear for $n = 0$, so assume that $n \geq 0$, $T_n \subseteq S_n$ and consider $b \in R$ such that $\theta(b) = n + 1$. By DA, for each $a \in R$ there exists $q \in R$ such that $\theta(a - bq) < \theta(b) = n + 1$, so $a - bq \in T_n$ and $a \in T_n + bR \subseteq S_n + bR$. This holds for all $a \in R$, so $S_n + bR = R$ and $b \in S_{n+1}$. Hence $T_{n+1} \subseteq S_{n+1}$ and by induction, $T_n \subseteq S_n$ for all n . Since $\bigcup T_n = R$, we see that (6) holds, so R is Euclidean, and $\phi(a) \leq \theta(a)$ for all $a \in R$. ■

If a ring satisfies the division algorithm relative to some $\mathbf{N}$ -valued function θ on R , then by Th. 1.1, ϕ (given by (7)) is defined and is the smallest $\mathbf{N}$ -valued function for which the algorithm holds. Later, in 2.8 we shall see that for any $a \in R$, $\theta(a)$ is an upper bound for the number of steps in the Euclidean algorithm for any pair a, b , so in a sense ϕ gives the 'fastest' Euclidean algorithm.

Even when (6) fails, we can continue the sequence transfinitely, putting

$$S_{\alpha+1} = S_\alpha \cup S'_\alpha, \quad S_\lambda = \bigcup_{\alpha < \lambda} S_\alpha \quad \text{at a limit ordinal } \lambda. \quad (10)$$

If $S_\tau = R$ for some ordinal τ , we say that R is *transfinitely Euclidean* and

define an ordinal function ϕ on R by

$$\phi(x) = \min\{\alpha \mid x \in S_\alpha\}. \quad (11)$$

We note that if $x \in \bigcup_{\alpha < \lambda} S_\alpha$, then $x \in S_\alpha$ for some $\alpha < \lambda$, hence $\phi(x)$ is never a limit ordinal.

Clearly R is Euclidean if and only if we can take $\tau = \omega$. In any case, to check whether R is transfinitely Euclidean, we need only consider ordinals not exceeding $|R|$.

As before, we have

THEOREM 1.2. *If R is a transfinitely Euclidean ring, then R satisfies the division algorithm relative to the ordinal valued function ϕ defined on R by (11). This function again satisfies (8), and (9) holds if and only if R is an integral domain.*

Conversely, if R satisfies the division algorithm relative to an ordinal valued function θ on R , then R is transfinitely Euclidean and $\theta(x) \geq \phi(x)$ for all $x \in R$. ■

In many Euclidean rings the remainder r in the division algorithm (1) is uniquely determined and we record the condition for this to happen.

PROPOSITION 1.3. *Let R be a Euclidean ring relative to the function θ , then the remainder in the division (1) is unique if and only if*

- (i) $\theta(a-b) \leq \max\{\theta(a), \theta(b)\}$ for all $a, b \in R$ and
- (ii) $\theta(a) \leq \theta(ab)$ for all $a, b \in R$ such that $ab \neq 0$.

Proof. Suppose that remainders are unique. Given $a, b \in R$, put $x = a, y = a - b$, then $x = y \cdot 0 + a = y \cdot 1 + b$, hence by uniqueness, $\theta(a) \geq \theta(y)$ or $\theta(b) \geq \theta(y)$. To establish (ii), if $\theta(ab) < \theta(a)$ but $ab \neq 0$, then $0 = ab - ab = a \cdot 0 + 0$, which contradicts uniqueness.

Conversely, when (i), (ii) hold and $\theta(r) < \theta(y)$, $\theta(s) < \theta(y)$ and $x = yq_1 + r = yq_2 + s$, then on writing $q = q_2 - q_1$, if $yq = r - s \neq 0$, we have $\theta(yq) = \theta(r - s) < \theta(y) \leq \theta(yq)$, which is a contradiction, hence $r = s$. ■

In fact, the uniqueness of the remainder ensures that we have the fastest algorithm:

PROPOSITION 1.4. *Let R be a Euclidean ring relative to θ . If the remainder in the division algorithm is unique and the values of θ and their limits form an initial segment of the ordinals, then θ is the least function for which the algorithm holds.*

Proof. Let ϕ be the least function for which the algorithm holds; then $\phi(x) \leq \theta(x)$ for all $x \in R$ and we must show that equality holds. Assume the contrary; then

$$\phi(a) < \theta(a)$$

for some $a \in R$, and we can choose a so that $\phi(a)$ is as small as possible. If we take $b \in R$ such that $\theta(b) = \phi(a)$, then there is a unique element $r = b - aq$ such that $\theta(r) < \theta(a)$. In fact, since $\theta(b) < \theta(a)$, we have $r = b$ by uniqueness, and for any $q \neq 0$, $\theta(b - aq) \geq \theta(a) > \phi(a)$. Now take $c = b - aq$ such that $\phi(c) < \phi(a)$, then $\phi(c) < \phi(a) < \theta(a) \leq \theta(c)$, and this contradicts the choice of a . ■

Sometimes it is convenient to put $\phi(0) = -1$, $\phi(1) = 0$. With this definition we can show that for Euclidean domains with unique remainder the least function defining the algorithm is in fact a degree-function.

PROPOSITION 1.5. *Let R be a Euclidean domain and ϕ the least function defining the algorithm. Then*

$$\phi(ab) \geq \phi(a) + \phi(b) \quad \text{for } a, b \neq 0, \quad (12)$$

with equality provided that remainders are unique.

Proof. By Th. 1.1, $\phi(ab) \geq \phi(a)$; now fix $c \neq 0$, we shall show that R is Euclidean relative to the function $\theta(x) = \phi(cx) - \phi(c)$. Given a, b , where $b \neq 0$, we have $cb \neq 0$ and by the Euclidean algorithm,

$$ca = cb \cdot q + s, \quad \phi(s) < \phi(cb). \quad (13)$$

Put $a - bq = r$, then $s = cr$. By cancelling c in (13) we obtain $a = bq + r$, and by subtracting $\phi(c)$ from the inequality we have $\theta(r) < \theta(b)$; thus R is Euclidean relative to θ . By Th. 1.1, we have $\theta(x) \geq \phi(x)$, which is (12). If the remainders for ϕ are unique, then they are also unique for θ , so $\theta = \phi$ by Prop. 1.4, hence we have equality in (12). ■

The principal applications of the division algorithm are to two classes of rings:

- (i) rings of algebraic integers,
- (ii) polynomial rings over fields.

In (i) the role of θ is played by the norm; in (ii) one uses the degree of the polynomial. Of these only the latter is a degree-function; since we shall mainly be concerned with generalizations of (ii), we shall concentrate on rings satisfying the division algorithm relative to a degree-function.

Let R be any ring with a degree-function d . Then as we have seen in 0.10, R is an integral domain; moreover, any unit of R has degree 0, for if $ab = 1$, then $d(a) + d(b) = d(ab) = d(1) = 0$, hence $d(a) = d(b) = 0$. In particular, if R is a field, then every non-zero element has degree 0, so d is trivial and the division algorithm holds relative to d .

Next consider the ring of polynomials $K[x]$ in a single indeterminate x over a field K ; of course, this satisfies the division algorithm relative to the usual degree-function. More generally, this holds for the skew polynomial ring $K[x; \alpha, \delta]$, where α is any endomorphism of K and δ an α -derivation. This is most easily verified in the form A: If $a = a_0 + \alpha a_1 + \dots + \alpha^n a_n$, $b = b_0 + \alpha b_1 + \dots + \alpha^m b_m$ ($a_n, b_m \neq 0, n \geq m$), then $\deg(a - b \cdot b_m^{-1} \alpha^{-m} a_n) < n = \deg(a)$. We note that $K[x; \alpha, \delta]$ does not satisfy the left division algorithm, unless α is an automorphism of K .

We now show that the examples just given exhaust the rings with a right division algorithm relative to a degree-function.

THEOREM 1.6. *Let R be a ring with a degree-function $d: R \rightarrow \mathbf{N} \cup \{-\infty\}$. Then R satisfies the right division algorithm relative to d if and only if either (i) R is a field or (ii) R is of the form $K[x; \alpha, \delta]$ for some field K , with endomorphism α and α -derivation δ , where $d(x) > 0$.*

Here d is trivial in (i), and in (ii) is a multiple of the usual degree: $d(f) = d(x) \cdot \deg f$.

Proof. Suppose that R satisfies condition A relative to d and write $K = \{a \in R \mid d(a) \leq 0\}$. By the properties of the degree-function, K is a subring of R . Given $a \in K^\times$, we have $d(a) = 0$, hence there exists $b \in R$ such that $d(ab - 1) < d(a) = 0$, so $ab = 1$, and $d(b) = 0$, i.e. $b \in K^\times$. Thus every non-zero element in K has a right inverse in K , whence K is a field. If R has no elements of positive degree, then $R = K$ and case (i) follows. Otherwise we take an element x , say, of least positive degree in R and assert that every element of R is of the form

$$\sum_0^n x^i a_i \quad (a_i \in K, n \geq 0). \quad (14)$$

For if this were not so, there would be elements not of this form; let b be an element of least degree which is not of this form (14). By DA, there exists $q \in R$ such that $d(b - xq) < d(x)$, and by the definition of x it follows that $b - xq \in K$. Thus for some $a \in K$ we have

$$b = xq + a, \quad (15)$$

and $d(q) < d(x) + d(q) = d(xq) \leq \max\{d(b), d(a)\} = d(b)$. Therefore, by the choice of b , q must be of the form $q = \sum x^i a_i$ ($a_i \in K$). Inserting this expression

in (15), we obtain

$$b = \sum x^{i+1} a_i + a,$$

which contradicts the assumption that b is not of the form (14). Moreover, the form (14) for any element of R is unique, for otherwise we should have a relation

$$x^n c_n + \dots + c_0 = 0,$$

say, where $c_n \neq 0$. Hence

$$d(x^n) = d(x^n c_n) \leq \max \{d(x^i c_i) \mid i = 0, 1, \dots, n-1\},$$

i.e. $nd(x) \leq (n-1)d(x)$, in contradiction with the assumption $d(x) > 0$. Finally we have, for any $a \in K$,

$$ax = xa_1 + a_2,$$

where $d(a_2) < d(x)$ and $d(a_1) \leq d(ax) - d(x) = 0$, hence $a_1, a_2 \in K$. By the uniqueness of the form (14) it follows as in 0.10 that $a \mapsto a_1$ is an endomorphism α , say of K and $a \mapsto a_2$ is an α -derivation δ , thus $R = K[x; \alpha, \delta]$ as asserted.

To prove the converse, consider a skew polynomial ring $K[x; \alpha, \delta]$ with a degree-function d . We have already seen that the elements of $K^\times$ must have degree 0, and by hypothesis $d(x) = \lambda > 0$, hence the degrees $d(x^n) = n\lambda$ are all different for different n . From the properties of degree-functions proved in 0.10 it follows that for $a_n \neq 0$,

$$d(x^n a_n + \dots + a_0) = \max \{d(x^i a_i) \mid i = 0, 1, \dots, n\} = n\lambda.$$

Thus all degrees are multiples of λ . We may therefore divide the degrees by λ and so obtain the usual degree-function on $K[x; \alpha, \delta]$. As we saw earlier, the right division algorithm holds for this degree-function. ■

Exercises 2.1

1. Show that the ring of integral quaternions over the integers satisfies the Euclidean algorithm relative to the norm function (q is integral if its minimal equation has coefficients in $\mathbb{Z}$). Does this still hold for the ring of quaternions with integer coefficients? (Recall that every principal ideal domain is integrally closed in its field of fractions.)

2. (Sanov [67]) Let R be a commutative Euclidean domain relative to an $\mathbb{N}$ -valued function ϕ and on the matrix ring R_n define $|A| = \phi(\det A)$. Show that for any $A, B \in R_n$ with $|A| \neq 0$, there exist $P, Q \in R_n$ such that

$$B = AQ + P, \quad 0 < |P| < |A| \quad \text{or} \quad P = 0.$$

Use this result to obtain a reduction to triangular form for matrices over R .

3°. (Samuel) Determine all rings with a division algorithm in which the number of atomic factors in the complete factorization of an element serves as value function.

4. (Hasse [28]) Show that a commutative integral domain R is a principal ideal domain if and only if there is an $\mathbb{N}$ -valued function ϕ on $R^\times$ such that (i) $a|b$ implies $\phi(a) \leq \phi(b)$, with equality if and only if $aR = bR$ and (ii) if neither of a, b divides the other, then there exist $p, q, r \in R$ such that $pa + qb = r$, $\phi(r) < \min\{\phi(a), \phi(b)\}$.

5°. Generalize Ex. 4 to obtain a characterization of Bezout domains.

6°. Investigate rings with a positive *real*-valued function satisfying the conditions (i) and (ii) of Ex. 4.

7. Let R be a Euclidean domain with fastest algorithm given by ϕ . Given $a \in R^\times$, show that $\phi(xa) = \phi(x)$ for some $x \in R^\times$ if and only if a is a unit.

8. Show that a Euclidean ring (with division algorithm (1)) is a principal right ideal ring.

9. If R is a Euclidean domain and T is a right Ore set, show that R_T is again Euclidean.

10. (Samuel) If R is a Euclidean domain, show that $R[[x]][x^{-1}]$ is also Euclidean. (Hint: Define $\phi(\sum x^i a_i) = \phi(a_s)$, where $s = \min\{i | a_i \neq 0\}$.)

11. (Dress [71]) Let R be an integral domain. If $R[[x]][x^{-1}]$ is a Euclidean domain with function ϕ , show that R is Euclidean with function $\phi_R(a) = \min\{\phi(f) | d(f) \in Ra\}$, where $d(f)$ is the coefficient of the least power occurring in f .

12*. (Lenstra [74]) Show that for a commutative ring R with unique remainder algorithm, $U(R) \cup \{0\} = k$ is a field. Deduce that a commutative ring has a unique remainder algorithm with function θ if and only if θ is a degree-function or $R = \mathbb{F}_2 \times \mathbb{F}_2$.

13°. (S. Singh) Find an integral domain which is Euclidean and has a non-multiplicative least value function.

14. (Lemmlain [54]) Let R be an integral domain with a function ϕ assuming non-negative integer values such that $\phi(0) = 0$, $\phi(x) > 0$ for $x \neq 0$, and there exists n_0 such that for any $x \in R$ satisfying $\phi(x) > n_0$ there exists $y \in R$ such that $\phi(y - qx) \geq \phi(x)$ for all $q \in R$. Show that R is not Euclidean.

2.2 The n -term weak algorithm

Any generalization of the Euclidean algorithm will necessarily depend on the form of the value function. We shall not make the most general choice, but take our function to be a *filtration*. This means that we have a map from R to $\mathbb{N} \cup \{-\infty\}$ with the properties:

- V.1. $v(x) \geq 0$ for all $x \neq 0$, $v(0) = -\infty$,
 V.2. $v(x-y) \leq \max\{v(x), v(y)\}$,
 V.3. $v(xy) \leq v(x) + v(y)$,
 V.4. $v(1) = 0$.

Such a function is also called a *pseudo-valuation*. If equality holds in V.3, we have a degree-function as defined in 0.10 ($-v(x)$ is then a valuation, as usually defined). Even in the general case we shall call $v(x)$ the *degree* of x .

Given any filtration, let us write R_h for the set of elements of degree at most h ; then the R_h are subgroups of the additive group of R such that

- (i) $0 = R_{-\infty} \subseteq R_0 \subseteq R_1 \subseteq \dots$,
 (ii) $\bigcup R_h = R$,
 (iii) $R_i R_j \subseteq R_{i+j}$,
 (iv) $1 \in R_0$.

Conversely, any series of subgroups R_h of the additive group of R satisfying (i)–(iv) leads to a filtration v , given by $v(x) = \min\{h \mid x \in R_h\}$, as is easily seen. We remark that every ring has the *trivial* filtration

$$v(x) = \begin{cases} 0 & \text{if } x \neq 0, \\ -\infty & \text{if } x = 0. \end{cases}$$

Let R be a filtered ring, with filtration v ; we shall say that the family (a_i) ($i \in I$) of elements of R is *right v -dependent* if there exist elements $b_i \in R$, almost all 0, such that

$$v(\sum a_i b_i) < \max\{v(a_i) + v(b_i)\},$$

or if some $a_i = 0$. Otherwise the family (a_i) is *right v -independent*. We note that any right v -independent family will be right linearly independent over R . Of course, the converse need not hold; in fact, linear dependence is just the special case of v -dependence obtained by taking v to be the trivial filtration.

An element a of R is said to be *right v -dependent* on a family (a_i) if $a = 0$ or if there exist $b_i \in R$, almost all 0, such that

$$v(a - \sum a_i b_i) < v(a) \quad \text{and} \quad v(a_i) + v(b_i) \leq v(a) \text{ for all } i.$$

In the contrary case a is said to be *right v -independent* of (a_i) . Note that dependence on a family (a_i) is unaffected by adjoining or removing 0 from the family. *Left v -dependence* is defined analogously.

We note that if a is right v -dependent on a set B , then it is right v -dependent on the set of members of B of degree at most $v(a)$. We also note the transitivity: If a is right v -dependent on a set B and each element of B is right v -dependent on a set C , then a is right v -dependent on C .

If one member of a family is right v -dependent on the rest, the family will be right v -dependent, but the converse is not generally true. In fact the converse constitutes the 'weak algorithm', as expressed in the following

Definition. A ring R with a filtration v is said to satisfy the n -term weak algorithm relative to v , if in any right v -dependent family of at most n members, say $a_1, \dots, a_m$, ($m \leq n$), where

$$v(a_1) \leq \dots \leq v(a_m),$$

some a_i is right v -dependent on $a_1, \dots, a_{i-1}$. If R satisfies the n -term weak algorithm for all n , we shall say that R satisfies the weak algorithm for v .

For example, the 1-term weak algorithm states that $v(ab) = v(a) + v(b)$ for all $a, b \neq 0$, i.e. that v is a degree-function, and the 2-term weak algorithm states that v is a degree-function and for any $a, b \in R$ such that $v(a) \geq v(b)$, we have $v(a - bc) < v(a)$ for suitable $c \in R$, whenever a, b are right v -dependent. In other words, the division algorithm holds for any pair of right v -dependent elements.

Strictly speaking, the notion here defined should be described as the weak right algorithm, but we shall see in 2.3 that the property of satisfying the (n -term) weak algorithm for a given filtration is left-right symmetric.

We note that for any $n' \leq n$, the n -term weak algorithm entails the n' -term weak algorithm. Thus in any filtered ring R we shall be able to prove more about R , the larger n is. Let us define the *dependence number* of R relative to the filtration v , written $\lambda_v(R)$, as the greatest integer n for which the n -term weak algorithm holds, or ∞ if no such n exists. Thus $\lambda_v(R) = \infty$ means that R satisfies the weak algorithm for v ; $\lambda_v(R) \geq 1$ means that the 1-term weak algorithm holds, i.e. v is a degree-function. In particular, such a ring will be an integral domain. For this reason the notion of dependence as defined here is of no interest for rings other than integral domains. In fact, we shall almost exclusively be concerned with integral domains in this book.

In any filtered ring R , the set $R_0 = \{a \in R \mid v(a) \leq 0\}$ is clearly a subring. If moreover, $\lambda_v(R) \geq 2$, then for any $a \in R_0$ the pair $\{1, a\}$ is right v -dependent, so if $a \neq 0$, then 1 is right v -dependent on a , i.e. there exists $b \in R_0$ such that $v(1 - ab) < v(1) = 0$. Thus $ab = 1$ and a is invertible; hence R_0 is a field whenever R satisfies the 2-term weak algorithm for v . For this reason, in considering filtered rings, we usually confine our attention to rings where R_0 is a field.

To illustrate the notion of dependence, let us consider the commutative case or, a little more generally, the case of Ore domains. In a right Ore domain, any set of more than one element is clearly right v -dependent. There-

fore, if $\lambda_v(R) \geq 2$, the familiar division algorithm holds, in the form A. Conversely, if the classical division algorithm holds in R , then any element of R is right v -dependent on any non-zero element not of higher degree, and this in turn shows that $\lambda_v(R) = \infty$. These results may be summed up in

PROPOSITION 2.1. *If R is a filtered right Ore domain, then there are exactly three possibilities:*

- (i) $\lambda_v(R) = 0$: v is not a degree-function,
- (ii) $\lambda_v(R) = 1$: v is a degree-function, but the division algorithm does not hold,
- (iii) $\lambda_v(R) = \infty$: v is a degree-function and the division algorithm holds in R . ■

In contrast to this result, for non-commutative rings λ_v can have any positive integer value, as we shall see in 2.11. All this is, of course, in strict parallel with the n -fir condition (cf. Prop. 1.1.7 and 1.2.2).

We now come to the connexion between the weak algorithm and semifirs. First we need a general result on filtered rings:

LEMMA 2.2. *Let R be a filtered ring. Then any n -tuple of elements of R can be reduced by a member of $E_n(R)$ acting on the right, to an n -tuple, each of whose non-zero terms is v -independent of the remaining terms.*

Proof. If a non-zero term of an n -tuple is right v -dependent on the rest, then by adding to it an appropriate right linear combination of the remaining terms, we can reduce its degree without affecting the degrees of the other terms. This clearly corresponds to right action by a member of $E_n(R)$. We repeat the process until no such terms remain. It must terminate since the set of values assumed by v is well-ordered: $-\infty, 0, 1, \dots$ and hence so is the sequence of degrees $(v(a_1), \dots, v(a_n))$ in the lexicographic ordering. ■

Now let R be a ring satisfying the n -term weak algorithm relative to a filtration v . Given any m -tuple ($m \leq n$), whose non-zero elements are right v -dependent, by the n -term weak algorithm, one of the non-zero terms is right v -dependent on the rest. By Lemma 2.2 we can apply a member of $E_m(R)$ so as to obtain an m -tuple whose non-zero elements are right v -independent and hence right linearly independent. This establishes

THEOREM 2.3. *If R is a filtered ring satisfying the n -term weak algorithm, then R is a strong E_n -ring, and in particular an n -fir. ■*

Let us consider more closely the structure of a right ideal α in a filtered ring R , where R_0 is a field. A subset B of α will be called a *weak v -basis* for α if (i) all elements of α are right v -dependent on B , and (ii) no element of B is right v -dependent on the rest of B . It is easily seen, using the well-ordering of the range of v , that a weak v -basis of α generates α as right ideal; but in general it need be neither v -independent nor a minimal generating set.

When R_0 is a field K , every right ideal α of R has a weak v -basis, which may be constructed as follows: For any integer $h \geq 0$, $\alpha_h = \alpha \cap R_h$ is clearly a right K -space; moreover the set α'_h of all elements of α_h that are right v -dependent on α_{h-1} is also a right K -space. For evidently α'_h is closed under right multiplication by elements of K ; closure under addition is clear if the sum has degree h , while if it does not, the sum lies in α_{h-1} . Now for each $h \geq 0$, choose a minimal set B_h spanning α_h over α'_h , i.e. representatives for a right K -basis of α_h/α'_h and put $B = \bigcup B_h$. By induction on h it follows that every member of α_h is right v -dependent on B , for all h . However, no element of B is right v -dependent on the rest of B , by the minimality condition in our choice of elements of B_h . Thus B is a weak v -basis for α .

Conversely, every weak v -basis of α must have the property that the elements of degree h in it form a right K -basis of $\alpha_h \pmod{\alpha'_h}$. Hence any two weak v -bases of a right ideal α have the same number of elements in each degree h , viz. $\dim_K(\alpha_h/\alpha'_h)$. We shall call this the *number of v -generators in degree h* and denote it by $r_h(\alpha)$. We remark that when the weak v -basis of α is right v -independent, then α is free of rank $\sum r_h(\alpha)$.

If α and β are two right ideals of R such that $\alpha \subset \beta$, and s is the least degree for which $\alpha_s \neq \beta_s$, then $r_h(\alpha) = r_h(\beta)$ for $h < s$, while $r_s(\alpha) < r_s(\beta)$ provided that $r_s(\alpha)$ is finite.

Although every right ideal in a filtered K -ring R has a right weak v -basis, this basis need not be v -independent. But if α happens to have a v -independent generating set, B say, then B must be a weak v -basis. For, given $a \in \alpha$, let us write $a = \sum b_i c_i$ ($b_i \in B$, $c_i \in R$). By the v -independence of the b_i we have $v(a) = \max\{v(b_i) + v(c_i)\}$, and so $v(a - \sum b_i c_i) = -\infty$ is a relation of v -dependence of a on B . Hence all elements of α are v -dependent on B , while clearly no element of B is v -dependent on the rest.

So far, R was any filtered ring in which R_0 is a field. We now impose the weak algorithm; then any weak v -basis of a right ideal α is right v -independent, and so α is free, with any weak v -basis as free generating set. Of course, α may have a basis which is not a weak v -basis. But by treating in turn the basis elements of degrees $0, 1, 2, \dots$ we find that such a basis can always be transformed to a weak v -basis by a (possibly infinite) sequence of elementary transformations, and so it has the same number of elements. Thus α has unique rank and we have proved

THEOREM 2.4. *Every filtered ring with weak algorithm is a right fir, and each right ideal has a v -independent basis. ■*

We note that the uniqueness of the rank of $\mathfrak{a}$ (IBN) also follows from Th. 2.3. From the symmetry of the weak algorithm (to be proved in 2.3) it will follow that a ring with weak algorithm is in fact a two-sided fir.

From Prop. 1.2.2 (or by an easy direct proof) we obtain

COROLLARY 2.5. *Every ring satisfying the right division algorithm relative to a degree-function is a principal right ideal domain. ■*

By contrast we do not have a two-sided conclusion here, owing to the asymmetry of the division algorithm.

We next turn to the case of a ring with n -term weak algorithm. As in Th. 2.3, we find again that this is an n -fir, but in addition we shall see that it also satisfies left and right ACC_n . The result will follow from a general property of weak v -bases in filtered rings:

LEMMA 2.6. *Let R be a filtered ring such that R_0 is a field. Then for any integer n , R satisfies ACC on right ideals $\mathfrak{a}$ such that $\Sigma_h r_h(\mathfrak{a}) \leq n$.*

Proof. Let $\mathfrak{a}$ be a right ideal with $\Sigma_h r_h(\mathfrak{a}) \leq n$; then $\mathfrak{a}$ has a weak v -basis $a_1, \dots, a_m$ ($m \leq n$). We associate with $\mathfrak{a}$ the n -tuple $(v(a_1), v(a_2), \dots, v(a_m), \infty, \dots, \infty)$ (with $(n-m)$ ∞ 's) as 'indicator'. Clearly the indicator is independent of the choice of weak v -basis. If $\mathfrak{a} \supset \mathfrak{b}$, the indicator of $\mathfrak{a}$ will be smaller, under the lexicographic ordering, than that of $\mathfrak{b}$. Since the set of these indicators is well-ordered, the ideals satisfy ACC . ■

PROPOSITION 2.7. *Let R be a filtered ring with n -term weak algorithm, and in the case $n = 1$ assume also that R_0 is a field[†]. Then R is an n -fir and satisfies left and right ACC_n .*

Proof. By Lemma 2.2 (and the proof of Th. 2.3) we see that the n -generator right ideals of R will have v -independent generating sets of at most n elements. These will be weak v -bases, hence such right ideals are free of unique rank and, by Lemma 2.6, they satisfy ACC . The assertion for left ideals follows by symmetry, if we accept the fact, soon to be proved (Cor. 3.2) that the n -term weak algorithm is left-right symmetric. ■

[†] For $n \geq 2$ this is automatic (as we have seen), but the proviso cannot be omitted (cf. Ex. 1).

Exercises 2.2

1. Let R be a filtered ring. If $\lambda_v(R) > 1$, show that R_0 is a field. Give an example of a ring with 1-term weak algorithm but not satisfying left or right ACC_1 . (Hint: Use Prop. 2.1.)
2. Define filtered modules over a filtered ring R and introduce the notion of weak algorithm for filtered modules. Show that every module satisfying the weak algorithm is free; what does the existence of a non-zero module with weak algorithm imply about R ?
3. Investigate rings satisfying the weak algorithm relative to the trivial filtration.
- 4°. Generalize Hasse's characterization of principal ideal domains (Ex. 1.4) to firs.
5. Show that R has a filtration v such that $\lambda_v(R) > 0$ if and only if R is an integral domain.
- 6°. Investigate the notion of a weak algorithm relative to a function ϕ more general than a filtration.

2.3 The associated graded ring

By a *graded ring* H we understand a family of disjoint abelian groups $H_0, H_1, \dots$ indexed by $\mathbb{N}$, with biadditive 'multiplication' maps $H_i \times H_j \rightarrow H_{i+j}$ satisfying the obvious associativity relations on $H_i \times H_j \times H_k$, such that H_0 has a ring structure for which each H_i is an H_0 -bimodule. More precisely, what we have defined is an *externally graded ring*. We write $H = \bigcup H_i$ and note that even the zero-elements of the H_i are taken to be distinct. Thus for each $a \in H$ there is a unique integer $d(a)$, called the *degree* of a , such that $a \in H_{d(a)}$.

The graded rings defined above are $\mathbb{N}$ -graded; here $\mathbb{N}$ may be replaced by any commutative monoid or group (usually ordered). Thus we may speak of $\mathbb{Z}$ -graded rings or also $\mathbb{Z}^2$ -graded rings; this is used at one point in 6.8.

When we speak of a *relation*

$$\sum a_i b_i = 0 \quad (1)$$

holding in H , this will, of course, entail that the $d(a_i) + d(b_i)$ are equal for all i . A family of elements $a_i \in H$ satisfying such a relation, where not all the b_i vanish, is said to be *right linearly dependent*. The triviality of such relations is defined as for ordinary rings (cf. 1.1).

With every filtered ring a graded ring may be associated as follows. Let R be a filtered ring and for each $i \geq 0$, write $\text{gr}_i R$ for the abelian group R_i/R_{i-1} ; thus $\text{gr}_0 R$ is simply R_0 . This family has a natural graded ring structure: given $\alpha \in \text{gr}_i R, \beta \in \text{gr}_j R$, take $a \in \alpha, b \in \beta$, then a is determined $(\text{mod } R_{i-1})$ and $b(\text{mod } R_{j-1})$, hence ab is an element of R_{i+j} which is determined $(\text{mod } R_{i+j-1})$.

Hence the class in $\text{gr}_{i+j} R$ obtained in this way depends only on α and β (not on a and b) and may be denoted by $\alpha\beta$. It is not hard to verify that with this multiplication the $\text{gr}_i R$ form a graded ring, which we shall denote by $\text{gr } R$. Given $a \in R^\times$, the image of a in $\text{gr}_{v(a)} R$ will be written $\bar{a}$ and will be called the *leading term* of a ; $\bar{0}$ is not defined.

From the definition in 2.2 it is clear that a family of elements of R is right v -dependent if and only if it contains 0 or the leading terms of its elements are right linearly dependent in $\text{gr } R$. Further, an element of R is right v -dependent on a family if and only if it is 0 or its leading term is a right linear combination of the leading terms of the family.

This shows that the filtered ring R satisfies the (n -term) weak algorithm if and only if $\text{gr } R$ satisfies the following condition:

Given any right linearly dependent sequence $a_1, \dots, a_m$ ($m \leq n$), numbered so that $d(a_1) \leq \dots \leq d(a_m)$, some a_i is a right linear combination of $a_1, \dots, a_{i-1}$.

In the next result we shall reformulate this condition in other equivalent ways, but in order to do so we shall need a matrix description of those operations on finite sequences of elements of H , which add to each term a right linear combination of the terms preceding it. Given a sequence $D = (d_1, \dots, d_n)$ of non-negative integers $d_1 \leq \dots \leq d_n$, let us define $\text{Tr}_D(H)$ as the set of all $n \times n$ matrices (x_{ij}) such that for $i < j$, x_{ij} is some element of $H_{d_j - d_i}$, for $i = j$ it is $1 \in H_0$ and for $i > j$ it is zero (considered as an operation from H_{d_i} to H_{d_j}). Such matrices clearly form a group which can act from the right on $H_{d_1} + \dots + H_{d_n}$ in the desired manner.

Similarly, given $D' = (d'_1, \dots, d'_n)$, where $d'_1 \geq \dots \geq d'_n$, the set of all matrices which can act from the left on $H_{d_1} + \dots + H_{d_n}$ by adding to each term a left linear combination of subsequent terms will form a group of upper uni-triangular matrices, which we shall call $\text{Tr}_{D'}^0(H)$. If $d = d_1 + d'_1 = \dots = d_n + d'_n$, where the d_i are non-decreasing and the d'_i non-increasing, we see that $\text{Tr}_D(H) = \text{Tr}_{D'}^0(H)$; and if we further write the relation (1), where $d(a_i) = d_i$, $d(b_i) = d'_i$, in matrix form as $a \cdot b = 0$, we see that for $P \in \text{Tr}_D(H)$, we have $aP \cdot b = a \cdot Pb$.

Let us call the relation (1) *well-arranged* if $d(a_1) \leq \dots \leq d(a_n)$, or equivalently, $d(b_1) \geq \dots \geq d(b_n)$. A well-arranged relation is said to be *Tr-trivializable* if there exists $P \in \text{Tr}_D(H)$, $D = (d(a_1), \dots, d(a_n))$, such that

$$aP \cdot P^{-1}b = 0$$

is a trivial relation.

We can now state some equivalent ways of expressing the n -term weak algorithm in $\text{gr } R$:

THEOREM 3.1. *Let H be a non-zero graded ring and n a positive integer. Then the following conditions are equivalent:*

(a) *given a right linearly dependent sequence $a_1, \dots, a_m \in H$ ($m \leq n$) such that $d(a_1) \leq \dots \leq d(a_m)$, some a_i is a right linear combination of $a_1, \dots, a_{i-1}$,*

(b) *given $a_1, \dots, a_m \in H$ ($m \leq n$) with $d(a_1) \leq \dots \leq d(a_m)$, there exists an upper unitriangular matrix P such that the non-zero terms of $(a_1, \dots, a_m)P$ are right linearly independent;*

(c) *every well-arranged relation $\sum a_i b_i = 0$ of at most n terms holding in H is Tr -trivializable.*

$(a^0)-(c^0)$ *The left-right duals of (a)-(c).*

The proof is straightforward. ■

Definition. A graded ring H satisfying the equivalent conditions of Th. 3.1 is said to possess *n -term weak algorithm*. If H has n -term weak algorithm for all n , we say that it satisfies the *weak algorithm*.

The remarks preceding Th. 3.1 show that the n -term weak algorithm (and hence the weak algorithm itself) holds in a filtered ring R if and only if it holds in the associated graded ring $\text{gr } R$. From the left-right symmetry in Th. 3.1 we therefore obtain

COROLLARY 3.2. *The n -term weak algorithm, for graded or filtered rings, is left-right symmetric.* ■

While for an ordinary ring R the condition that R be a strong Tr_2 -ring implies that R is a field, the conditions of Th. 3.1 do not imply that all the non-zero elements of H are invertible (in fact, no elements other than those of H_0 can possibly be invertible). The reason lies in our restriction of the Tr -triviality assumption to well-arranged relations. In a graded ring with 2-term weak algorithm, the relation $a \cdot 1 + (-1) \cdot a = 0$ will be Tr -trivializable if $a \in H_0$ but not otherwise, so only H_0 need be a field.

We conclude this section by describing a condition for a subring to inherit the weak algorithm. If R is a ring and S a subring of R , let us call S *left closed* in R if for any $v_1, \dots, v_n \in S$ which are left linearly independent over R , if $\sum c_i v_i \in S$, then $c_j \in S$ for $j = 1, \dots, n$. Right closed subrings are defined analogously.

THEOREM 3.3. *Let R be a graded ring satisfying the n -term weak algorithm (for some n), and let S be a subring which is left or right closed in R . Then S again has n -term weak algorithm.*

Proof. By the symmetry of the n -term weak algorithm it is enough to consider the case where S is left closed in R . We shall verify condition (a⁰) of Th. 3.1 for elements of S that are linearly dependent over R ; thus let $v_1, \dots, v_m$ ($m \leq n$) be elements of S which are left linearly dependent over R . Then there is a non-trivial relation

$$\sum c_i v_i = 0 \quad (c_i \in R, \text{ not all } c_i \text{ vanish}). \quad (2)$$

We shall use induction on m . By renumbering and using the n -term weak algorithm in R , we may assume that $c_1, \dots, c_r$ in (2) are right linearly independent over R , while $c_{r+1}, \dots, c_m$ are linearly dependent on them. Moreover, $r \geq 1$, because (2) is non-trivial. Thus we have

$$c_j = \sum_1^r c_i a_{ij} \quad (a_{ij} \in R, j = r+1, \dots, m). \quad (3)$$

Substituting from (3) into (2), we obtain

$$\sum_1^r c_i (v_i + \sum_j a_{ij} v_j) = 0.$$

By the linear independence of $c_1, \dots, c_r$ over R we have $v_i + \sum_j a_{ij} v_j = 0$, for $i = 1, \dots, r$. Now either $v_{r+1}, \dots, v_m$ are left linearly dependent over R , then the result follows by induction, because $r > 0$, or $v_{r+1}, \dots, v_m$ are left linearly independent over R , and then $a_{ij} \in S$ because S is left closed in R . Thus v_i (for $i \leq r$) is left linearly dependent (over S) on $v_{r+1}, \dots, v_m$ and the result follows. ■

COROLLARY 3.4. *Let R be a filtered ring with weak algorithm and S a subring such that $\text{gr } S$ is left (or right) closed in $\text{gr } R$, then S also has weak algorithm.* ■

Exercises 2.3

1. Carry out all the verifications that for a filtered ring R , $\text{gr } R$ is a graded ring.
2. Develop a theory of graded modules over graded rings; in particular, define the free module on a given set of free generators (with preassigned degrees) and prove the universal property of such modules.
3. Define a *graded semifir* as a graded ring in which every relation is trivializable. Show that (i) every filtered ring with a weak algorithm is a graded semifir (relative to the grading induced by the filtration) and (ii) every graded semifir is a semifir.
4. Show that the property (of a subring) of being left closed is transitive and is preserved by intersections. Give an example of a subring of $k\langle x, y \rangle$ which is left but not right closed.
- 5⁰. Investigate rings graded by the positive real numbers.

2.4 Weak algebra bases in filtered rings and free algebras

In this section we shall prove an analogue of Th. 1.6 which will describe the rings with a weak algorithm. If R is a filtered ring, we define the 'formal degree' of an expression $\sum_i a_{i1} \dots a_{in_i}$ as

$$\max_i \{v(a_{i1}) + \dots + v(a_{in_i})\}.$$

Clearly the actual degree of an element of R never exceeds the formal degree of any expression for it. We also note that the definition of v -independence of a family states that the degree of elements represented by certain expressions should equal the formal degree of these expressions.

Let R be a filtered ring for which R_0 is a field K . A set X in R is called a *weak algebra basis* if the monomials in X span R as a right K -space and no element of X is right v -dependent on the rest. Such a set X may be constructed as follows:

For each $h > 0$ denote by R'_h the right K -subspace of R_h spanned by the products ab , where $a, b \in R_{h-1}$ and $v(a) + v(b) \leq h$. Now choose a minimal set X_h spanning $R_h \pmod{R'_h}$ over K , i.e. a set of representatives for a right K -basis of R_h/R'_h , and put $X = \bigcup X_h$. To show that X has the properties stated above, suppose that an element $x \in X$ is right v -dependent on the rest:

$$x \equiv \sum x_j b_j \pmod{R_{h-1}}, \quad (1)$$

where $h = v(x)$. Any terms $x_j b_j$ with $v(x_j) < h$ lie in R'_h , so (1) implies that

$$x \equiv \sum x_j \beta_j \pmod{R'_h},$$

where $\beta_j \in K$ and $v(x_j) = h$ whenever $\beta_j \neq 0$. But this contradicts the construction of X ; thus no element of X is right v -dependent on the rest. Now an easy induction on the degree shows that the monomials in X span R as right K -space; more precisely, the monomials of formal degree at most h span R_h . Thus we see that every filtered ring R for which R_0 is a field has a weak algebra basis. As in the case of weak v -bases of right ideals, we see that the cardinal of a weak algebra basis (and more precisely, the number of elements of a given degree) is independent of the choice of basis.

We now show that the weak algebra basis is v -independent precisely when the weak algorithm holds. Given any set X in a K -ring R , if R is spanned as right K -space by the monomials in X , we can define a filtration on R by assigning to each $x \in X$ an arbitrary degree and to each element a of R the minimum of the formal degrees of the right K -linear expressions in X representing a . We shall denote by $x_I = x_1 x_2 \dots x_n$ a monomial in X .

THEOREM 4.1. *Let R be a filtered ring. Then R satisfies the weak algorithm if and only if R_0 is a field K and R has a right v -independent weak algebra basis X , whose elements have positive degrees.*

When this is so, the monomials in any weak algebra basis X form a right K -basis for R and the degree of any expression $\sum x_i \alpha_i$ is its formal degree.

Proof. Suppose that R satisfies the weak algorithm; then any weak algebra basis X of R is clearly right v -independent, and each $x \in X$ has positive degree.

Conversely, let X be a weak algebra basis satisfying these conditions. Since X has no elements of degree 0, it follows that $R_0 = K$. Suppose that the monomials in X are linearly dependent, say $\sum x_i \alpha_i = 0$. By splitting off the left-hand factor from X in each x_i we can write this as

$$\sum x \alpha_x + \alpha = 0, \quad (x \in X, \alpha_x \in R, \alpha \in K).$$

By the v -independence of X , each $\alpha_x = 0$ and so $\alpha = 0$. Now an induction on the formal degree shows that the given relation was trivial. Thus the monomials in X are right K -linearly independent and hence form a K -basis for R .

We complete the proof by showing that R satisfies the left-hand analogue of the weak algorithm, which we know is equivalent to the weak algorithm itself. Let us consider how monomial terms, i.e. scalar multiples of monomials, multiply in R . The product $(x_1 \dots x_i \alpha)(y_1 \dots y_j \beta)$ can be written $(x_1 \dots x_i)(\alpha y_1 \dots y_j \beta)$. If we write the second factor as a right K -linear combination of monomials, little can be said about the terms that will occur, except that we know a bound for their degrees. However, in the product, all terms will clearly begin with $x_1 \dots x_i$.

Let us fix a monomial $x_1 \dots x_h$ of degree r , and define the *right transduction* for this monomial as the right K -linear map $a \mapsto a^*$ of R into itself which sends any monomial of the form $x_1 \dots x_h b$ to b and all other monomials to 0. Thus a^* is the 'right cofactor' of $x_1 \dots x_h$ in the canonical expression for a . For any $a \in R$, $v(a^*) \leq v(a) - r$. Further, if $a, b \in R$, then

$$(ab)^* \equiv a^*b \pmod{R_{v(b)-1}}. \quad (2)$$

This is clear if a is a monomial term of degree at least r ; in fact we then have equality. If a is a monomial term of degree less than r , the right-hand side of (2) is 0, and so (2) holds as a congruence. Now the general case follows by linearity.

Assume now that $b_1, \dots, b_n$ is a left v -dependent family, i.e.

$$v(\sum a_i b_i) < d = \max_i \{v(a_i) + v(b_i)\}.$$

Taking the b_i ordered so that $v(b_1) \geq \dots \geq v(b_n)$, we must show that some b_i is left v -dependent on those that follow. By omitting terms if necessary we may assume that $v(a_i) + v(b_i) = d$ for all i , hence $v(a_1) \leq \dots \leq v(a_n)$.

Let $x_1 \dots x_n$ be a product of maximal degree $r = v(a_1)$ occurring in a_1 with a non-zero coefficient α and denote the right transduction for $x_1 \dots x_n$ by $*$. Consider now $\sum a_i^* b_i$; the i th term differs from $(a_i b_i)^*$ by a term of degree $< v(b_i) \leq v(b_1)$. Hence the sum will differ by a term of degree less than $v(b_1)$ from $(\sum a_i b_i)^*$, which has degree $\leq v(\sum a_i b_i) - r < d - r = v(b_1)$. Therefore $v(\sum a_i^* b_i) < v(b_1)$, and this gives a relation of left v -dependence of b_1 on the remaining b_i , since $a_1^* = \alpha \in K^\times$. ■

An example of a filtered ring to which this theorem applies is the free D -ring $D_K \langle X \rangle$ on a K -centralizing set X over a field D containing a subfield K , where we assign arbitrary positive integral degrees to the elements of X (e.g. we may take them all to be 1), and use the induced formal degree. In particular, free algebras satisfy the weak algorithm; more precisely we have

PROPOSITION 4.2. *Let R be an algebra over a commutative field k with a filtration v such that $R_0 = k$. Then R is the free associative k -algebra on a set X and v is the formal degree induced from $v: X \rightarrow \mathbb{N}_+$ if and only if R satisfies the weak algorithm.*

For by hypothesis, k is contained in the centre of R and the form of the elements described in Th. 4.1 shows that R is the free k -algebra on X . ■

By Th. 2.4 we obtain

COROLLARY 4.3. *Any free associative algebra over a commutative field is a two-sided fir.* ■

In fact, the remarks preceding Prop. 4.2 show that any free D -ring $D_K \langle X \rangle$, where D is a field with subfield K , is a two-sided fir. In 2.6 we shall meet the most general ring with weak algorithm.

In connexion with Th. 4.1 we note that, given a field K and a K -ring R with a subset X such that the monomials in X form a right K -basis for R , if we assign arbitrary degrees to members of X and give elements of R their formal degrees when expressed in terms of this basis, this will not necessarily define a filtered ring structure on R . The main reason is that for $\alpha \in K$, $x \in X$, the element αx , when expressed as a right linear combination of monomials in X , may not have the same formal degree as x . In the above example of free associative algebras, this cannot happen. But in general it may not be possible

to assign a suitable filtration; e.g. we may have $\alpha x = x^2 + y$ ($\alpha \in K, x, y \in X$). In the proof of Th. 4.1 essential use was made of the fact that v was given as a filtration on R .

Exercises 2.4

1. State and prove analogues of Th. 4.1 and Prop. 4.2 for graded rings with weak algorithm.
2. Give a direct proof that the centre of a non-Ore integral domain with a weak algorithm is a field (cf. 6.3).
3. If R is a ring with weak algorithm, show that R_0 is complemented by a right ideal. Give examples to show that R_0 may not have a complement which is a two-sided ideal. (Hint: Try skew polynomial rings.)
4. Let R be a K -ring with a subset X whose monomials form a right K -basis for R . Given a filtration on R such that $R_0 \subseteq K$, show that X is right v -independent if and only if every element of positive degree in R is v -dependent on X . Deduce that R satisfies the weak algorithm relative to this filtration.
5. Let $R = k\langle X \rangle$ be a free algebra; by regarding R as the universal associative envelope of the free Lie algebra on X , define a filtration on R for which the weak algorithm is not satisfied.
6. In $\mathbb{Z}\langle X \rangle$ define the *content* $c(f)$ of f by $c(0) = 0$, while for $f \neq 0$, $c(f)$ is the HCF of all the coefficients. Prove Gauss's lemma in the form $c(fg) = c(f)c(g)$. (Hint: Imitate the proof in the commutative case.)

2.5 The Hilbert series of a filtered ring

Let R be a filtered ring, where R_0 is a field K , and consider $\text{gr } R = \{\text{gr}_n R\}$, where $\text{gr}_n R = R_n/R_{n-1}$. If each $\text{gr}_n R$ is finite-dimensional as right K -space, say $\dim_K(R_n/R_{n-1}) = \alpha_n$, then we can form the formal power series

$$H(R:K) = \sum \alpha_n t^n. \quad (1)$$

It is called the *Hilbert series* of R . Using Th. 4.1, we can calculate the Hilbert series of any ring with weak algorithm:

PROPOSITION 5.1. *Let R be a filtered ring with weak algorithm, where $R_0 = K$. Let $X = \cup X_n$ be a weak algebra basis of R and define $\lambda_n = |X_n| = \dim_K(R_n/R'_n)$ in the notation of 2.4; further, put $H(X) = \sum \lambda_n t^n$. Then*

$$H(R:K) = (1 - H(X))^{-1}. \quad (2)$$

Proof. We saw that a right K -basis of R_n/R_{n-1} is formed by the set of all monomials x_I of degree n . Each sequence $(n_1, \dots, n_r)$ such that $n_1 + \dots + n_r = n$ gives rise to $\lambda_{n_1} \dots \lambda_{n_r}$ monomials of degree n , hence

$$\alpha_n = \sum \lambda_{n_1} \dots \lambda_{n_r}$$

where the summation is over all ordered partitions of n . It follows that

$$H(R:K) = \sum \lambda_{n_1} t^{n_1} \dots \lambda_{n_r} t^{n_r},$$

i.e. (2). ■

In a moment we shall give another proof of (2). Meanwhile we note the special case of free algebras:

COROLLARY 5.2. *Let $R = k\langle x_1, \dots, x_r \rangle$, where all the x_i have degree 1. Then $H(X) = rt$, and therefore $H(R:k) = (1-rt)^{-1}$. ■*

We return to a filtered K -ring and consider the Hilbert series of a right ideal. Given any right ideal α of R , let $\beta_n = r_n(\alpha) = \dim_K(\alpha_n/\alpha'_{n-1})$ (cf. 2.2) and $\gamma_n = \dim_K(\alpha_n/\alpha_{n-1})$. We note that β_n and γ_n are bounded by α_n , hence we may define the Hilbert series

$$H(\alpha:R) = \sum \beta_n t^n, \quad H(\alpha:K) = \sum \gamma_n t^n.$$

If R satisfies the weak algorithm, these series are related by the formula

$$H(\alpha:K) = H(\alpha:R)H(R:K). \quad (3)$$

For if (e_λ) is a weak v -basis of α , then a K -basis of α_n/α_{n-1} is given by the family of all $e_\lambda x_I$ of degree n . Hence we have $\gamma_n = \sum \beta_i \alpha_{n-i}$ and (3) follows.

We can use (3) to give another proof of (2). We note that (2) is essentially a statement about $\text{gr } R$, so we may assume R to be a graded ring. Then XR is the augmentation ideal, so that we have $R = K \oplus XR$ as K -spaces, and $H(XR:R) = H(X)$, $H(R:K) = 1 + H(XR:K)$. Inserting these values in (3) (for $\alpha = XR$) we obtain $H(R:K) = 1 + H(X) \cdot H(R:K)$, from which (2) follows.

We now turn to modules over rings with a weak algorithm, and derive a presentation which in some cases provides information about the characteristic of the module. We recall that for any ring R and any set B , the free left R -module on B is written $R^{(B)}$; if (b) denotes the generator indexed by $b \in B$, its elements have the form $\sum r_b(b)$ ($r_b \in R, b \in B$).

THEOREM 5.3. *Let R be a filtered ring with weak algorithm, where $R_0 = K$, and M a left R -module. If B is a left K -basis of M , and X a weak algebra basis of*

R , then there is an exact sequence

$$0 \longrightarrow R^{(X \times B)} \xrightarrow{\beta} R^{(B)} \xrightarrow{\alpha} M \longrightarrow 0, \quad (4)$$

where $\alpha: (b) \mapsto b$ ($b \in B$) and $\beta: (x, b) \mapsto x(b) - \sum \lambda_i(b_i)$ if $xb = \sum \lambda_i b_i$ in M ($\lambda_i \in K, b, b_i \in B$).

Proof. We have $\beta\alpha = 0$ by the definition of α and β . The cokernel of β is the R -module with generators (b) and defining relations $x(b) = \sum \lambda_i(b_i)$, hence the K -space spanned by the (b) is already an R -module and so is all of $\text{coker } \beta$. Thus the natural surjection $\text{coker } \beta \rightarrow M$ is an isomorphism; so far we have used no hypothesis beyond the fact that B is a left K -basis of M .

It remains to show that β is injective. Consider any non-zero element of $R^{(X \times B)}$:

$$s = \sum \lambda_{x_1, \dots, x_m, b} x_1 \dots x_{m-1} (x_m, b).$$

Put

$$n = \max \{v(x_1 \dots x_m) \mid \lambda_{x_1, \dots, x_m, b} \neq 0 \text{ for some } b\},$$

and consider $s\beta$. By the definition of β we have $(x, b)\beta = x(b) +$ terms of lower degree, hence the terms of highest degree (viz. n) in $s\beta$ are

$$\sum \lambda_{x_1, \dots, x_m, b} x_1 x_2 \dots x_m (b),$$

summed over all terms with $v(x_1 \dots x_m) = n$. This is non-zero, hence $s\beta \neq 0$ and so β is injective. ■

If X and B are finite, the characteristic of M as R -module is, by (4),

$$\chi_R(M) = (1 - |X|) \dim_K(M). \quad (5)$$

For example, if $R = k\langle x_1, \dots, x_r \rangle$, $\mathfrak{a}$ is a left ideal of finite rank d and $\dim_K(R/\mathfrak{a}) = m$, then taking $M = R/\mathfrak{a}$, we have $\chi_R(M) = 1 - d$ and (5) becomes

$$d - 1 = (r - 1)m. \quad (6)$$

This formula, due to Lewin, is the precise analogue of Schreier's formula for groups.

Let R be any k -algebra, generated by r elements over k , and write F for the free k -algebra of rank r , so that $R \cong F/\mathfrak{n}$, for some ideal $\mathfrak{n}$ of F . Any left ideal $\mathfrak{a}$ of R corresponds to a left ideal $\mathfrak{A}$ of F containing $\mathfrak{n}$ such that $R/\mathfrak{a} \cong F/\mathfrak{A}$ as left k -spaces. In particular, if $\mathfrak{a}$ is of finite codimension m over k , then so is $\mathfrak{A}$ and by (6), $\mathfrak{A}$ is then free as left F -module, of rank $d = (r - 1)m + 1$. Hence $\mathfrak{a} = \mathfrak{A}/\mathfrak{n}$ can be generated by d elements and we obtain

COROLLARY 5.4. *Let R be a k -algebra generated by r elements over k . Then any left ideal α of finite codimension m in R over k can be generated by $(r-1)m+1$ elements. ■*

Exercises 2.5

1. Let R, S be filtered algebras over a commutative field k and $T = R \otimes S$ their tensor product. Show that if the Hilbert series of R, S are defined, then so is that of T and $H(T:k) = H(R:k)H(S:k)$.

2. Let R be a filtered ring, where R_0 is a field K . Show that if $H(R:K)$ is defined, then it has an inverse $L(R:K)$ (as a power series with integer coefficients). Verify that $L(K[x]:K) = 1-t$. If R, S are filtered rings satisfying the above condition and U is their coproduct over K (cf. e.g. Cohn [77], Ch. 5), show that U also satisfies this condition and that $L(U:K) = L(R:K) + L(S:K) - 1$.

3*. Let $M = C \otimes_R C$ and consider the tensor C -ring $T = C\langle M \rangle$. Show that T is the coproduct of an ordinary polynomial ring over C and a complex-skew polynomial ring over C . How does this generalize to finite Galois extensions?

4. (Lewin [69]) Show that any two-sided ideal in a free algebra R has the same rank as left and as right R -module.

5. Let $R = k\langle x_1, \dots, x_r \rangle$ be a free algebra and α a non-zero principal left ideal. Then the formula (6) gives

$$\dim_k(R/\alpha) = (1-r)^{-1}(1-1) = 0.$$

Explain this paradox.

6. Examine the relation (3) between Hilbert series when $R = K[x; \alpha, \delta]$ is a skew polynomial ring. Do the same for the Hilbert series of the opposite ring when α is not surjective.

7. Define the notion of a filtered module M (over a filtered ring R) satisfying the weak algorithm. Show that any submodule M' with the induced filtration again satisfies the weak algorithm. Under suitable hypotheses define a Hilbert series $H(M:R)$ and show that $H(M:R_0) = H(M:R)H(R:R_0)$. If $M'' = M/M'$ has the induced filtration, show that

$$H(M'':R_0) = [H(M:R) - H(M':R)] \cdot H(R:R_0).$$

8*. For any K -ring R the universal derivation bimodule $\Omega_{R/K}$ may be defined as the kernel of the multiplication map $x \otimes y \mapsto xy$ in R (cf. Ex. 0.10.14), so that we have an exact sequence (cf. Cartan and Eilenberg [56], Lewin [74], Bergman and Dicks [75])

$$0 \longrightarrow \Omega_{R/K} \xrightarrow{\lambda} R \otimes_K R \longrightarrow R \longrightarrow 0,$$

where $\lambda: dx \mapsto x \otimes 1 - 1 \otimes x$. When $R = K\langle X, Y, Y^{-1} \rangle$ is the mixed free algebra,

verify that $\Omega_{R/K} = R^{(X \cup Y)}$ and hence obtain the exact sequence

$$0 \rightarrow (R \otimes_K M)^{(X \cup Y)} \rightarrow R \otimes_K M \rightarrow M \rightarrow 0.$$

When G is the free group on Y and $X = \emptyset$, then $R = KG$ is the group algebra. Let H be any subgroup of G and put $M = K[H \backslash G]$, where $H \backslash G = \bigcup Hg$ is the set of right cosets; deduce Schreier's formula in the form $1 - \text{rk } H = (G : H)(1 - \text{rk } G)$.

9*. Let R be a K -ring with a subset X of r elements such that the monomials in X span R as left K -space. Show that for any left R -module M there is a sequence (4), exact except possibly at $R^{(X \times B)}$. If M is s -generated as left R -module, deduce that $M \cong R^s/N$, where N is $(mr+s)$ -generated. If $\mathfrak{a}$ is a left ideal of finite codimension m in R over K , show that $\mathfrak{a}$ can be generated by $(mr+1)$ elements.

2.6 The construction of rings with a weak algorithm

In Th. 4.1 we obtained a description of filtered rings with weak algorithm, in terms of a v -independent generating set, but this is not very explicit. Although it enables one to write down many examples, it does not provide a method for constructing *all* rings with a weak algorithm. We shall now describe such a method, but in order to do so we need another concept.

By a *truncated filtered ring* of height h (briefly, *h -truncated ring*), R_h , we shall mean a finite chain of abelian groups

$$0 = R_{-\infty} \subseteq R_0 \subseteq R_1 \subseteq \dots \subseteq R_h,$$

with a function called *multiplication* defined in $\bigcup \{R_i \times R_j \mid i+j \leq h\}$ such that:

T.1. For $i+j \leq h$, multiplication restricted to $R_i \times R_j$ is a biadditive function with values in R_{i+j} .

T.2. For $i+j+k \leq h$, multiplication is associative on $R_i \times R_j \times R_k$.

T.3. R_0 contains the neutral element for multiplication, 1.

Note that we have used the same symbol, R_h , for the last term of the defining chain and for the total structure.

By a morphism $f: R_h \rightarrow S_k$ of truncated filtered rings we mean a map f respecting addition, multiplication and unit element, such that $R_i f \subseteq S_i$; in particular, $h \leq k$.

Let R_h be a truncated filtered ring and let $a \in R_h$. We define the *degree*, $v(a)$, of a as the least i such that $a \in R_i$. From the definition of multiplication in R_h we see that ab is defined precisely when $v(a) + v(b) \leq h$.

With every filtered ring and every integer h there is associated a truncated

filtered ring of height h , obtained by 'forgetting the terms of degree greater than h '. This 'truncation' functor has a left adjoint, associating with every truncated filtered ring R_h the universal filtered ring $U(R_h)$ generated by R_h . More generally, for every h -truncated ring and each $k \geq h$ there is a universal k -truncated ring.

Clearly we can define ' v -dependence' for a family of elements of a truncated filtered ring, and ' v -dependence' of an element on a family, exactly as for ordinary filtered rings. Of course, the value of $\max_i \{v(a_i) + v(b_i)\}$ in the relations considered must not exceed h . Th. 4.1 is then easily seen to go through in this context.

Given a truncated filtered ring R_h satisfying the weak algorithm, let us denote the field R_0 by K and let X be constructed as in Th. 4.1. We denote by R'_{h+1} the right K -space which has for right K -basis the monomials in X of formal degree at most $h+1$ and for degree function the formal degree obtained by expressing elements in terms of this basis, using the degrees in R_h of the elements of X . We claim that R'_{h+1} is the universal truncated filtered ring of height $h+1$ for R_h .

Indeed, the space R'_{h+1} will clearly have the desired universal property if it can be given a truncated filtered ring structure extending that of R_h . To obtain this structure it is clear how we must define addition, right multiplication by elements of K and left multiplication by elements of X , so it remains to define left multiplication of monomials in X by elements of K , i.e. to give a left K -space structure to R'_{h+1} . Moreover, we are given this structure on R_h , so it suffices to define products αu , where $\alpha \in K$ and u is a monomial of degree $h+1$ in X .

Such a monomial can be written as $u = xw$, where $x \in X$ and w is a monomial of lower degree. In R_h we have $\alpha x = \sum x_I \beta_I$, where $v(x_I) \leq v(x)$, and for each I we have $\beta_I w = \sum x_J \gamma_{JI}$, $v(x_J) \leq v(w)$. We then put

$$\alpha u = \alpha x w = \sum \sum x_I x_J \gamma_{JI}.$$

Routine calculations show that this leads to a truncated filtered ring structure of height $h+1$ on our space R'_{h+1} ; here the main point requiring verification is the associativity for triple products of monomial terms. We leave these verifications to the reader.

By Th. 4.1 (in its extended form, for truncated filtered rings), R'_{h+1} will now also satisfy the weak algorithm, and the h -truncation of R'_{h+1} is just R_h . We now wish to find the *most general* form for an $(h+1)$ -truncated filtered ring R_{h+1} with weak algorithm, having R_h as its h -truncation. If R_{h+1} is such a ring, it is clear from the method of construction of Th. 4.1, that the set X constructed for R_h can be enlarged to a corresponding set for R_{h+1} . Since monomials in the elements of this set must be right K -linearly independent,

R_{h+1} will have R'_{h+1} embedded in it, i.e. the map given by the universal property will be injective.

Let us take any K -bimodule R_{h+1} containing R'_{h+1} as subbimodule, and extend v to it by setting it equal to $h+1$ outside R_h . Then we have made R_{h+1} into a truncated filtered ring of height $h+1$. For the only multiplications that need to be defined on the elements of degree $h+1$ are their products with members of $R_0 = K$, and the conditions T.1–3 that they must satisfy are just the conditions for a K -bimodule. Further, R_{h+1} , as a truncated filtered ring, will satisfy the weak algorithm, because on enlarging X by adjoining a minimal generating set for the right K -space $R_{h+1} \pmod{R'_{h+1}}$ we get a generating set for R_{h+1} satisfying Th. 4.1.

Since any filtered ring R with a weak algorithm can be written in a unique way as $\cup R_h$, where each R_h is a truncated filtered ring of height h satisfying the weak algorithm, and equal to the h -truncation of R_{h+1} , it follows that R may be obtained in this way. Thus we have proved

THEOREM 6.1. *Any filtered ring with weak algorithm can be constructed by the following steps:*

(0) choose an arbitrary field K to be R_0 ,

⋮

(h) given R_{h-1} , form the universal extension R'_h as above, let R_h be any K -bimodule containing R'_h as subbimodule, and consider R_h as truncated filtered ring,

⋮

(∞) let R be the filtered ring $\cup R_h$. ■

Note that at step (1) we have $R'_1 = R_0$, so this step is simply: choose a K -bimodule containing K . Of course, the structure of K -bimodules over a field K is itself a non-trivial topic. Thus K -bimodules are effectively $(K^0 \otimes_{\mathbb{Z}} K)$ -modules and a ring of the form $K^0 \otimes K$ can have a highly complicated ideal structure.

As an illustration of Th. 6.1, consider a skew polynomial ring $K[x; \alpha, \delta]$. Here R_h differs from R'_h only for $h = 1$. More generally, if we apply the above construction with $R'_h = R_h$ for $h > 1$, the ring obtained will be the universal K -ring on a certain K -bimodule R_1 containing K as a subbimodule. In the special case where K is complemented (as bimodule), say $R_1 = K \oplus M$, this is just $K\langle M \rangle$, the tensor K -ring on the K -bimodule M .

As a special case we obtain

THEOREM 6.2. *Let D be a skew field, K a subfield and X any set; then $D_K\langle X \rangle$, the free D -ring on a K -centralizing set X , satisfies the weak algorithm, and hence is a fir.*

For this ring is of the form $D\langle M \rangle$, where $M = (D \otimes_K D)^{(X)}$. ■

A slight variant of this result yields an analogue of the well known HNN-construction for groups. Let D be a skew field which is a k -algebra (over a commutative field k), with subfields A, B which are k -algebras, isomorphic via a k -linear isomorphism ϕ ; then the ring

$$R = D\langle t \mid at = t(a\phi), a \in A \rangle$$

satisfies the weak algorithm. For let $M = D \otimes_A D$, where D is considered as an A -bimodule with the usual right multiplication but with left multiplication

$$a \cdot u = (a\phi)u \quad (a \in A, u \in D).$$

If we write t for $1 \otimes 1$, the elements of M can be written as sums $\sum u_i t v_i$ ($u_i, v_i \in D$), with the defining relations $at = t(a\phi)$ ($a \in A$). Now M is a D -bimodule and $D\langle M \rangle$ is therefore a ring with weak algorithm. To complete the analogy with the group case, we need to find a field containing $D\langle M \rangle$. This will be done in Ch. 7, where it is shown that every semifir can be embedded in a field.

Exercises 2.6

1. Carry out the verifications needed to show that R'_{h+1} as defined in the text is a truncated filtered ring.
2. State and prove an analogue of Th. 4.1 for truncated filtered rings.
3. Let K be a field with endomorphism α and α -derivation δ . Show that $K[x; \alpha, \delta]$ is the tensor K -ring on the K -bimodule ${}_x K$ and deduce that R is a ring with weak algorithm and hence a fir, which is right principal (but not left principal unless α is an automorphism of K).
4. For any filtered ring R write $T_h R$ for the h -truncation obtained from it, and for an h -truncated filtered ring R_h denote by UR_h the universal filtered ring. Given an h -truncated filtered ring R_h , show that the canonical map $R_h \rightarrow T_h UR_h$ is surjective but not necessarily injective.
5. Let F be a prime field. Show that every F -bimodule is a direct sum of copies of F as F -bimodule. Deduce that every filtered F -ring with weak algorithm is a free algebra.

2.7 Generators and relations for $\mathbf{GE}_2(R)$

Our next objective is to treat the analogue of the Euclidean algorithm that exists in rings with 2-term weak algorithm. One of the main consequences is that such a ring is always a strong $\mathbf{E}_2$ -ring (cf. 1.3); moreover, there is a convenient normal form for the elements of $\mathbf{GE}_2(R)$, which leads to a presentation of $\mathbf{GL}_2(R)$. Since many of the formulae are valid in quite general rings, we shall digress in this section to discuss generators and relations in $\mathbf{GE}_2(R)$ for general rings.

For brevity let us write

$$P(x) = \begin{pmatrix} x & 1 \\ 1 & 0 \end{pmatrix}, \quad [\alpha, \beta] = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}, \quad C(\alpha) = [\alpha, -\alpha^{-1}],$$

and observe the following relations between these matrices, valid over any ring R :

$$P(x+y) = P(x)P(0)P(y), \quad x, y \in R, \quad (1)$$

$$P(\alpha)P(-\alpha^{-1})P(\alpha) = C(\alpha) \quad (2)$$

$$P(x)[\alpha, \beta] = [\beta, \alpha]P(\beta^{-1}x\alpha), \quad \alpha, \beta \in U(R). \quad (3)$$

We observe that $\mathbf{GE}_2(R)$ is generated by the $P(x)$ and $[\alpha, \beta]$ since

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} = P(x)P(0) \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} = P(0)P(x).$$

From (1)–(3) we obtain the following consequences:

$$P(0)^2 = I, \quad [P(1)P(-1)]^3 = -I, \quad (4)$$

$$P(x)^{-1} = P(0)P(-x)P(0) = \begin{pmatrix} 0 & 1 \\ 1 & -x \end{pmatrix}, \quad (5)$$

$$P(x)P(y)^{-1} = P(x-y)P(0), \quad (6)$$

$$P(x)P(y)^{-1}P(z) = P(x-y+z) \quad (7)$$

and

$$P(x)P(\alpha)P(y) = P(x+\alpha^{-1})C(\alpha)P(y+\alpha^{-1}), \quad (8)$$

where $x, y, z \in R$ and $\alpha \in U(R)$. Any element A of $\mathbf{GE}_2(R)$ can be brought to the form

$$A = [\alpha, \beta]P(a_1) \dots P(a_r), \quad (9)$$

using (3) and (5). If $a_i = 0$ for some $i \neq 1, r$, this relation can be shortened by (1), while if $a_i \in U(R)$ for $i \neq 1, r$, it can be shortened by using (8) and then (3)

to bring $C(\alpha)$ to the left. Thus in any ring R we can express any matrix A in $\mathbf{GE}_2(R)$ in the form (9), where $\alpha, \beta \in U(R)$, $a_i \in R$ and such that for $1 < i < r$, a_i is not 0 or a unit. Such an expression (9) for A is called a *standard form* for A . In the next section we shall see that in any ring R with 2-term weak algorithm there is a unique standard form for each $A \in \mathbf{GE}_2(R)$; more generally, this holds in any ring R with a degree-function such that R_0 is a field (cf. Cohn [66]). If R is a ring with unique standard form for $\mathbf{GE}_2$, then in any relation the left-hand side can be brought to standard form, and by uniqueness this must be I , i.e. any relation can be transformed to the trivial relation $I = I$ using only (1)–(3) and the obvious relation

$$[\alpha, \beta][\alpha', \beta'] = [\alpha\alpha', \beta\beta']. \quad (10)$$

This proves

PROPOSITION 7.1. *In any ring R with unique standard form for $\mathbf{GE}_2$ the relations (1)–(3) and (10) form a complete set of defining relations for $\mathbf{GE}_2(R)$. ■*

The sufficient condition given here is not necessary, since it does not hold for the ring of integers $\mathbf{Z}$ which, however, satisfies the conclusion.

Let us return to the expression (9); in order to obtain explicit formulae for the product of the P 's we define a sequence of polynomials p_n in non-commuting indeterminates $t_1, t_2, \dots$ with integer coefficients. The p_n are defined by the recursion formulae:

$$p_{-1} = 0, \quad p_0 = 1, \quad (11)$$

$$p_n(t_1, \dots, t_n) = p_{n-1}(t_1, \dots, t_{n-1})t_n + p_{n-2}(t_1, \dots, t_{n-2}). \quad (12)$$

For $n \geq 0$, the suffix of p_n just indicates the number of arguments, and so may be omitted when the arguments are given explicitly. We shall do so in what follows and only write the suffix when the arguments are omitted. We assert that

$$P(t_1) \dots P(t_r) = \begin{pmatrix} p(t_1, \dots, t_r) & p(t_1, \dots, t_{r-1}) \\ p(t_2, \dots, t_r) & p(t_2, \dots, t_{r-1}) \end{pmatrix}. \quad (13)$$

This is clear for $r = 1$ and in the general case follows by induction, since on writing $p_i = p(t_1, \dots, t_i)$, $p'_i = p(t_2, \dots, t_{i+1})$, we have

$$\begin{pmatrix} p_{r-1} & p_{r-2} \\ p'_{r-2} & p'_{r-3} \end{pmatrix} \begin{pmatrix} t_r & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} p_r & p_{r-1} \\ p'_{r-1} & p'_{r-2} \end{pmatrix}.$$

From the symmetry of (13) it is clear that the p 's may also be defined by (11)

and

$$p_n(t_1, \dots, t_n) = t_1 p_{n-1}(t_2, \dots, t_n) + p_{n-2}(t_3, \dots, t_n). \quad (14)$$

Either definition shows that p_n may be described as the sum of $t_1 t_2 \dots t_n$ and all terms obtained by omitting one or more pairs of adjacent factors $t_i t_{i+1}$. This mode of forming the p_n might be called the *leapfrog construction*; the number of terms in p_n is f_n , the n th Fibonacci number. The first few polynomials are:

$$p_1 = t_1, p_2 = t_1 t_2 + 1, p_3 = t_1 t_2 t_3 + t_1 + t_3, p_4 = t_1 t_2 t_3 t_4 + t_1 t_2 + t_1 t_4 + t_3 t_4 + 1.$$

Equivalently, p_n may be described as the polynomial part of the formal product (when expanded):

$$(t_1 + t_2^{-1})(t_2 + t_3^{-1}) \dots (t_{n-1} + t_n^{-1}) t_n. \quad (15)$$

When the t 's are allowed to commute, the p 's just reduce to the continuant polynomials, used in the study of continued fractions, and we shall use the term *continuant polynomial* also to describe the p 's in the general case.

From (5) it follows that the inverse of $P(t_1) \dots P(t_n)$ is

$$P(0)P(-t_n) \dots P(-t_1)P(0) \\ = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} p(-t_n, \dots, -t_1) & p(-t_n, \dots, -t_2) \\ p(-t_{n-1}, \dots, -t_1) & p(-t_{n-1}, \dots, -t_2) \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (16)$$

Clearly,

$$p(-t_1, \dots, -t_n) = (-1)^n p(t_1, \dots, t_n), \quad (17)$$

hence (16) reduces to

$$[P(t_1) \dots P(t_n)]^{-1} = (-1)^n \begin{pmatrix} p(t_{n-1}, \dots, t_2) & -p(t_{n-1}, \dots, t_1) \\ -p(t_n, \dots, t_2) & p(t_n, \dots, t_1) \end{pmatrix}. \quad (18)$$

Comparing this formula with (13), we obtain

LEMMA 7.2. *The continuant polynomials satisfy*

- (i) $p(t_1, \dots, t_n)p(t_{n-1}, \dots, t_2) - p(t_1, \dots, t_{n-1})p(t_n, \dots, t_2) = (-1)^n$ and
- (ii) $p(t_1, \dots, t_n)p(t_{n-1}, \dots, t_1) - p(t_1, \dots, t_{n-1})p(t_n, \dots, t_1) = 0$. ■

Of course, this lemma can also be proved directly by induction; it corresponds to the well known relations between successive convergents to a continued fraction.

We shall use these formulae to analyse comaximal relations in GE_2 -rings.

We recall that in any ring R a relation

$$ab' = ba'$$

is comaximal if there exist $c, d, c', d' \in R$ such that $da' - cb' = ad' - bc' = 1$. Suppose now that R is weakly 2-finite. Then by Prop. 0.6.5, there exists $A \in \text{GL}_2(R)$ such that

$$A = \begin{pmatrix} a & b \\ * & * \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} * & -b' \\ * & a' \end{pmatrix}. \quad (19)$$

Similarly, if R is a 2-Hermite ring, then every equation of comaximality $ad' - bc' = 1$, arises, by Th. 0.4.1, from a pair of mutually inverse matrices

$$A = \begin{pmatrix} a & b \\ * & * \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} d' & * \\ -c' & * \end{pmatrix}. \quad (20)$$

This leads to the following explicit formulae for comaximal relations in GE_2 -rings:

PROPOSITION 7.3. *Let R be a ring, $x_1, \dots, x_n, y, z \in R$ and $\alpha, \beta \in \text{U}(R)$.*

(i) *if R is a weakly 2-finite GE_2 -ring, then every comaximal relation in R has the form*

$$\alpha p(x_1, \dots, x_n) p(x_{n-1}, \dots, x_1) \beta = \alpha p(x_1, \dots, x_{n-1}) p(x_n, \dots, x_1) \beta; \quad (21)$$

(ii) *if R is a 2-Hermite GE_2 -ring, then every equation of comaximality can be written as*

$$\begin{aligned} \alpha p(x_1, \dots, x_n) p(x_{n-1}, \dots, x_2) \alpha^{-1} (-1)^n \\ - \alpha p(x_1, \dots, x_{n-1}) p(x_n, \dots, x_2) \alpha^{-1} (-1)^n = 1 \text{ and} \end{aligned} \quad (22)$$

(iii) *if R is a strong E_2 -ring, then every equation $rs = uv$, where r, s are not both 0 and u, v are not both 0, can be written*

$$yp(x_1, \dots, x_n) \cdot p(x_{n-1}, \dots, x_1) z = yp(x_1, \dots, x_{n-1}) \cdot p(x_n, \dots, x_1) z. \quad (23)$$

Proof. (i) In a weakly 2-finite ring every comaximal relation $ab' = ba'$ arises from a pair of mutually inverse matrices (19). Since R is a GE_2 -ring, we can write A in the form (9):

$$\begin{aligned} A &= [\alpha, \beta^{-1}] P(x_1) \dots P(x_n), \\ A^{-1} &= P(0) P(-x_n) \dots P(-x_1) P(0) [\alpha^{-1}, \beta], \end{aligned} \quad (24)$$

and now (21) follows on combining (19) and (24).

(ii) This follows similarly from (24) and the form (20) for matrices arising from a relation of comaximality.

(iii) Since a strong $\mathbf{E}_2$ -ring is a 2-fir, every relation $rs = uv$ is obtained from a comaximal relation $ab' = ba'$, in the form $ya \cdot b'z = yb \cdot a'z$. The result now follows by applying (i) and remembering that 2-firs are weakly 2-finite. ■

The significance of this proposition becomes clearer if we make the following definitions.

In any ring R , two elements a, a' are said to be *GL-related* if there exists $A \in \mathbf{GL}_2(R)$ such that A has a as the $(1, 1)$ -entry and A^{-1} has a' as the $(2, 2)$ -entry. If such A can be found in $\mathbf{GE}_2(R)$ we say that a, a' are *GE-related*. Thirdly, if A can be found in $\mathbf{E}_2(R)$, then a, a' are said to be *E-related*.

Clearly E-related elements are GE-related, and GE-related elements are GL-related. Moreover, by (13) and (18), a, a' are E-related if and only if there exist $x_1, \dots, x_n \in R$ such that $a = p(x_1, \dots, x_n)$, $a' = p(x_n, \dots, x_1)$ (the necessary change of sign in (18) can be effected by changing the sign of the second row of A and the second column of A^{-1}). By Cor. 0.6.4 we see that a, a' are GL-related if and only if they are stably associated. In $\mathbf{GE}_2$ -rings, we further have

PROPOSITION 7.4. *Let R be a $\mathbf{GE}_2$ -ring; then for any $a, a' \in R$, the following assertions are equivalent:*

- (a) a is GL-related to a' ,
- (b) a is GE-related to a' ,
- (c) a is E-related to an associate of a' .

Proof. The equivalence of (a) and (b) is immediate. When (b) holds, then by (9), $a = \alpha p(x_1, \dots, x_n)$, $a' = p(x_n, \dots, x_1)\beta^{-1}$, whence (c), and the converse is clear. ■

In order to compare E-related elements we make use of the following formulae, which follow from the leapfrog construction of continuants.

Let $t_1, \dots, t_n$ be any elements of a ring and α a unit, then if n is odd,

$$p(t_1\alpha, \alpha^{-1}t_2, t_3\alpha, \dots, \alpha^{-1}t_{n-1}, t_n\alpha) = p(t_1, \dots, t_n)\alpha, \quad (25)$$

$$p(\alpha t_1, t_2\alpha^{-1}, \alpha t_3, \dots, t_{n-1}\alpha^{-1}, \alpha t_n) = \alpha p(t_1, \dots, t_n), \quad (26)$$

while if n is even,

$$p(t_1\alpha, \alpha^{-1}t_2, t_3\alpha, \dots, t_{n-1}\alpha, \alpha^{-1}t_n) = p(t_1, \dots, t_n), \quad (27)$$

$$p(\alpha t_1, t_2\alpha^{-1}, \alpha t_3, \dots, \alpha t_{n-1}, t_n\alpha^{-1}) = \alpha p(t_1, \dots, t_n)\alpha^{-1}. \quad (28)$$

Further, an easy calculation shows that for any n :

$$p(t_1, \dots, t_n) = p_{n+1}(t_1, \dots, t_{n-1}, t_n - 1, 1). \quad (29)$$

This formula allows us to change the parity of n in any representation of an element by a continuant, as in the proof of the next result:

PROPOSITION 7.5. *In any ring R , if a is E-related to a' and α is a unit, then (i) αa is E-related to $\alpha a'$, (ii) αa is E-related to $\alpha a'$ and (iii) a is E-related to $\alpha^{-1} a' \alpha$.*

Proof. Let $a = p(x_1, \dots, x_n)$, $a' = p(x_n, \dots, x_1)$, then (i) follows by (25) if n is odd; if n is even, we can replace it by $n+1$, using (29) for a and the left-hand analogue of (29) for a' , and then applying the preceding argument. Similarly (ii) follows from (26). To prove (iii) we first ensure that n is even (using (29)) and then apply (27) and (28). ■

We note that this proposition may be used to give another proof of Prop. 7.4.

Exercises 2.7

1. Prove the identity $p(t_1, \dots, t_n) = p(t_1, \dots, t_r)p(t_{r+1}, \dots, t_n) + p(t_1, \dots, t_{r-1})p(t_r, \dots, t_n)$, for any n and $1 \leq r \leq n$.

2. (Brenner [55]) Show that the matrices $\begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$ generate a free group if α is a real number such that $\alpha \geq 2$.

3*. Find all pairs of algebraic integers α, β such that $(P(\alpha)P(\beta))^n = I$ for some n .

4*. Show that in $k\langle x, y, z, t \rangle$ the matrix $\begin{pmatrix} x & y \\ z & t \end{pmatrix}$ cannot be written as a product of elementary and diagonal matrices.

5*. Let $R = A_1(k)$ be the Weyl algebra on x and y over a field k of characteristic not 2. Show that the matrix $\begin{pmatrix} yx & x^2 \\ y^2 & xy \end{pmatrix}$ is invertible but is not in $\text{GE}_2(R)$. (Hint: Compute the inverse in $k(y)[x; 1, \cdot]$ and use the filtration by degree for the last part.)

6. (Helm [83]) Show that (1)–(3) and (10) form a complete set of defining relations for $\text{GE}_2(R)$ provided that for any two non-units a, b in R there exists a unit α such that $a + \alpha$ and $b + \alpha^{-1}$ are units.

7. Show that any local ring R is a **GE**-ring and that there is a relation of the form $[\alpha, \beta]P(a)P(b)P(c)P(d) = I$, where b, c are neither zero nor units, unless R is a field.

8. Let R be a totally ordered ring. Given $a_1, \dots, a_r \in R$ such that $a_i > 0$ for $1 \leq i \leq r$, show that $p(a_1, \dots, a_r) > 0$. Show that this still holds if $a_1 \geq 0$ and $a_i > 0$ for $2 \leq i \leq r$, provided that $r \geq 2$.

9*. Let R be a totally ordered ring such that $a > 0$ implies $a \geq 1$. Show that R has a unique form for $\mathbf{GE}_2$ (cf. Cohn [66']).

10*. Let D be a field with a central subfield k . Show that the only element of $D_k\langle x \rangle$ E-related to x is x itself. If $p \in k[x]$ is irreducible over k , show that p is an atom in $D_k\langle x \rangle$. (Hint: Note that $px = xp$ and use the natural homomorphism $D_k\langle x \rangle \rightarrow D[x]$ to show that x cannot be comaximal with any non-unit factor of p .)

11*. Show that $k[x]$ is 1-inert in $D_k\langle x \rangle$. (Hint: Use Ex. 10.)

12°. (Bergman [67]) In a free algebra, is every element E-related to a square itself a square (cf. Ex. 4.3.10)?

13°. If two elements of a ring are GL-related to a third, give an explicit proof that they are GL-related to each other (cf. 0.6).

14°. In a free algebra, does there exist an infinite family of pairwise E-related elements?

15*. In a free algebra R , show that GL-related elements are $\mathbf{E}'_2$ -related, where $\mathbf{E}'_2$ is the derived group of $\mathbf{E}_2(R)$ (cf. Cohn [66'], Th. 9.3).

2.8 The 2-term weak algorithm

We shall now develop the usual Euclidean algorithm, using the 2-term weak algorithm. In particular, everything that is said will apply to classical Euclidean domains in which the algorithm is defined relative to a degree-function, as well as to free algebras, which possess a weak algorithm by Prop. 4.2.

Let R be any filtered ring with 2-term weak algorithm. Given an equation

$$ab' = ba' \neq 0 \quad (1)$$

in R , if we choose $q_1 \in R$ such that $v(a - bq_1)$ is minimal, we find by an easy induction (as in 2.1) that $v(a - bq_1) < v(b)$. Thus

$$a = bq_1 + r_1, \quad v(r_1) < v(b), \quad (2)$$

where q_1, r_1 are unique, by Prop. 1.3. Substituting from (2) into (1), we find $r_1b' = (a - bq_1)b' = b(a' - q_1b')$. If we put $r'_1 = a' - q_1b'$, this may be written

$$r_1b' = br'_1. \quad (3)$$

By (3) and (2), $v(b) + v(r'_1) = v(r_1) + v(b') < v(b) + v(b')$, hence $v(r'_1) < v(b')$, so

there is complete symmetry (as we know there must be, by Cor. 3.2). It may happen that $r_1 = 0$, but by (3) this is so only if $r'_1 = 0$. If this is not the case, we can apply the same reasoning to (3) and thus obtain the familiar chain of equations of the Euclidean algorithm. More precisely, we obtain two such chains, one for left and one for right division:

$$\begin{aligned} a &= bq_1 + r_1, & a' &= q_1b' + r'_1, & r_1b' &= br'_1, \\ b &= r_1q_2 + r_2, & b' &= q_2r'_1 + r'_2, & r_2r'_1 &= r_1r'_2, \\ r_1 &= r_2q_3 + r_3, & r'_1 &= q_3r'_2 + r'_3, & r_3r'_2 &= r_2r'_3, \\ &\vdots & &\vdots & &\vdots \end{aligned} \quad (4)$$

Note that whereas the remainders r_i, r'_i on the two sides are in general distinct, the quotients q_i are the same. The degrees of the remainders decrease strictly,

$$v(b) > v(r_1) > \dots, \quad v(b') > v(r'_1) > \dots, \quad (5)$$

so the remainders must vanish eventually. Let n be the least integer such that $r_{n+1} = 0$. Since $r_{n+1}r'_n = r_n r'_{n+1}$, it follows that $r'_{n+1} = 0$; if we had $r'_k = 0$ for some $k \leq n$, then by symmetry $r_k = 0$, which contradicts the definition of n . Hence both chains in (4) end at the same step, i.e. r'_{n+1} is the first vanishing remainder of the right-hand division, and the last two rows of (4) read

$$\begin{aligned} r_{n-2} &= r_{n-1}q_n + r_n, & r'_{n-2} &= q_n r'_{n-1} + r'_n, & r_n r'_{n-1} &= r_{n-1} r'_n, \\ r_{n-1} &= r_n q_{n+1}, & r'_{n-1} &= q_{n+1} r'_n, & r_{n+1} &= r'_{n+1} = 0. \end{aligned} \quad (6)$$

From (4), (6) and the inequalities (5) we see that $v(q_i) > 0$ for $i = 2, 3, \dots, n+1$, while $v(q_1) > 0$ if and only if $v(a) > v(b)$.

Let us again write $P(x) = \begin{pmatrix} x & 1 \\ 1 & 0 \end{pmatrix}$, for any $x \in R$; then we can express equations (4) and (6) as follows:

$$(a \ b) = (r_n \ 0)P(q_{n+1})P(q_n) \dots P(q_1), \quad (7)$$

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = P(q_1)P(q_2) \dots P(q_{n+1}) \begin{pmatrix} r'_n \\ 0 \end{pmatrix}. \quad (8)$$

These equations make it evident that r_n is a common left factor of a and b and since the P 's are invertible, it is actually a highest common left factor (HCLF). Likewise r'_n is a HCRF of a' and b' . In particular, it follows that R is a strong $\mathbf{GE}_2$ -ring.

In the algorithm (4) the remainders, and hence the quotients are unique provided that the inequalities (5) are satisfied, and this will be so if we

perform the algorithm in the fewest possible number of steps. In this case, moreover, $q_2, \dots, q_{n+1}$ have positive degree, and when $v(a) > v(b)$, then q_1 also has positive degree. On changing notation and bearing in mind (13) of 2.7 we thus obtain

PROPOSITION 8.1. *Let R be a filtered ring with 2-term weak algorithm. If a, b are right commensurable elements of R , then there is an expression*

$$a = up(x_1, \dots, x_n), \quad b = up(x_1, \dots, x_{n-1}), \quad u, x_1, \dots, x_{n-1} \in R, \quad (9)$$

and if n is chosen minimal, then $x_1, \dots, x_{n-1}$ are non-zero non-units and the expression (9) for a, b is unique. Moreover, x_n is non-zero if and only if $v(a) \geq v(b)$, with equality if and only if x_n is a unit.

For the proof we need only observe that n can be reduced if x_i ($1 \leq i < n$) is 0 or a unit; the uniqueness now follows by the uniqueness of the remainders. ■

Let

$$ab' = ba' \neq 0 \quad (10)$$

be a dependence relation between a and b . If (9) holds, we have of course

$$a' = p(x_n, \dots, x_1)v, \quad b' = p(x_{n-1}, \dots, x_1)v$$

for some $v \in R^\times$, so the relation (10) is of the form (21) of 2.7 (with α, β replaced by u, v).

We remark that when $q_2, \dots, q_{n+1}$ are all of positive degree in the algorithm (4), then the degrees of the remainders must be strictly decreasing. Thus the expression (9) will also be unique if instead of prescribing n to be minimal we require $x_1, \dots, x_{n-1}$ to be non-zero non-units. Thus in every invertible matrix

$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we can reduce the first row by (7) and so write A uniquely as

$$\begin{pmatrix} \alpha & 0 \\ u & \beta \end{pmatrix} P(x_1) \dots P(x_n), \quad \text{where } x_1, \dots, x_{n-1} \in R^\times \setminus U(R),$$

and where $\alpha, \beta \in U(R)$ because A is invertible. Now

$$\begin{pmatrix} \alpha & 0 \\ u & \beta \end{pmatrix} = [\alpha, \beta] P(0) P(\beta^{-1}u),$$

hence

$$A = [\alpha, \beta] P(0) P(\beta^{-1}u) P(x_1) \dots P(x_n),$$

and this form is unique, with the proviso that the first two P 's are to be

omitted if $u = 0$, or transformed by (8) of 2.7 if u is a unit. Summing up, we have

PROPOSITION 8.2. *Any filtered ring R with 2-term weak algorithm is a strong E_2 -ring and the standard form for $GE_2(R)$ is unique. ■*

Exercises 2.8

1. Let R be a filtered ring with 2-term weak algorithm and let X be a generating set of $R^\times$, as a monoid. Write down a presentation for $R^\times$ as monoid, in terms of X .

2. Let $R = k\langle x_1, x_2, \dots \rangle$, where the free generating set is finite or countable. Show that $GL_2(R) \cong GL_2(k[x])$. (Hint: Find a k -linear map which preserves the defining relations, cf. Cohn [66].)

3. In a filtered ring R with 2-term weak algorithm, show that a, b are right comaximal if and only if there exist $u, v \in R$ and a unit λ such that $au - bv = \lambda$.

4. Let R be as in Ex. 3. Given two right comaximal elements a, b , show that the equation $ax - by = f$ has a solution (x, y) for any $f \in R$. More precisely, there exist $a', b', c', d' \in R$ such that $ab' = ba', ad' - bc' = 1$, $v(c') < v(a') = v(a)$, $v(d') < v(b') = v(b)$, and the general solution (x, y) has the form $x = d'f + b'g$, $y = c'f + a'g$, where g is arbitrary in R .

5. Let F be a free group. Show that the group algebra is a strong E -ring. (Hint: If F is free on X , show that every matrix over kF is associated to a matrix over $k\langle X \rangle$.)

6*. (Bergman [71]) Let R be a filtered ring with 2-term weak algorithm and S any monoid of (ring-) endomorphisms of R . Show that the set of fixed points under the action of S is a strong E_2 -ring.

7. In $R = k\langle x, y \rangle$ show that $(1 - xy)R \cap (1 - yx)R = 0$

8. (C. Reutenauer) In $R = \mathbb{Z}\langle X \rangle$ show that $p(a_1, \dots, a_r)$ and $p(a_r, \dots, a_1)$ have the same content. (Hint: Use Gauss's lemma, cf. Ex. 4.6 and apply induction on r to Prop. 7.3.)

2.9 The inverse weak algorithm

The classical division algorithm, as described for the polynomial ring $k[x]$ in 2.1, depended essentially on the degree-function $d(a)$ defined in this ring. If instead we use the order function $o(a)$ (cf. 0.10) we have an analogous statement, with the opposite inequality:

Given $a, b \in k[x]$ such that $o(b) < o(a) < \infty$, there exist q and a_1 such that

$$a = bq + a_1, \quad o(a_1) > o(a). \quad (1)$$

The process can be repeated, but since $\mathbf{N}$ has no maximal element, there is no reason why the process should terminate. However, we can pass to the completion of the ring $k[x]$, namely the formal power series ring $k[[x]]$. Here a repetition of the step (1) leads to a convergent process, and in fact one can make deductions about the divisibility in $k[[x]]$ which are quite similar to (and often stronger than) the consequences of the classical division algorithm. For example, the ring $k[[x]]$ displays such simple divisibility behaviour that its connexion with the algorithm (1) is usually forgotten. In the non-commutative case we do, however, obtain non-trivial results from the 'inverse algorithm'.

By an *inverse filtration* on a ring R we shall mean a function v such that

- I.1. $v(x) \in \mathbf{N}$ for $x \in R^\times$, $v(0) = \infty$,
- I.2. $v(x - y) \geq \min\{v(x), v(y)\}$,
- I.3. $v(xy) \geq v(x) + v(y)$.

If equality holds in I.3, we have an order function as defined in 0.10.

Writing $R_n = \{x \in R \mid v(x) \geq n\}$, we have the inverse filtration

$$R = R_0 \supseteq R_1 \supseteq R_2 \supseteq \dots, \quad R_i R_j \subseteq R_{i+j}, \quad \cap R_n = 0; \quad (2)$$

we can again construct the associated graded ring $\text{gr } R = \{\text{gr}_n R\}$, where $\text{gr}_n R = R_n/R_{n+1}$ ($n = 0, 1, \dots$). To give an example, let R be any ring with an ideal α such that $\cap \alpha^n = 0$, then (2) holds with $R_n = \alpha^n$; this is called the α -*adic* filtration on R .

Let R be any inversely filtered ring. If $\text{gr } R$ has (n -term) weak algorithm, the ring R is said to have (n -term) *inverse weak algorithm*. Of course, this notion can also be defined in terms of v -dependence, as before, if the necessary changes are made (which amount to reversing the inequalities in the definition of v -dependence). As before, the inverse weak algorithm is left-right symmetric.

For any inversely filtered ring R we define the *inverse dependence number* $\mu_v(R)$ as the greatest integer n for which the n -term inverse weak algorithm holds, or ∞ if no such n exists.

Given $a \in R^\times$, if $v(a) = n$, we write $\bar{a} = a + R_{n+1} \in \text{gr}_n R$; $\bar{0}$ is not defined. If R satisfies 2-term inverse weak algorithm, then $\text{gr}_0 R = R/R_1$ is a field (and hence R_1 is a maximal ideal). This means that the following general principle applies to such rings (in the case of the ordinary weak algorithm there is a corresponding principle, which we were able to use without stating it formally, because in that case our ring actually contained a field):

Exchange principle. Let R be an inversely filtered ring such that R/R_1 is a field. Given $a, a' \in R$ and $A \subseteq R$, if $v(a) \leq v(a')$ and a is right v -dependent on $A \cup \{a'\}$, then either a is right v -dependent on A or a' is right v -dependent on $A \cup \{a\}$.

For $a = 0$ this holds trivially; taking $a \neq 0$, we have by hypothesis, $a_i \in A \cup \{a'\}$, $b_i \in R$ such that

$$\bar{a} = \sum_1^n \bar{a}_i \bar{b}_i \quad \text{in } \text{gr } R;$$

thus $v(a_i) + v(b_i) = v(a)$. If no a_i equals a' , this shows a to be right v -dependent on A . Otherwise let $a' = a_1$, say; since $v(a) \leq v(a')$, we have equality here and so $v(b_1) = 0$, $\bar{b}_1 \in \text{gr}_0 R$ and by hypothesis $\bar{b}_1$ is a unit, say $\bar{b}_1 \bar{c} = 1$. It follows that

$$\bar{a}_1 = \bar{a}_1 \bar{b}_1 \bar{c} = \bar{a} \bar{c} - \sum_2^n \bar{a}_i \bar{b}_i \bar{c},$$

and so $a' = a_1$ is right v -dependent on $A \cup \{a\}$. ■

Clearly the principle holds whenever $\mu_v(R) \geq 2$; by assuming it explicitly we shall find that some of our results can be extended to arbitrary inversely filtered rings.

Let R be an inversely filtered ring. The chain (2) may be taken as the neighbourhood base at 0 for a topology on R , and we can form the *completion* of R , denoted by $\hat{R}$. Explicitly we have $\hat{R} = C/\mathfrak{N}$, where C is the ring of all Cauchy sequences in R and $\mathfrak{N}$ the ideal of sequences converging to 0 (cf. e.g. Cohn [77], p. 306). The ring $\hat{R}$ again has an inverse filtration and there is a natural embedding $R \rightarrow \hat{R}$ respecting the filtration; we shall usually take this embedding to be an inclusion. If this inclusion is an isomorphism, R is said to be *complete*; for example, $\hat{R}$ is always complete. In any case, the induced map of graded rings $\text{gr } R \rightarrow \text{gr } \hat{R}$ is easily seen to be an isomorphism, hence $\mu_v(\hat{R}) = \mu_v(R)$; thus R satisfies the n -term weak algorithm if and only if $\hat{R}$ does.

Frequently $\hat{R}$ is a local ring; we record a sufficient condition for this to happen:

PROPOSITION 9.1. *Let R be a complete inversely filtered ring. If R/R_1 is a field, then R is a local ring.*

Proof. It will be enough to show that any element not in R_1 is invertible. Let $a \in R \setminus R_1$, then $\bar{a}$ has a right inverse, i.e. there exists $b \in R$ such that $\bar{a}\bar{b} = 1$. Hence $ab = 1 - c$, where $c \in R_1$, so $c^n \in R_n$ and $\sum c^n$ is convergent. Now $ab(\sum c^n) = 1$; thus every element not in R_1 has a right inverse. Let $au = 1$,

then $u \notin R_1$ so $uv = 1$, and here $v = auv = a$, which shows that a has an inverse. ■

Before we can apply the inverse weak algorithm we still need two general reduction lemmas. To obtain the best results we have to take our rings complete:

LEMMA 9.2. *Let R be a complete inversely filtered ring and $a, a_1, \dots, a_n$ any elements of R . Then there exist $b_1, \dots, b_n \in R$ such that $v(b_i) \geq v(a) - v(a_i)$ and $a - \sum a_i b_i$ is either 0 or right v -independent of $a_1, \dots, a_n$.*

Proof. Assume that we can find $b_i^{(k)}$ such that $v(b_i^{(k)}) \geq v(a) - v(a_i)$ and

$$v(a - \sum a_i b_i^{(k)}) \geq v(a) + k;$$

if $a - \sum a_i b_i^{(k)}$ is right v -independent of the a_i , the result follows; otherwise we can subtract a right linear combination of the a_i to get a higher value of k . If this holds for all k , then $b_i^{(k)} \rightarrow b_i$ by completeness, and $a = \sum a_i b_i$. ■

We can now obtain an analogue of Lemma 2.2, but in a much stronger form:

LEMMA 9.3. *Let R be a complete inversely filtered ring such that R/R_1 is a field, and let $a_1, \dots, a_n \in R$. Then there exists $P \in \mathbf{P}_n(R)/\mathbf{Tr}_n(R)$ such that $(a_1, \dots, a_n)P = (a'_1, \dots, a'_r, 0, \dots, 0)$, where none of the a'_i is right v -dependent on the rest.*

Proof. Let $a'_1 = a_i$ be any element of least value. Applying Lemma 9.2, we can modify the a_j ($j \neq i$) by right multiples of a'_1 so as to make them 0 or right v -independent of a'_1 . This can only increase their values, so $v(a'_1)$ will still be minimal. Let a'_2 be of least value among the resulting elements other than a'_1 ; by another application of Lemma 9.2 we can make all the elements other than a'_1, a'_2 zero or right v -independent of a'_1, a'_2 . Continuing this process, we get a sequence $a'_1, a'_2, \dots, a'_n$ which will clearly be the image under a unitriangular matrix of a certain ordering of $a_1, \dots, a_n$. Since

$$v(a'_1) \leq v(a'_2) \leq \dots \leq v(a'_n)$$

by construction, all zeros will occur at the end. Now suppose that some non-zero a'_i is right v -dependent on the remaining a'_j . By the exchange principle we conclude that some a'_j is right v -dependent on those preceding it; but this contradicts the construction. ■

We now impose the weak algorithm to obtain an analogue of Th. 2.3.

THEOREM 9.4. *Let R be a complete inversely filtered ring with n -term inverse weak algorithm. Then any sequence of at most n elements of R can be reduced by a permutation matrix followed by an upper unitriangular matrix to a sequence of v -independent terms followed by zeros. In particular, R is a strong E_n -ring and an n -fir.*

Proof. The case $n = 1$ is clear; for $n > 1$ the result follows from Lemma 9.3, using the algorithm. ■

In particular, this shows that any complete inversely filtered ring with inverse weak algorithm is a semifir, but without completeness this need not hold (cf. Ex. 16).

'Weak v -bases' for ideals of inversely filtered rings can be defined as before (cf. 2.2) and constructed similarly; the definition, the construction and Lemma 2.6 can all be stated in $\text{gr } R$, using the right ideal of leading terms of members of the right ideal $\mathfrak{a}$ under construction. However, if B is a weak v -basis of the right ideal $\mathfrak{a}$, it is no longer true that B generates $\mathfrak{a}$, but merely a dense right ideal in $\mathfrak{a}$. As in Prop. 2.7, we thus obtain

PROPOSITION 9.5. *Let R be a complete inversely filtered ring with n -term inverse weak algorithm, and in the case $n = 1$ assume also that R/R_1 is a field. Then R has right and left ACC_n . ■*

In 2.2 we saw that a filtered ring with weak algorithm is a fir (Th. 2.4); this is not to be expected here. For a ring with inverse weak algorithm we find instead that it is a kind of 'topological fir'.

PROPOSITION 9.6. *Let R be a complete inversely filtered ring with inverse weak algorithm. Then any right ideal $\mathfrak{a}$ of R contains a right v -independent set B such that BR is dense in $\mathfrak{a}$.*

Proof. Let B be a weak v -basis for $\mathfrak{a}$. By the inverse weak algorithm it is right v -independent, and any $a \in \mathfrak{a}$ is right v -dependent on B . It follows that for any v , there exist elements $c_b^{(v)} \in R$ such that

$$v(a - \sum bc_b^{(v)}) > v,$$

where b runs over B , and the sum is finite for any given v . As $v \rightarrow \infty$, $c_b^{(v)} \rightarrow c_b$ and we obtain $a = \sum bc_b$, where the sum on the right may be infinite, but is convergent in the sense that only finitely many terms of value $\leq v$ occur for any v ; this just means that BR is dense in $\mathfrak{a}$. ■

The construction of the 'weak algebra basis' in the remarks preceding Th. 4.1 was essentially carried out in $\text{gr } R$ and so can be repeated here. But instead of finite sums we now must allow infinite convergent series of monomial terms, with coefficients chosen from a set of representatives of R/R_1 .

From Prop. 9.6 we have (as in 2.4)

THEOREM 9.7. *Let R be an inversely filtered ring, then R satisfies the inverse weak algorithm if and only if R/R_1 is a field K and R has a right v -independent weak v -basis for R_1 . In this case, if $\bar{K}$ is a set of representatives of K in R (with 0 represented by itself, for simplicity), and $X = \{x_i\}$ is a weak v -basis for R_1 , then any element of the completion $\hat{R}$ can be uniquely represented as a convergent series*

$$\sum x_I \alpha_I, \quad \text{where } \alpha_I \in \bar{K}, \quad (3)$$

where $I = (i_1, \dots, i_n)$ runs over all finite suffix sets and $x_I = x_{i_1} \dots x_{i_n}$. Conversely, all such expressions represent elements of $\hat{R}$. ■

Here (3) is convergent if and only if for each v , the set $\{I \mid v(x_I) \leq v \text{ and } \alpha_I \neq 0\}$ is finite. We remark that the finite sums (3) form a dense subgroup of $\hat{R}$.

The most important example of a ring with inverse weak algorithm is the power series ring in a number of non-commuting indeterminates over a field, and we indicate briefly how the inverse weak algorithm can be used to characterize such rings.

Let R be an inversely filtered K -ring, where K is a field. Then R/R_1 is a K -ring in a natural way; if it is equal to K , we shall call R a *connected* inversely filtered K -ring. This just means that $R = K \oplus R_1$.

For simplicity let us take k to be a commutative field. Let $R = k\langle X \rangle$ be the free k -algebra on a set X , and denote by v the usual order-function on R (given by the terms of least degree, as in 0.10). This defines an inverse filtration on R for which R is a connected k -algebra. More generally, we may assign different degrees to the various elements of X , thus we take $X = \cup X_i$ itself to be graded (the previous case is just the special case $X = X_1$); this is particularly useful when X is infinite. Then we have a function $v: X \rightarrow \mathbb{N}_+$ and the resulting completion $k\langle\langle X \rangle\rangle$ is the power series ring in the graded set X over k .

The power series ring (over a field) has the following characterization, analogous to the characterization of free algebras in 2.4.

PROPOSITION 9.8. *If k is a commutative field and R a complete inversely filtered connected k -algebra, then R is a power series ring in a graded set X over k if and only if R satisfies the inverse weak algorithm.*

Proof. If $R = k\langle\langle X \rangle\rangle$ has the inverse filtration as defined above, then $\text{gr } R = k\langle X \rangle$, and this satisfies the weak algorithm, hence R satisfies the inverse weak algorithm.

Conversely, if R satisfies the inverse weak algorithm, then by Th. 9.7, R_1 has a right v -independent weak v -basis X , hence the k -algebra generated by X is free on X and is dense in R , therefore R is equal to its completion $k\langle\langle X \rangle\rangle$. ■

COROLLARY 9.9. *In a power series ring $k\langle\langle X \rangle\rangle$ in a graded set X over k , any closed subalgebra satisfying the inverse weak algorithm is again a power series ring in a graded set over k . ■*

Here 'closed' refers, of course, to the topology: every convergent series of terms in the subalgebra has a sum in the subalgebra.

COROLLARY 9.10. *Let R be a complete inversely filtered connected k -algebra. Then R is a power series ring in a single variable over k if and only if $R \neq k$ and of any two non-zero elements of R , the one of higher value is right v -dependent on the other.*

Proof. Any ring R satisfying the hypotheses has inverse weak algorithm and so, by Prop. 9.8, is of the form $k\langle\langle X \rangle\rangle$ for some $X \neq \emptyset$. If X contains more than one element, say $x, y \in X$, $x \neq y$, then neither of x, y is right v -dependent on the other. This contradicts the hypothesis, hence $|X| = 1$. Conversely, $k[[x]]$ clearly satisfies the hypotheses. ■

The rings in Cor. 9.10 turned out to be commutative. If we consider generally commutative inversely filtered rings (or even Ore domains), we again find that the weak algorithm already follows from the 2-term weak algorithm. Let us define a right *principal valuation ring* (sometimes abbreviated as right PVR) as an integral domain R with a non-unit p such that every non-zero element of R has the form $p^r u$ ($r \geq 0$, $u \in U(R)$). When R is commutative, this clearly reduces to the usual definition of PVR as discrete rank 1 valuation ring. We have

PROPOSITION 9.11. *Let R be a complete inversely filtered right Ore domain with 2-term inverse weak algorithm. Then R has inverse weak algorithm and is either a field or a right principal valuation ring.*

Proof. If $R_1 = 0$, then $R = \text{gr}_0 R$ is a field. Otherwise take $p \in R_1$ such that $v(p)$ has least value. Given any $a \in R$, a and p are right linearly dependent, hence right v -dependent, and so by Lemma 9.2, either $a = pc$ or $p = ac$. If a

is a non-unit, then $v(a) \geq v(p)$, thus in the second case $v(c) = v(p) - v(a) \leq 0$, hence c is then a unit, so we have in any case $a \in pR$ (unless a is a unit). If we write

$$a = p^r u, \quad (4)$$

then $v(a) \geq rv(p)$, hence r is bounded, and if we choose it maximal in (4), then u must be a unit. Thus R is then a right PVR and this clearly satisfies the inverse weak algorithm. ■

A second important case (which actually includes the case of power series rings in a finite ungraded set) arises from rings with an α -adic filtration. Given a ring R and an ideal α of R , let us write

$$K = \text{gr}_0 R = R/\alpha, \quad M_n = \text{gr}_n R = \alpha^n/\alpha^{n+1}, \quad (5)$$

so that K is a ring and M_n is a K -bimodule. We first give a general condition for $\text{gr } R$ to be a tensor ring.

THEOREM 9.12. *Let R be a ring and α an ideal in R . If in the notation (5), M_1 is free as right K -module and*

$$\alpha \otimes_R \alpha \cong \alpha^2, \quad (6)$$

then the graded ring associated with the α -adic filtration, $\text{gr } R = \bigoplus M_n$, is a tensor ring:

$$\text{gr } R \cong K \langle M_1 \rangle.$$

Proof. Since $M_1 = \alpha/\alpha^2$ is free as right K -module, we can take a K -basis and lift it back to a subset Y of α . This gives a map $f: R^{(Y)} \rightarrow \alpha$, defined by $(a_y) \mapsto \sum y a_y$, such that the induced map $f \otimes_R K: K^{(Y)} \rightarrow M_1$ is an isomorphism. For each n we have the induced homomorphism

$$f \otimes_R R/\alpha^n: (R/\alpha^n)^{(Y)} \rightarrow \alpha/\alpha^{n+1}, \quad (7)$$

which we shall show by induction to be an isomorphism, for all n . This holds for $n = 0$, so assume that $n \geq 0$ and that (7) is an isomorphism. Applying $\otimes_{R/\alpha^n} \alpha/\alpha^{n+1}$, we find that

$$f \otimes \alpha/\alpha^{n+1}: (\alpha/\alpha^{n+1})^{(Y)} \rightarrow \alpha \otimes_R \alpha/\alpha^{n+1} \cong \alpha \otimes_R \alpha \otimes_R R/\alpha^n \cong \alpha^2 \otimes_R R/\alpha^n \cong \alpha^2/\alpha^{n+2},$$

where the last step is by (6). We thus have a commutative diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & (\alpha/\alpha^{n+1})^{(Y)} & \longrightarrow & (R/\alpha^{n+1})^{(Y)} & \longrightarrow & (R/\alpha)^{(Y)} \longrightarrow 0 \\ & & \downarrow & & \downarrow & & \downarrow \\ 0 & \longrightarrow & \alpha^2/\alpha^{n+2} & \longrightarrow & \alpha/\alpha^{n+2} & \longrightarrow & \alpha/\alpha^2 \longrightarrow 0 \end{array}$$

We have just seen that the first column is an isomorphism, the third column is an isomorphism by the induction hypothesis, hence the middle column is an isomorphism. Thus (7) holds with n replaced by $n+1$, and hence for all n .

Thus $\mathfrak{a}/\mathfrak{a}^{n+1}$ is free as right $R/\mathfrak{a}^n$ -module, and so the sequence

$$0 \rightarrow \mathfrak{a}^{n-1}/\mathfrak{a}^n \rightarrow R/\mathfrak{a}^n$$

remains exact under the operation $\mathfrak{a}/\mathfrak{a}^{n+1} \otimes_{R/\mathfrak{a}^n} - \cong \mathfrak{a} \otimes_R -$. Hence we have an embedding

$$\mathfrak{a} \otimes_R \mathfrak{a}^{n-1}/\mathfrak{a}^n \rightarrow \mathfrak{a}/\mathfrak{a}^{n+1}.$$

Here the image is $M_n = \mathfrak{a}^n/\mathfrak{a}^{n+1}$, so we have an isomorphism

$$\mathfrak{a} \otimes_R M_{n-1} \cong M_n.$$

Since $\mathfrak{a} \otimes_R - \cong \mathfrak{a}/\mathfrak{a}^2 \otimes_{R/\mathfrak{a}} - \cong M_1 \otimes_K -$, it follows that

$$M_1 \otimes_K M_{n-1} \cong M_n,$$

therefore by induction, $M_n \cong M_1^{\otimes n}$, and the result follows. ■

It is easy to derive conditions for the $\mathfrak{a}$ -adic filtration to satisfy the inverse weak algorithm:

COROLLARY 9.13. *Let R be a ring with an ideal $\mathfrak{a}$ such that (i) $R/\mathfrak{a}$ is a field, (ii) $\mathfrak{a} \otimes_R \mathfrak{a} \cong \mathfrak{a}^2$ and (iii) $\bigcap \mathfrak{a}^n = 0$. Then R satisfies the inverse weak algorithm relative to the $\mathfrak{a}$ -adic filtration.*

Proof. By Th. 9.12, $\text{gr } R \cong K\langle M_1 \rangle$, where $K = R/\mathfrak{a}$ is a field. Hence $\text{gr } R$ satisfies the inverse weak algorithm, by Th. 9.7, and so does R itself, by 2.3. ■

By Cor. A.5, $\mathfrak{a} \otimes_R \mathfrak{a} \cong \mathfrak{a}^2$ holds, for example, if $\mathfrak{a}$ is flat as right (or left) R -module. In 5.11 we shall see that in a fir R , any proper ideal $\mathfrak{a}$ satisfies $\bigcap \mathfrak{a}^n = 0$, and $\mathfrak{a}$, being free, is then flat, so Cor. 9.13 applies to a fir whenever (i) is satisfied.

We note a second case in which the conditions may be applied.

COROLLARY 9.14. *Let R be an augmented k -algebra, with augmentation ideal $\mathfrak{a}$, and let V be a complement of $\mathfrak{a}^2$ in $\mathfrak{a}$, as k -space. If*

$$\mathfrak{a} \otimes_R \mathfrak{a} \cong \mathfrak{a}^2,$$

then the natural map $f: k\langle V \rangle \rightarrow R$ is injective.

Proof. Write $S = k\langle V \rangle$, and denote the augmentation ideal by $\mathfrak{b}$. We have

$\cap b^n = 0$, so if f is not injective, there exists $c \in b^n$, $c \notin b^{n+1}$ for some n and $cf = 0$. But the induced map $b^n/b^{n+1} \rightarrow a^n/a^{n+1}$ decomposes into a sequence of isomorphisms

$$b^n/b^{n+1} \cong V^{\otimes n} \cong (a/a^2)^{\otimes n} \cong a^n/a^{n+1},$$

which is a contradiction because $cf \in a^{n+1}$. ■

We now come to an important relation between a ring with inverse weak algorithm and its completion.

THEOREM 9.15 (Inertia theorem). *Let R be a fir which is inversely filtered, with inverse weak algorithm, and let $\hat{R}$ be its completion. Then R is totally inert in $\hat{R}$.*

Proof. As before we define

$$R_1 = \{x \in R \mid v(x) > 0\}, \quad \hat{R}_1 = \{x \in \hat{R} \mid v(x) > 0\},$$

so that $\hat{R}_1$ is the completion of R_1 (or also its closure in $\hat{R}$). By the weak algorithm, $R/R_1 \cong \hat{R}/\hat{R}_1$ is a field K , and R_1 as a right ideal of R has a weak v -basis X , say. Then X is also a weak v -basis of $\hat{R}_1$, so each element r of $\hat{R}_1$ can be written uniquely as a convergent series

$$r = \sum x r_x \quad \text{where } r_x \in \hat{R}. \quad (9)$$

We note that $v(r) > \min\{v(r_x) \mid r_x \neq 0, x \in X\}$ and $r \in R_1$ if and only if all the r_x lie in R and are almost all zero.

Let $A \subseteq \hat{R}^n$, $B \subseteq {}^n\hat{R}$ be such that $AB \subseteq R$; we may enlarge A, B to be maximal in the sense that each consists of all the rows (columns) mapped into R by multiplication with the other. Then A is a left R -submodule of $\hat{R}^n$ and B a right R -submodule of ${}^n\hat{R}$. Further, the image $\bar{A}$ of A in $(\hat{R}/\hat{R}_1)^n = K^n$ is a left K -space of dimension s , where $s \leq n$. By $\hat{R}$ -column operations on A (and corresponding row operations on B) we may assume that A contains $e_1, \dots, e_s$, part of the standard basis for row vectors, while any component after the first s in any element of A is a non-unit in $\hat{R}$, i.e. has positive order.

Consider the case $s < n$. We claim that for all $a = (a_1, \dots, a_n) \in A$, $a_n = 0$. For if not, let us choose $a \in A$ so as to minimize $v(a_n)$. By adding left R -multiples of $e_1, \dots, e_s$ to a we may suppose that each a_i has positive order. Hence we can write

$$a = \sum x a_x, \quad \text{where } a_x \in \hat{R}^n. \quad (10)$$

We claim that all the a_x lie in A . For, given $b \in B$, we have $\sum x a_x b = ab \in R$, hence $a_x b \in R$ and by the maximality of A we find that $a_x \in A$. By (10) the n th

component of a_x must have lower order than a_n . This contradicts the minimality of $v(a_n)$, so $a_n = 0$ for all $a \in A$. We can now omit the final component in A, B and reach the desired conclusion by induction on n .

There remains the case $s = n$. Then $R^n \subseteq A \subseteq \hat{R}^n$, hence $B \subseteq {}^nR$. By symmetry we may also assume that the image $\bar{B}$ of B in ${}^n(\hat{R}/\hat{R}_1) = {}^n(R/R_1)$ has dimension n , and so is all of ${}^n(R/R_1)$. If $b \in B$ lies in the kernel of this map, then in terms of a left weak v -basis Y of R_1 we have $b = \sum b_y y$, where $b_y \in {}^nR$ and the b_y are almost all 0. As before we see that $b_y \in B$, hence

$${}^n(R/R_1) = \bar{B} = B/BR_1 \cong B \otimes_R R/R_1. \quad (11)$$

Now R is a fir, so B is a free right R -module, of rank n , by (11). If we take any right R -basis of B , we have n columns forming an $n \times n$ matrix P over R . By (11) the columns of $P \pmod{R_1}$ form a K -basis of ${}^nK (\cong {}^n(R/R_1))$, i.e. P is invertible $\pmod{R_1}$ and hence invertible over $\hat{R}$. Now $B = P({}^nR)$, so $P^{-1}B = {}^nR$ and hence $AP = R^n$, as we wished to show. ■

Combining the result with Cor. 9.13 (and using Cor. 5.11.4) we obtain

COROLLARY 9.16. *Let R be a fir and α an ideal such that R/α is a field. Then R is totally inert in its α -adic completion. ■*

The most important application of this result is to free algebras and their completions. In that case the theorem states that $k\langle X \rangle$ is totally inert in $k\langle\langle X \rangle\rangle$; of course, in this case no appeal to the results of 5.11 is necessary.

There are other subalgebras of $k\langle\langle X \rangle\rangle$ that we may sometimes wish to consider, in particular the algebra of all rational, and that of all algebraic power series.

The least subalgebra of $k\langle\langle X \rangle\rangle$ containing X and closed under taking inverses when they exist in $k\langle\langle X \rangle\rangle$ is denoted by $k_{\text{rat}}\langle X \rangle$ and is called the algebra of *rational power series*. It is clear that every element with non-zero constant term has an inverse, hence $k_{\text{rat}}\langle X \rangle$ is a local ring. In Ch. 7 we shall see that $k_{\text{rat}}\langle X \rangle$ consists of the components of the solutions u of the matrix equations

$$u = Bu + b, \quad B \in \mathfrak{M}_n(k\langle X \rangle), \quad b \in {}^nk\langle X \rangle, \quad (12)$$

where B has zero constant term.

An element f of $k\langle\langle X \rangle\rangle$ is said to be *algebraic* if it is of the form $f = \alpha + u_1$, where $\alpha \in k$ and u_1 is a component of the solution u of a system of equations

$$u_i = \phi_i(u, x), \quad i = 1, \dots, n, \quad (13)$$

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. In fact, every such system (13) has a unique solution by power series without constant terms, and the algebraic power series form a subalgebra, denoted by $k_{\text{alg}}\langle X \rangle$.

THEOREM 9.17. *Any system (13), where each ϕ_i is a polynomial without constant term or linear term in the u 's, has a unique solution with components of positive order in $k\langle\langle X \rangle\rangle$, and the set $k_{\text{alg}}\langle X \rangle$ is a subalgebra and a local ring, the non-units being the elements of positive order.*

Proof. Denote by $u_i^{(v)}$ the component of degree v in u_i , then by equating homogeneous components in (13) we find

$$u_i^{(v)} = \phi_i^{(v)}(u, x).$$

Here $\phi_i^{(v)}$ is the sum of the terms of degree v ; by hypothesis, for any term $u_j^{(\mu)}$ occurring in $\phi_i^{(v)}$ we have $\mu < v$, therefore the components $u_i^{(v)}$ are uniquely determined in terms of the $u_j^{(\mu)}$ with $\mu < v$, while $u_i^{(0)} = 0$, by hypothesis. Thus (13) has a unique solution satisfying $o(u_i) > 0$.

If $u_i = \phi_i(u, x)$ ($i = 1, \dots, m$), $v_j = \psi_j(v, x)$ ($j = 1, \dots, n$) are two such systems, then to show that $u_1 + v_1, u_1 v_1, \Sigma u_1^i = (1 - u_1)^{-1}$ are algebraic, we combine the above systems of equations for u_i, v_j with $w = \phi_1 + \psi_1, w = \phi_1 \psi_1, w = \phi_1 + u_1 w$, respectively. Thus we have indeed a subalgebra. To prove the last assertion, let f be an element of order 0 in $k_{\text{alg}}\langle X \rangle$, say $f = \alpha + u, \alpha \in k^\times$, then $f = \alpha(1 - v)$, where $v = -\alpha^{-1}u$, hence $f^{-1} = \alpha^{-1}(1 - v)^{-1}$, and this shows f to be a unit in $k_{\text{alg}}\langle X \rangle$. ■

We evidently have

$$k\langle X \rangle \subseteq k_{\text{rat}}\langle X \rangle \subseteq k_{\text{alg}}\langle X \rangle \subseteq k\langle\langle X \rangle\rangle.$$

It is not hard to see that all the inclusions here are strict; this is already clear when $|X| = 1$.

In 2.4, p. 105 we defined the notion of transduction for $k\langle X \rangle$, and it is clear how this extends to $k\langle\langle X \rangle\rangle$. Let us show that both $k_{\text{rat}}\langle X \rangle$ and $k_{\text{alg}}\langle X \rangle$ admit all transductions. It will be enough to examine the (left or right) cofactor of a generator x since the cofactor of a monomial can then be obtained by iteration.

Let u_1 be a component of a solution of (12), where B has zero constant term, and for a given $x \in X$ let $p \mapsto p'$ be the transduction 'left cofactor of x '. Clearly the constant term of u equals that of b . Let us modify u by subtracting its constant term, and modify b accordingly, then we get a system of the same

form (12), but in which the solution u has no terms of degree 0. If we now apply ' we find

$$u' = [Bu + b]' = Bu' + b',$$

hence u' satisfies a system as before, and so is again rational.

Next consider the algebraic case. Thus let u be a solution of (13), where ϕ_i contains no constant terms or linear terms in the u 's. Let us modify u by subtracting its linear terms, thus we may assume that u contains no terms of degree less than 2. Further, we may take the alphabet X to be finite. Each u_i has the form $u_i = \sum u_{ix}x$, where the u_{ix} have no constant terms, and we can write $\phi_i = \sum \phi_{ij}u_j + \sum \phi_{ix}x$, hence the equations (13) become

$$u_{ix} = \sum \phi_{ij}u_{jx} + \phi_{ix}.$$

If we express ϕ_{ij}, ϕ_{ix} in terms of the x and u_{ix} we obtain a system as before, except that some of the ϕ_{ix} may have constant terms. But this would lead to constant terms in u_{ix} and hence linear terms in u_i , contradicting the fact that there are no such terms. So we have indeed a system of the form (13) for u_{ix} , showing that the u_{ix} are again algebraic.

This result may be applied as follows:

PROPOSITION 9.18. *Any subring of $k\langle\langle X \rangle\rangle$ containing k and admitting inverses for all elements of zero order and all X -transductions is a semifir. In particular, $k_{\text{rat}}\langle X \rangle$ and $k_{\text{alg}}\langle X \rangle$ are semifirs.*

Proof. Let R be a subring satisfying the conditions and consider a relation in R :

$$\sum a_i b_i = 0, \quad \text{where } a_i, b_i \in R. \quad (14)$$

If some a_i has a non-zero constant term, it is a unit in R and we can trivialize (14). Otherwise let p be a monomial of shortest length occurring in any a_i , say in a_1 and let $f \mapsto f'$ be the transduction 'right cofactor of p ', then we have

$$\sum a'_i b_i = 0,$$

and here $a'_1 = \alpha \in k$, where α is the coefficient of p in a_1 ; moreover $a'_i \in R$ by hypothesis. So a'_1 is a unit and this allows b_1 to be expressed as a right linear combination of $b_2, \dots, b_r$, therefore (14) can be trivialized. ■

In Ch. 7 we shall see that $k_{\text{rat}}\langle X \rangle$ is actually a fir; by contrast $k_{\text{alg}}\langle X \rangle$ is not a fir when $|X| > 1$ (cf. Prop. 5.10.9).

Exercises 2.9

1. State and prove an analogue of Prop. 2.1 for the inverse dependence number.
2. Let R be any ring and $n \geq 2$; show that the following two properties are equivalent, where t is the trivial filtration:
 - (i) in a t -dependent family of at most n elements, any element a is t -dependent on those of degree $\leq t(a)$,
 - (ii) R is a local ring and an n -fir.
3. Investigate rings satisfying the inverse weak algorithm relative to the trivial filtration.
4. Find an extension of Lemma 9.3 to the case of an infinite family $\{a_i\}$.
5. Verify the inverse weak algorithm for the following rings and their completions:
 - (i) $\mathbb{Z}$ with $v = v_p$, the p -adic valuation,
 - (ii) $k\langle X \rangle$, where k is a field and v the degree of the least non-zero terms.
6. Show that every complete inversely filtered connected K -ring with an inverse division algorithm (1) is a power series ring in one indeterminate x , with commutation rule

$$ax = x\alpha^x + x^2\alpha^{2x} + \dots,$$

where α is an endomorphism of K and $(\delta_1, \delta_2, \dots)$ is a higher α -derivation (cf. 8.7).

7*. Let A be the group algebra over k of the free group on X and define an order function with values in $\mathbb{Z} \cup \{\infty\}$ in terms of the total degree: $d(x_1^{e_1} \dots x_n^{e_n}) = \sum e_i$. Show that A has a completion $\hat{A}$ relative to this function, and that $k\langle\langle X \rangle\rangle$ can be embedded in $\hat{A}$.

8. Let R be a complete inversely filtered ring with $\mu_v(R) \geq 2$. Show that $R^\times$ is a rigid monoid.

9. Let R be as in Ex. 8. If $a, b, a' \in R^\times$ are such that $v(a) \geq 0$ and $ab = ba'$, find $u, v \in R$ and $r \geq 0$ such that $a = uv$, $a' = vu$, $b = a'u = ua'^r$. Deduce that in a power series ring the eigenring of a non-zero element is an Artinian local ring.

10. If X is a finite set, the elements of $k\langle\langle X \rangle\rangle$ can be described as infinite power series $\sum x_i \alpha_i$, where x_i ranges over all monomials in X . Show that for infinite X this is no longer true, but when X is countable, there is a ring whose elements are all the (finite or) infinite power series in X , and which is obtained as the completion of $k\langle X \rangle$ by assigning suitable degrees to the elements of X . (This still holds for arbitrary sets X , if we allow transfinite degrees.)

11. (Jooste [71]) Let R be a complete inversely filtered ring with $\mu_v(R) \geq n$. Show that the kernel of any derivation of R is an n -fir. (Hint: Use Ex. 2.)

12*. (Jooste [71]) Let R be a complete inversely filtered ring with inverse weak algorithm, and let d be any derivation of R , with kernel N . Show that N is a semifir and that R is flat as an N -module.

13. Show that $k\langle\langle X \rangle\rangle$ is faithfully flat over $k_{\text{rat}}\langle X \rangle$ and $k_{\text{alg}}\langle X \rangle$ (it may be assumed that a ring R is left (right) faithfully flat over a subring S if R/S is flat as left (right) S -module, cf. Bourbaki [72]).

14*. Let G be the free group on a set Y . Show that if X is a set with a bijection $\phi: Y \rightarrow X$, then kG may be embedded in $k\langle\langle X \rangle\rangle$ by mapping $y \mapsto 1 - y\phi, y^{-1} \mapsto \sum (y\phi)^n$. Show further that this embedding is totally inert. (Hint: Verify that every matrix over kG is stably associated to a matrix over $k\langle Y \rangle$ and use the inertia theorem.)

15. Let $R = k\langle\langle x_1, x_2, \dots \rangle\rangle$ be the power series ring in countably many indeterminates and write α_n for the set of elements of order $\geq n$. Verify that each α_n is an ideal and that $R = \alpha_0 \supset \alpha_1 \supset \dots$. Show also that $\alpha_i \alpha_j \subset \alpha_{i+j}$. (Hint: For example, for $i = j = 1$ show that $\sum x_n^{n+1}$ belongs to the right- but not the left-hand side.)

16. Show that $k\langle\langle X \rangle\rangle \otimes_k k(t)$ satisfies the inverse weak algorithm but is not a semifir (unless $X = \emptyset$). What is the completion? (Hint: Look at the case $|X| = 1$.)

17°. Investigate inversely filtered k -algebras with inverse weak algorithm which are not connected.

18°. (Bergman) Let R be a free algebra and $\hat{R}$ its power series completion. If an element of R is a square in $\hat{R}$, is it associated (in $\hat{R}$) to a square of an element of R ?

2.10 The transfinite weak algorithm

In 2.1 we saw that the classical division algorithm can be defined for any ring R with a function from R to $\mathbb{N}$, or more generally, for any ordinal-valued function. However, in defining the weak algorithm we used the additivity of the values $v(x)$ in an essential way, and so were limited to $\mathbb{N}$. Nevertheless, the definition of the weak algorithm can be modified so as to apply to ordinal-valued functions. The resulting notion is not left-right symmetric (in contrast to the ordinary weak algorithm); this makes it suited for studying rings which lack left-right symmetry in some respect. Instead of filtrations we need to consider more general functions whose precise form is suggested by Th. 1.2.

Let R be a ring with a function w defined on it, satisfying the following conditions:

T.1. w maps $R^\times$ to a segment of the ordinals, $w(1) = 0$, $w(0) = -1$,

T.2. $w(a-b) \leq \max\{w(a), w(b)\}$,

T.3. $w(ab) \geq w(a)$ for any non-zero $b \in R$.

It follows from T.3 that R must be an integral domain. Moreover, Prop. 1.3

shows that when a division algorithm of the form $a = bq + r$, with $w(r) < w(b)$ exists, then the remainder r is unique.

We shall refer to a function w satisfying T.1–3 as a *transfinite degree function* on R . Given such a function w on R , we shall say that a family $\{a_i\}$ of elements of R is *right w -dependent* if some a_i is 0 or if there exist elements $b_i \in R$, almost all 0, such that

$$w(\sum a_i b_i) < \max_i \{w(a_i b_i)\}.$$

Otherwise the family $\{a_i\}$ is called *right w -independent*. An element $a \in R$ is said to be *right w -dependent* on a family $\{a_i\}$ if a is 0 or if there exist $b_i \in R$, almost all 0, such that

$$w(a - \sum a_i b_i) < w(a), \quad w(a_i b_i) \leq w(a) \text{ for all } i.$$

Otherwise a is *right w -independent* of the a_i . The notions of *left w -dependence* are defined correspondingly.

We shall only be concerned with *right w -dependence* and so we often omit the distinguishing adjective. We also note that a is w -dependent on the empty set if and only if $a = 0$.

Definition. A ring with a transfinite degree function w is said to satisfy the *right transfinite weak algorithm* if in any right w -dependent family $a_1, \dots, a_n$ such that $w(a_1) \leq \dots \leq w(a_n)$, some a_i is right w -dependent on $a_1, \dots, a_{i-1}$.

Let R be a ring with transfinite degree function w satisfying the right transfinite weak algorithm; it follows as before that the set $K = \{x \in R \mid w(x) \leq 0\}$ is a field.

Given a right ideal $\mathfrak{a}$ of R , let us well-order the elements of $\mathfrak{a}$ in any way so that the degree is increasing and omit any element right w -dependent on earlier elements. The resulting set B is a right w -independent basis of $\mathfrak{a}$, for it is clearly a generating set, and if $w(\sum_1^n a_i b_i) < \max_i \{w(a_i b_i)\}$, where $a_i \in B$ and the a_i occur in the order $a_1, \dots, a_n$, so that $w(a_1) \leq \dots \leq w(a_n)$, then by the transfinite weak algorithm, for some i ,

$$a_i = \sum_1^{i-1} a_j c_j + a'_i, \quad \text{where } w(a'_i) < w(a_i).$$

Hence a_i is right w -dependent on earlier elements, contradicting the fact that $a_i \in B$. Thus B is right w -independent, and *a fortiori* right R -linearly independent, hence a basis of $\mathfrak{a}$. To show that R has IBN, suppose that $\mathfrak{a}$ is generated by $a_1, \dots, a_n$ where $w(a_1) \leq \dots \leq w(a_n)$ say, and that these elements are linearly dependent. Then (1) holds for some i and we can replace a_i by a'_i . By induction on the n -tuple $(w(a_1), \dots, w(a_n))$ we find that $\mathfrak{a}$ can be generated by

fewer than n elements, hence by Th. 1.1.1, a has unique rank, and so R has IBN. Remembering Th. 1.2.3, we obtain

THEOREM 10.1. *Any ring with a transfinite degree function satisfying the right transfinite weak algorithm is a right fir, and hence satisfies right ACC_n for all n . Moreover, every right ideal has a right w -independent basis. ■*

To obtain examples of this construction we shall take certain monoid rings and we begin by looking at the appropriate monoids. A monoid S is said to satisfy right ACC_1 if any ascending chain of principal right ideals $a_1S \subset a_2S \subset \dots$ breaks off. Let S be any conical rigid monoid with right ACC_1 . On S we can define a partial preordering by left divisibility:

$$u \leq v \quad \text{if and only if } v = us \text{ for some } s \in S. \quad (1)$$

Since S is conical and has cancellation, this is in fact a partial ordering, and it satisfies the minimum condition, by the right ACC_1 on S . Moreover, for any $s \in S$, the lower segment generated by s , i.e. $\{x \in S \mid x \leq s\}$ is totally ordered, by rigidity, and hence, by ACC_1 is an ordinal number, which we denote by $w(s)$ and call the *transfinite degree* (defined by left divisibility). From the definition it is clear that we have

$$w(u) \leq w(v) \quad \text{implies } w(cu) \leq w(cv), \text{ for } u, v, c \in S, \quad (2)$$

$$w(b) \leq w(c) \quad \text{implies } w(bu) \leq w(cu), \text{ for } b, c, u \in S. \quad (3)$$

Now consider the monoid ring $R = kS$; we extend w to R by putting

$$w(\sum \lambda_s s) = \max\{w(s) \mid \lambda_s \neq 0\}.$$

As usual, by a *leading term* of $a \in kS$ we mean a term of highest degree. Then it is easily seen that (2), (3) still hold for $b, c \in kS, u, v \in S$. Moreover, it is clear that T.1, 2 hold, and T.3 also follows, for if $a = \sum \lambda_i u_i$, $b = \sum \mu_j v_j$ and $w(ab) < w(a)$, let u_1 be a leading term of a . Then $w(ab) < w(u_1) \leq w(u_1 v_j)$ for all j ; thus each term $u_1 v_j$ must be cancelled by a sum of other terms $u_i v_k$. Choose j so that v_j is not a right factor of any v_k ; we have $u_1 v_j = u_i v_k$ where by the choice of u_1 (and rigidity) $u_1 = u_i s$, $s \neq 1$, hence $v_k = sv_j$, a contradiction. Thus we have a transfinite degree function. We claim that it satisfies the right transfinite weak algorithm.

Let $a_1, \dots, a_n$ be a right w -dependent family, thus

$$w(\sum a_i b_i) < \max_i \{w(a_i b_i)\}; \quad (4)$$

if one of the a_i is 0, there is nothing to prove, so we may exclude that case. We suppose that the a_i are so numbered that

$$w(a_1) \leq \dots \leq w(a_n). \quad (5)$$

We have to show that some a_i is right w -dependent on $a_1, \dots, a_{i-1}$. By omitting superfluous terms we may again suppose that

$$w(a_1 b_1) = \dots = w(a_n b_n) = \alpha.$$

Since cancellation holds in S , the left cofactor of an element u in a product vu is well-defined, and for a given $u \in S$ we can define the left transduction $a \mapsto a^*$ in R , where for any $a \in R$, a^* is the left cofactor of u .

We shall need two properties of transductions.

LEMMA 10.2. *Let S be a conical rigid monoid with right ACC_1 and let $R = kS$ be the monoid ring with the transfinite degree function w defined by left divisibility. If $*$ denotes the left transduction with respect to an element $u \in S$, then*

(i) *for any $a, b \in R$,*

$$(ab)^* = ab^* + \text{terms of degree} < w(a) \text{ and} \quad (6)$$

(ii) *if $a, b, c \in R$, where c includes u among its leading terms, and $w(a) < w(bc)$, then $w(a^*) < w((bc)^*)$.*

Proof. (i) By linearity we need only check this when $b \in S$. If $b = tu$ for some $t \in S$, then $ab = atu$ and so $(ab)^* = at = ab^*$, hence (6) holds then. Now write $a = a_1 + a_2$, where a_1 consists of all terms in a such that $a_1 b$ has u as right factor, say $a_1 b = cu$. Then

$$(ab)^* = (a_1 b)^* + (a_2 b)^* = c;$$

now $b, u \in S$ and by the rigidity of S , either $b = tu$, a case that has been dealt with, or $u = tb$, for some $t \in S$, and we may assume that $t \neq 1$. Then $b^* = 0$ and $a_1 b = ct b$, so $a_1 = ct$ and $w(c) < w(a_1) \leq w(a)$, hence (6) is again satisfied.

(ii) Write $a = a_1 u + a_2$, where no term in a_2 has u as right factor; then $a^* = a_1$, $(bc)^* = b + \text{lower terms}$ and $w(a_1 u) \leq w(a) < w(bc) = w(bu)$, where we have used (2). Hence by (3) $w(a^*) = w(a_1) < w(b) = w((bc)^*)$. ■

We shall apply the lemma to (4), where (5) is assumed to hold. Among the leading terms of b_n pick one, u say, which is not a right factor of any other leading term of b_n ; it is clear that u cannot be a right factor of any non-leading term of b_n . Denote the coefficient of u in b_n by λ . If $*$ is the left transduction with respect to u , then by (6) we have

$$\sum a_i b_i^* = (\sum a_i b_i)^* + \text{terms of degree} < w(a_n). \quad (7)$$

Since $w(\sum a_i b_i) < w(a_n b_n)$ and $(a_n b_n)^* = \lambda a_n + \text{terms of lower degree}$, it follows by (ii) that $w((\sum a_i b_i)^*) < w(a_n)$, and with (7) this shows that

$$w(\sum a_i b_i^*) < w(a_n);$$

since $b_n^* = \lambda \neq 0$, this shows a_n to be right w -dependent on $a_1, \dots, a_{n-1}$, and so we obtain

THEOREM 10.3. *Let S be a conical rigid monoid with right ACC_1 . Then the monoid algebra kS satisfies the right transfinite weak algorithm with respect to the left divisibility order, and hence is a right fir. ■*

As an example to illustrate Th. 10.3 consider the monoid S generated by y, x_i ($i \in \mathbb{Z}$) subject to the relations

$$yx_i = x_{i-1}.$$

It is easily verified that the elements of S can be uniquely expressed in the form

$$x_{i_1} \dots x_{i_r} y^m \quad i_p \in \mathbb{Z}, \quad r, m \geq 0.$$

Therefore S is rigid, conical, with right ACC_1 ; so the monoid algebra $R = kS$ is a right fir, but it is not a left fir because the left ACC_1 does not hold in R :

$$Rx_0 \subset Rx_1 \subset Rx_2 \subset \dots$$

We note that y is right large in R , in fact we have $Ry \subseteq yR$, i.e. y is right invariant in R (cf. 4.2).

Exercises 2.10

1. Show that in a ring with right transfinite weak algorithm, any element of degree 0 is a unit and any element of degree 1 is an atom.

2. Let R be a ring with right transfinite weak algorithm. If $p \in R$ is an element of degree 1, and $\mathfrak{a}$ is a proper right ideal containing p , then $\mathfrak{a}$ has a weak w -basis including p .

3. Prove an analogue of Lemma 2.6 for the transfinite degree function.

4. Let S be the monoid on x_i, y_i, z ($i \in \mathbb{Z}$) with defining relations $x_i = zx_{i-1}$, $y_i = zy_{i-1}$. Show that the monoid ring $R = kS$ is a right fir in which $\cap z^n R = x_1 R + y_1 R$, and deduce that the lattice of principal right ideals containing $x_1 R$ is not complete.

5. Show that the only monoid with right cancellation in which (1) is a well-ordering is $\mathbb{N}$. What are the monoids with this property but instead of right cancellation having left cancellation? (cf. Cohn [61].)

6°. (Samuel [68]) Let R be a ring with transfinite algorithm. Does R have an integer-valued algorithm? What is the answer if R is commutative and the residue class field $(\text{mod } pR)$, for each atom p , is known to be finite?

7°. Show that the monoid algebra of a conical rigid monoid is a semifir.

2.11 Computation of the dependence number

Many situations require counter-examples in the form of an n -fir that is not an $(n+1)$ -fir. Usually it is easy to see that the proposed example is not an $(n+1)$ -fir, but it is less easy to prove that it is an n -fir and our aim in this section is to derive a result which in certain cases allows the dependence number to be estimated and hence shows the given ring to be an n -fir for appropriate n .

For example, to construct a ring which is weakly n -finite but not weakly $(n+1)$ -finite, one would take a k -algebra R on $2(n+1)^2$ generators a_{ij}, b_{ij} ($i, j = 1, \dots, n+1$), with defining relations which in matrix form can be written

$$AB = I, \quad A = (a_{ij}), \quad B = (b_{ij}). \quad (1)$$

Th. 11.2, below, shows R to be an n -fir and hence weakly n -finite. If it were an $(n+1)$ -fir, it would be weakly $(n+1)$ -finite, but that is not so, since in R we have

$$BA \neq I. \quad (2)$$

Intuitively it is clear that the relation (1) does not entail the relation $BA = I$, and this can be made into a rigorous proof using a normal form argument. Another method is to interpret A, B as endomorphisms of an infinite-dimensional k -space V . Let V have a basis (u_{vi}) indexed by $\mathbb{N} \times \{n+1\}$ and define

$$u_{vi}\alpha = \sum a_{ij}u_{v+1j}, \quad u_{vi}\beta = \begin{cases} \sum_j b_{ij}u_{v-1j} & \text{if } v > 0, \\ 0 & \text{if } v = 0. \end{cases}$$

Then it is easily verified that $\alpha\beta = 1, \beta\alpha \neq 1$; since R can be represented in $\text{End}(V)$ by mapping $A \mapsto \alpha, B \mapsto \beta$, it follows that $BA \neq I$. Actually it is enough to have one algebra with elements A, B satisfying (1), (2); then (1), (2) follow for the above matrix ring by the transfer principle described in Bergman [74].

Let R be a k -algebra (k a commutative field), generated by a set U ; thus if $F = k\langle U \rangle$, we have $R \cong F/\alpha$, where α is the ideal of relations holding in R . On F we have the usual degree-function $d(f)$. Let us write $a \mapsto \bar{a}$ for the natural homomorphism $F \rightarrow R$, then we can define a filtration v on R by

$$v(r) = \inf\{d(a) \mid \bar{a} = r\}.$$

Given $r, s \in R$, choose $a, b \in F$ such that $\bar{a} = r, \bar{b} = s$ and $v(r) = d(a), v(s) = d(b)$, then

$$\begin{aligned} v(r+s) &\leq d(a+b) \leq \max\{d(a), d(b)\}, \\ v(rs) &\leq d(ab) = d(a) + d(b). \end{aligned}$$

Hence we obtain

$$v(r+s) \leq \max\{v(r), v(s)\}, \quad v(rs) \leq v(r) + v(s);$$

clearly $v(1) = 0$, thus we have indeed a filtration on R .

Consider a set of defining relations for R . If one of them is linear, say $\sum \alpha_i u_i + \beta = 0$ ($u_i \in U$, $\alpha_i, \beta \in k$), we can use it to eliminate one of the generators u_i , because k is a field. Thus we may assume that there are no linear relations; now relations of higher degree may be replaced by quadratic relations by introducing further generators. This means that the defining relations for R may all be taken to be of the form

$$\sum_0^n x_v y_v = b, \quad (3)$$

where $x_v, y_v \in U$ and b is an expression in the generators with no term having x_v as a left factor. We shall assume that all defining relations are of the form (3) with the same value of n ; this can always be achieved by incorporating supernumerary terms in b .

Now each element f of R can be expressed as a linear combination of products of the generators. Such an expression is said to be in reduced form for the suffix 0, or briefly, 0-reduced if no term in f contains a factor $x_0 y_0$, for any of the defining relations (3). Any $f \in R$ can be expressed in 0-reduced form by writing down any expression for f and then applying the moves

$$x_0 y_0 \mapsto b - \sum_{v \neq 0} x_v y_v \quad (4)$$

arising from (3), for all the defining relations, wherever possible. Under fairly mild conditions (which need to be checked in any particular case) a 0-reduced expression for f is reached in a finite number of moves. In general there may be more than one such reduced form for a given f ; if for each $f \in R$ there is just one 0-reduced form, we shall call it the *normal form* of f (for the suffix 0), and we also say that a normal form for the suffix 0 exists. Similarly a μ -reduced form and a normal form may be defined for each suffix $\mu = 0, 1, \dots, n$, using instead of (4) the moves

$$x_\mu y_\mu \mapsto b - \sum_{v \neq \mu} x_v y_v. \quad (4_\mu)$$

The normal form for μ , when it exists, will be denoted by $[f]_\mu$.

It will also be necessary to index the defining relations rather more precisely; we shall index the rows $(x_0, x_1, \dots, x_n)$ occurring in the relations (3) by a set I and the columns $(y_0, y_1, \dots, y_n)^T$ by a set J . Now we can state the result to be proved as follows:

THEOREM 11.1. *Let R be a k -algebra generated by a set $U = X \cup Y \cup Z$, where $X = (x_{iv})$ ($i \in I, v = 0, 1, \dots, n$), $Y = (y_{vj})$ ($j \in J, v = 0, 1, \dots, n$) and X, Y need*

not be disjoint, though all the x_{iv} are distinct, as are all the y_{vj} . Assume that R has a complete set of defining relations indexed by some subset of $I \times J$:

$$\sum_{v=0}^n x_{iv} y_{vj} = b_{ij},$$

where b_{ij} is an expression in all the generators, of degree at most 2, with no terms of degree 2 having x_{iv} as a left factor. Further, assume that a normal form exists for each $v = 0, 1, \dots, n$ satisfying

N_v . For any $u, v, w \in U$, if uv is v -reduced, then the terms of degree 3 in $u[vw]_v$ are in normal form for v .

Then $\lambda_v(R) \geq n$; thus R satisfies the n -term weak algorithm and hence is an n -fir.

Here v is, of course, the filtration derived from the given presentation as indicated earlier. Sometimes it is convenient to assign degrees greater than 1 to the y_{vj} ; as long as the degree of y_{vj} is independent of v , this does not affect the truth of the theorem, and the proof remains essentially the same.

Proof. Let us write

$$H^r = \{f \in R \mid v(f) \leq r\}.$$

We have to prove that the n -term weak algorithm holds; thus if

$$f_1 g_1 + \dots + f_m g_m \equiv 0 \pmod{H^{r-1}}, \quad v(f_\alpha) + v(g_\alpha) = r, \quad (5)$$

where $m \leq n$, then some g_α is left v -dependent on the g 's of degree not exceeding $v(g_\alpha)$.

In order to reduce (5) we shall work with the normal form for the suffix 0 and therefore write $[f]$ for $[f]_0$ in what follows. We shall also need to make use of transductions: given $f \in R$ and $u \in U$, we shall denote the right cofactor of u in the expression $[f]$ by $(^u f)$ and the left cofactor by (f^u) .

We note that under the given hypotheses N_v actually holds for any $u, v, w \in R$. By linearity we need only verify this when u, v, w are products of generators. Now the reducibility of a product $f.g$, where g is a reduced monomial with leftmost factor $y_{\mu j}$, depends only on μ (whether $\mu = v$) and on j (whether there is a relation (i, j) such that f has a rightmost factor $x_{i\mu}$), but (4_μ) shows that the highest-degree terms resulting from a reduction have the same 'left-hand data' as the original word.

Now return to (5): if $v(f_\alpha) = 0$ for some α , the result is clear, so let $v(f_\alpha) > 0$ for all α . We take the terms of highest degree in f_α and g_α in normal form; let

$$f_\alpha \equiv \sum u(^u f_\alpha) \quad (u \in U) \quad (6)$$

be the expression of f_α in normal form (mod terms of lower degree than f_α). Inserting this expression in (5), we obtain a congruence which may be written as

$$\Sigma u[(u f_\alpha) g_\alpha] \equiv 0 \pmod{H^{r-1}}. \quad (7)$$

If $v(f_\alpha) > 1$ for all α , then the left-hand side of (7) has all its terms of highest degree in normal form, for there can be no reduction in $u[(u f_\alpha) g_\alpha]$ unless this was already possible in f_α , by N_0 . From the uniqueness of the normal form it now follows that the coefficient of each u in (7) must vanish, and going back to the original form of this coefficient, we find that

$$\Sigma (u f_\alpha) g_\alpha \equiv 0 \pmod{H^{r-v(u)-1}}. \quad (8)$$

Now the result follows by induction on r .

There remains the case where for some α , say $\alpha = 1$, $v(f_1) = 1$ and $f_1 g_1$ is not 0-reduced. This means that for some $i \in I$, say $i = 0$, f_1 contains a term x_{00} and $x_{00} g_1$ is not 0-reduced. Let us write

$$f_\alpha = x_{00}(x_{00} f_\alpha) + f'_\alpha, \quad (9)$$

and

$$g_\alpha = \Sigma y_{0j}(y_{0j} g_\alpha) + g'_\alpha. \quad (10)$$

Substituting from (9) and (10) into (5), we obtain

$$\Sigma_\alpha (x_{00}(x_{00} f_\alpha) + f'_\alpha) (\Sigma y_{0j}(y_{0j} g_\alpha) + g'_\alpha) \equiv 0 \pmod{H^{r-1}}.$$

Now reduce the terms of the first sum and equate the terms of degree r with x_{00} as left-hand factor. We may assume the f_α so numbered that $(x_{00} f_\alpha)$ has degree 0 for $\alpha = 1, \dots, s$ and positive degree for $\alpha > s$. Then, since

$$x_{00} y_{0j} = -\Sigma x_{0v} y_{vj} + b_{0j},$$

we have

$$\begin{aligned} & - \sum_{\beta=1}^s \sum_{v=1}^n \Sigma_j (x_{00} f_\beta) (x_{0v} y_{vj} (y_{0j} g_\beta) + b_{0j} (y_{0j} g_\beta)) + \sum_{\beta=1}^s x_{00} (x_{00} f_\beta) g'_\beta \\ & + \sum_{\gamma=s+1}^m x_{00} (x_{00} f_\gamma) g_\gamma + \sum_{\alpha=1}^m f'_\alpha g_\alpha \equiv 0 \pmod{H^{r-1}}. \end{aligned}$$

Equating cofactors of x_{00} we find

$$\sum_{\beta=1}^s (x_{00} f_\beta) g'_\beta + \sum_{\gamma=s+1}^m (x_{00} f_\gamma) g_\gamma \equiv 0 \pmod{H^{r-2}}. \quad (11)$$

Now by hypothesis $(x_{00} f_1)$ is a non-zero scalar, λ say; write

$$g_1^* = g_1 + \lambda^{-1} \left(\sum_{\beta=2}^s (x_{00} f_\beta) g_\beta + \sum_{\gamma=s+1}^m (x_{00} f_\gamma) g_\gamma \right), \quad (12)$$

then by (11), on putting $\lambda^{-1} (x_{00} f_\beta) (y_{0j} g_\beta) = h_{j\beta}$, we find

$$g_1^* = \sum_{\beta j} y_{0j} h_{j\beta} + \text{terms of lower degree}. \quad (13)$$

Now (12) represents an elementary transformation of the g 's which does not affect g_α for $\alpha \neq 1$. Hence if $f_1 g_1$ is not 0-reduced, we can by an elementary transformation bring g_1 to the form g_1^* given by (13) without disturbing the other g 's. Since the y 's are all distinct, $f_1 g_1$ is now μ -reduced for any suffix $\mu \neq 0$. Next take $\mu = 1$; as before there is a term $f_\alpha g_\alpha$ which is not 1-reduced. By what has been said, $\alpha \neq 1$ and after suitable renumbering of the f 's and g 's we may take $\alpha = 2$. The same argument now shows that there is an elementary transformation bringing g_2 to the form $g_2^* = \sum y_{1j} h_{j\beta} + \text{terms of lower degree}$, without disturbing the other g 's. Thus we may suppose by induction on t that g_β is replaced by

$$g_\beta^* = \sum y_{\beta-1 j} h_{j\gamma} + \text{lower terms } (\beta = 1, \dots, t-1); \quad (14)$$

then some $f_\alpha g_\alpha$ is not t -reduced. This means that $\alpha \neq 1, \dots, t-1$ and by renumbering the f 's and g 's we may take $\alpha = t$. As before, we can bring g_t to a form g_t^* given by (14) without affecting the g_β ($\beta \neq t$). This applies for $t = 1, 2, \dots, m$ and since $m \leq n$, we eventually reach a contradiction. So we are reduced to the first case and the result follows. ■

The hypotheses of Th. 11.1, although rather cumbersome to state, are quite natural ones holding in many cases, and when they do hold they are usually easy to verify. We shall do so in one important general case, but first we must introduce another concept, the matrix reduction functor.

The operation of forming $n \times n$ matrix rings may be regarded as a functor $\mathfrak{M}_n$ which associates with any ring R the ring R_n of $n \times n$ matrices over R . This functor has a left adjoint, the n -matrix reduction $\mathfrak{B}_n$:

$$\text{Rg}(R, \mathfrak{M}_n(S)) \cong \text{Rg}(\mathfrak{B}_n(R), S), \quad (15)$$

where Rg is the category of rings and homomorphisms. To define $\mathfrak{B}_n$ explicitly, let $\mathfrak{F}_n(R)$ be the ring generated by R and n^2 elements e_{ij} satisfying the equations of a matrix basis (1), 0.1 and in case R is a k -algebra (as will be the case below) also the relations $\alpha e_{ij} = e_{ij} \alpha$ for all $\alpha \in k$. By Prop. 0.1.1, $\mathfrak{F}_n(R)$ is itself an $n \times n$ matrix ring over a well-defined ring which we shall denote by $\mathfrak{B}_n(R)$, thus

$$\mathfrak{F}_n(R) = \mathfrak{M}_n(\mathfrak{B}_n(R)).$$

To verify the universal property of $\mathfrak{M}_n$, take any ring R and any homomorphism $f: R \rightarrow \mathfrak{M}_n(S)$ to an $n \times n$ matrix ring. By construction we have a natural homomorphism $\lambda: R \rightarrow \mathfrak{F}_n(R)$ and there is a unique homomorphism $f': \mathfrak{F}_n(R) \rightarrow \mathfrak{M}_n(S)$ which maps the matrix basis of $\mathfrak{F}_n(R)$ to that of $\mathfrak{M}_n(S)$ and the element $a\lambda$ of $\mathfrak{F}_n(R)$ to $af \in \mathfrak{M}_n(S)$. Hence $\lambda f' = f$, and by Prop. 0.1.1 there is a unique homomorphism $g: \mathfrak{M}_n(R) \rightarrow S$ such that $\mathfrak{M}_n(g) = f'$. Conversely, given $g: \mathfrak{M}_n(R) \rightarrow S$, we can retrace our steps to reach a unique homomorphism $f: R \rightarrow \mathfrak{M}_n(S)$. Thus (15) is established.

THEOREM 11.2. *Let R be a non-zero k -algebra. Then $\mathfrak{M}_{n+1}(R)$ has a filtration v for which the n -term weak algorithm holds.*

Proof. We shall denote the elements of R by a, b , etc., and the corresponding elements $a\lambda, b\lambda$ of $\mathfrak{F}_{n+1}(R)$ (which are matrices) by $(a_{\mu\nu}), (b_{\mu\nu})$, etc. We shall prove the result by verifying the n -term weak algorithm relative to the filtration induced on $\mathfrak{M}_{n+1}(R)$ by a presentation derived from one of R as follows. As we have seen, we can take the defining relations of R to be of the form $\sum a_i b_i = c$, or if we single out one pair a_1, b_1 and drop the suffix 1,

$$ab = c' \left(= c - \sum_{i \neq 1} a_i b_i \right).$$

As a matrix equation this reads

$$\sum a_{rv} b_{vs} = c'_{rs},$$

and this is of the form (3). If we regard the moves (4) as a 'straightening process', we see that every expression can be 0-reduced, by an induction on the number of 'bad' factors $a_{r0} b_{0s}$. To show that we actually get a normal form we have to check that for any expression containing two 'bad' factors the results of the two ways of reducing them can themselves be reduced to a common value, and here it is clearly enough to consider the case where the two factors in question overlap (this is in effect an application of the diamond lemma, cf. Bergman [78] or Cohn [81, 84]). Thus we have to examine products of the form

$$a_{r0} b_{00} c_{0s}.$$

If we add certain (uniquely reducible) terms, we obtain

$$\sum a_{r\mu} b_{\mu\nu} c_{\nu s},$$

and here we can reduce either the first pair of factors, then we get $\sum_v (ab)_{rv} c_{vs}$ and by a further reduction this gives $(abc)_{rs}$; or reducing the second pair first we get $\sum_\mu a_{r\mu} (bc)_{\mu s}$, which again reduces to $(abc)_{rs}$. Thus we get equality (essentially by the associative law) and this establishes the existence of a normal form for the index 0; the same argument holds for $\mu = 1, \dots, n$.

It remains to verify N_v ; again we need only consider N_0 . Thus uv is 0-reduced, so a typical term will be $a_{\lambda\mu}b_{\nu\rho}$, where μ, ν are not both 0 if ab occurs in a defining relation. If vw is not reduced, then $\rho = 0$ and w will have the form $c_{0\sigma}$, so $[b_{\nu 0}c_{0\sigma}] = (bc)_{\nu\sigma} - \sum_{\kappa \neq 0} b_{\nu\kappa}c_{\kappa\sigma}$ and when we multiply on the left by $a_{\lambda\mu}$ it remains 0-reduced because μ, ν are not both 0. Thus all the conditions of Th. 11.1 are satisfied and so the n -term weak algorithm holds. ■

Instead of interpreting the elements of R as square matrices of order $n+1$ we can also take them to be rectangular matrices. Thus if each generator u is interpreted as $n_u \times m_u$ matrix, where the numbers n_u, m_u are such that the defining relations, as matrix expressions, make sense, and if further, $m_u, n_u > n$, then the resulting ring again has n -term weak algorithm. The proof is essentially the same and so is left to the reader.

It is clear from Th. 11.2 that the ring described at the beginning of this section (with matrices A, B satisfying (1), (2)) is an n -fir. In the same way we can construct a ring over which every finitely generated module can be generated by n elements, but where n cannot be replaced by $n-1$. We need only take an $(n+1) \times n$ matrix A and an $n \times (n+1)$ matrix B with the relation

$$AB = I. \quad (16)$$

By Th. 11.2, the resulting ring $U_{n, n+1}$ is an $(n-1)$ -fir, hence weakly $(n-1)$ -finite, but from 0.2 we know that on writing $R = U_{n, n+1}$, we have $R^n \cong R^{n+1} \oplus P$ for some R -module P , hence $R^N \cong R^{N+1} \oplus P$ for $N \geq n$ and it follows that every finitely generated R -module is n -generated. However, R^n cannot be generated by fewer than n elements because R is weakly $(n-1)$ -finite. Here the remark (on rectangular matrices) following Th. 11.2 is not needed since the relations (16) may be written $A^i B_j = \delta_{ij}$, where A^i is the i th row of A and B_j is the j th column of B . However, this remark is needed for the next example.

Let A be $m \times n$ and B be $n \times m$ and consider the ring $R = V_{m,n}$ with generators $a_{i\lambda}, b_{\lambda i}$ satisfying

$$AB = I, \quad BA = I.$$

If $m < n$, R is an $(m-1)$ -fir but does not have IBN; in fact, it may be shown that it is of type $(m, n-m)$ (cf. Ex. 2).

Exercises 2.11

1. Prove Th. 11.2 when the generators are interpreted as rectangular matrices.
2. If $V_{m,n}$ (as defined in the text) is of type (h, k) , show that $m \geq h, n \equiv m \pmod{k}$. Use

the fact that $V_{m,n}$ is an $(m-1)$ -fir to show that $m = h$. (The fact that $n-m = k$ can be shown by a trace argument, cf. Cohn [66] or Corner [69].)

3. (Bowtell [67]) Let A be $m \times n$ and B be $n \times m$, where $n > m > 2$. Consider the ring R with defining relations $(AB)_{ij} = 0$ for $i \neq j$, $i, j = 1, \dots, n$. Show that R is an $(m-1)$ -fir but is not embeddable in a field. (Note that Th. 11.2 does not apply as it stands.)

4. (Klein [67]) Let A be $n \times n$ and consider the ring R with defining relation $A^{n+1} = 0$. Show that $A^n \neq 0$ and that R is an $(n-1)$ -fir.

5. Let A_ν, B_ν ($\nu = 1, 2, \dots$) be $n \times n$ matrices and consider the ring R with defining relations

$$A_{\nu-1} = A_\nu B_\nu \quad (\nu = 2, 3, \dots).$$

Show that R is a semifir and that each A_ν is non-invertible. Moreover, show that $\lambda_\nu(R) \geq n-1$ and deduce that R satisfies right ACC_{n-1} but not right ACC_n .

6. Let A be $n \times n$ and consider the ring R with defining relation $A^2 = A$. Show that R is an $(n-1)$ -fir but that there is an n -generator projective module which is not free.

7. Construct an $(n-1)$ -fir with an n -generator module which is stably free but not free.

8. (Montgomery [83]) Let k be a commutative field and $E = k(\alpha)$ a simple commutative extension of degree $n+1$, with minimal equation $x^{n+1} + a_1 x^n + \dots + a_{n+1} = 0$ for α ($a_i \in k$). Let R be the k -algebra generated by $x_0, x_1, \dots, x_n, y_0, \dots, y_n$ with defining relations obtained by equating powers of α in $(\sum x_i \alpha^i)(\sum y_j \alpha^j) = 1$. Show that R is weakly n -finite, but that $R \otimes_k E$ is not weakly 1-finite.

9. (Malcev [37]) Let A, B be 2×2 matrices such that $AB = \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix}$. Show that the algebra with these generators and defining relations is an integral domain, but not embeddable in a field. Show that this ring is not embeddable in a 2-fir. (Hint: Use the partition lemma.)

10. (M. Kireczi) Use Ex. 2 and Ex. 0.2.14 to show that there is no homomorphism $V_{1,r} \rightarrow V_{m,n}$ when $r > 1$, $1 < m < n$.

Notes and comments on Chapter 2

The familiar Euclidean algorithm occurs in Euclid, Book VII, Propositions 1 and 2, as a method of finding the highest common factor of two integers. The extension to polynomials was not undertaken until the 16th century, when Simon Stevin in his *Arithmetic* (1585), Book II, Problem LIII, uses it to find the HCF of two polynomials. He remarks that this application is probably new, since Pedro Nuñez, writing only a few years earlier (*Libro de Algebra*, 1567) attempts to treat the same problem, but does not get beyond a few generalities.

There is a very extensive literature dealing with the Euclidean algorithm in algebraic number fields; most of this does not concern us, but Motzkin [49], who determines the imaginary quadratic extensions admitting a Euclidean algorithm with respect to any function, introduces the notion of derived set and proves most of Th. 1.1. For the 2nd edition we have changed the notion of derived set following Samuel [71], cf. also Rodoski [80]. Propositions 1.3–5, giving conditions for the algorithm to be defined by a degree-function, are taken from Lenstra [74], who also determines all rings with a transfinite unique remainder algorithm, cf. 8.7. The first commutative examples of a (genuinely) transfinite Euclidean algorithm were found by Hiblot [75]. Th. 1.6 is due to Jacobson [34] and was found again by Cohn [61].

An interesting generalization of the Euclidean algorithm is considered by Leutbecher [78], who defines a ring R to be *quasi-Euclidean* if there is a function of two arguments $\theta: R^2 \rightarrow \mathbb{N}$ such that for any $(a, b) \in R \times R^*$ there exists $q \in R$ such that $\theta(-b, a - bq) < \theta(a, b)$. He shows that this is sufficient to derive many of the usual consequences of the Euclidean algorithm and in particular proves that a ring R is quasi-Euclidean if and only if R is a GE_2 -ring and every matrix $A \in R_2$ is right associated to a matrix with $(1, 2)$ -entry 0.

The notion of weak algorithm was introduced by Cohn [61, 63] as a simplified and abstract version of what was observed to hold in the free product of fields (cf. Cohn [60]). It was rediscovered by Bergman [67] (and greatly generalized in the coproduct theorems of Bergman [74] which, however, go beyond the framework of this book). Our presentation is based on all of these sources; in particular the original definitions have been modified as suggested by Bergman [67] so as to be closer to the usual notion of dependence. The n -term weak algorithm was introduced by Cohn [66, 69], where Prop. 2.7 was proved. The proof in the text, using weak v -bases, is due to Bergman [67]. For the use of transductions in the study of rational and algebraic power series see Nivat [68] and Fliess [70].

The dependence number as defined here seems more natural than the notion defined in the first edition (which was larger by one). For any ring R the *dependence number* $\lambda(R)$ may be defined independently of any filtration as the supremum of the $\lambda_v(R)$ as v runs over all filtrations of R . This is a positive integer or ∞ , defined for any ring; e.g. $\lambda(R) \geq 1$ if and only if R is an integral domain. Bergman [67] gives examples of rings R such that $\lambda_v(R)$ is finite but unbounded, as v ranges over the filtrations on R ; thus $\lambda(R) = \infty$, but R has no filtration for which the weak algorithm holds.

The weak algorithm for graded rings (and the material in 2.3) is taken from Bergman [67] except for Th. 3.3, which is new. The results of 2.4, characterizing free algebras by the weak algorithm, are due to Cohn [61], though the presentation largely follows Bergman [67]. The analogue of Schreier's formula in 2.5 was obtained by Lewin [69] as a corollary of the theorem that submodules of free modules are free, which Lewin proves for free algebras by a Schreier-type argument; it was also found by D. Lane (unpublished). Its extension to rings with weak algorithm first appeared in the 1st edition; the present version owes simplifications to Bergman and Dicks (cf. also Dicks [74]). Hilbert series (sometimes called Poincaré series, or also *gocha*, in the 1st edition) have been studied intensively for commutative graded rings over the last decade, particularly the cases of rationality, cf. Roos [81]. The problem of constructing all rings with weak algorithm was raised by Cohn [63] and solved by Bergman [67], whose presentation we follow in 2.6.

The HNN-construction (Higman, Neumann and Neumann [49]) starts from a group G with two isomorphic subgroups A, B via an isomorphism $\phi: a \mapsto a\phi$ and has

as its aim the construction of a group G_1 containing G and an element t such that $at = t \cdot a\phi$ ($a \in A$). In fact, G_1 is obtained by freely adjoining t to G subject to these relations. Of course, it has to be proved that G is embedded in G_1 , see Lyndon and Schupp [77] for a brief proof. The analogue for rings mentioned after Th. 6.2 was introduced and used by Macintyre [79]; for the field analogue see Cohn [71] and Ex. 7.5.19. A systematic study of HNN-constructions of rings has been undertaken by Dicks [83].

The Euclidean algorithm in a (non-commutative) integral domain with a division algorithm was treated by Wedderburn [32]; the presentation in 2.8, valid for any ring with 2-term weak algorithm, follows Cohn [63], and the description of $\text{GE}_2(R)$ in 2.7 is taken from Cohn [66]. Prop. 7.1 is improved by Menal [79] who shows that R has unique standard form for GE_2 if and only if R is 'universal' for GE_2 (i.e. the conclusion of Prop. 7.1 holds) and the subring generated by $U(R)$ is a field.

The inverse weak algorithm (2.9) was introduced by Cohn [62], where 9.6–11 are proved; 9.12–13 are taken from Cohn [70'], but the present version of 9.12, with the present simple proof, is due to Dicks [74]. The 'exchange principle' and 9.2–5 are due to Bergman [67], who also proves Th. 9.15 for free algebras in the case of 1-inertia. The case of total inertia was new in the 1st edition, though with some gaps in the proof, which have now been filled, following Cohn and Dicks [76]. A special case of Th. 9.15 was proved by Tarasov [67], who showed that an atom (without constant term) in the free algebra remained one in the power series ring. For an application of the inertia theorem to free radical rings cf. Cohn [73]. Prop. 9.18 answers a question raised by Fliess (in correspondence) and has not been published before.

The transfinite weak algorithm is taken from Cohn [69] and was suggested by a method of Skornyakov [65], used there to construct one-sided fir. The construction in the text is based on that of Skornyakov. In the first edition the right transfinite weak algorithm in the example was checked directly; the proof of Th. 10.3, applying to a whole class of monoid rings, is new in this edition. More generally, Kozhukhov [82] has shown that the monoid algebra kS is a right fir if S is a rigid monoid with right ACC_1 such that the subgroup of units in S is a free group and for any unit u and non-unit a , $au \in Sa$ implies that $u = 1$. Kozhukhov shows further that these conditions on S are necessary as well as sufficient for kS to be a right fir.

Th. 11.1 was proved in Cohn [69'], generalizing a particular case from Cohn [66]. Th. 11.2 is a special case of results in Bergman [74]. The k -algebra with $n \times n$ matrices A, B satisfying $AB = BA = I$ is an $(n-1)$ -fir by Th. 11.2; in fact, Bergman [74] has shown that it is a fir, though clearly not a GE_2 -ring. The matrix reduction functor, probably familiar to many, has not received formal treatment until recently, cf. Bergman [74], Cohn [72'', 79].

3. Factorization

After reviewing the familiar notion of a commutative unique factorization domain in 3.1, we study its generalization obtained by looking at the lattice of factors. The resulting notion of non-commutative UFD, 3.2 and 3, is mainly of interest for the factorization of full $n \times n$ matrices over 2n-firs; thus it can be applied to study factorization in free algebras. Another class, the rigid UFD's 3.4, generalize valuation rings and are exemplified by free power series rings. We also examine various direct decomposition theorems (3.5 and 6), but throughout this chapter we only consider square (full) matrices, corresponding to torsion modules over semifirs. The factorization of rectangular matrices, which is much less well developed, will be taken up in Ch. 5.

3.1 The commutative case

As is well known, a commutative integral domain is called a *unique factorization domain* (UFD for short) if it is atomic and any expression of an element as a product of atoms is unique except for the order of the factors and the presence of units. This definition makes it clear that unique factorization is a property of the multiplicative monoid of the ring, even though aspects of the ring itself are usually needed to establish it. We therefore restate the definition in terms of monoids.

Any commutative monoid S has a preordering by divisibility:

$$a|b \text{ if and only if } b = ac \text{ for some } c \in S. \quad (1)$$

If $a|b$ and $b|a$ in a cancellation monoid S , then

$$a = bu \text{ for some unit } u \text{ of } S, \quad (2)$$

i.e. a and b are *associated*. Clearly (1) is a partial ordering on S if and only if 1 is the only unit of S , i.e. when S is conical. With every commutative cancellation monoid S , having a group U of units, we associate a conical monoid S/U whose elements are the classes of associated elements of S . Since the relation (2) between a and b clearly defines a congruence on S , the set of these classes forms a monoid in a natural way.

A commutative cancellation monoid S will be called a *UF-monoid* if the

associated conical quotient monoid S/U is free commutative. With this definition it is clear that a commutative ring is a UFD if and only if its non-zero elements form a UF-monoid under multiplication. In studying unique factorization in commutative rings we can therefore limit ourselves to UF-monoids.

To state the condition for unique factorization in monoids succinctly, let us define a *prime* of a commutative monoid S as an element p of S which is a non-unit and such that

$$p \mid ab \text{ implies } p \mid a \text{ or } p \mid b.$$

Clearly any associate of a prime is again a prime. Further, a prime in a cancellation monoid is necessarily an atom. For if $p = ab$, then $p \mid a$ or $p \mid b$, say $a = pq$, then $p = ab = pqb$ and by cancelling p we have $qb = 1$, so b is a unit. The converse is false: an atom need not be prime, e.g. consider the monoid generated by a, b with the defining relation $a^2 = b^2$; here a is an atom but not prime. In fact the converse, together with a finiteness condition, ensures that we have a UF-monoid.

For later applications it is useful to consider a slightly more general case. An element c of a monoid S is said to be *invariant*, if

$$cS = Sc.$$

If every element of S is invariant, we say that S is *invariant*. Clearly this includes all commutative monoids and it shares with them the property that the preordering by left divisibility (which is defined in any monoid) coincides with the preordering by right divisibility. For if $a = bc$, then also $a = cb_1$ for some $b_1 \in S$ and $a = c_1b$ for some $c_1 \in S$. In an invariant monoid we can therefore define primes as in commutative monoids and can again associate a conical monoid S/U with S , whose elements are the classes of associated elements. An invariant cancellation monoid S will be called a *UF-monoid* if its associated quotient monoid S/U is free commutative. This clearly generalizes the previous definition.

A monoid S is said to satisfy *right ACC*₁ when ACC holds for 'principal right ideals' aS ; *left ACC*₁ is defined similarly, but these two conditions, of course, coincide in an invariant monoid. As in 1.2 we see that an invariant monoid satisfying ACC₁ is atomic, but not in general conversely, even in the case of rings, cf. Grams [74], Zaks [82]. Now we have

THEOREM 1.1 *In any invariant cancellation monoid S the following conditions are equivalent:*

- (a) S is a UF-monoid,
- (b) S satisfies ACC₁ and any two elements have an HCF,

- (c) S satisfies ACC_1 and any two elements have an LCM,
 (d) S is atomic and every atom of S is prime.

Proof. None of the conditions is affected if we pass to the associated conical quotient monoid, so we may assume that 1 is the only unit in S . It is clear that a free commutative monoid is conical and satisfies (b), so (a) $\Rightarrow$ (b). To prove (b) $\Rightarrow$ (c), assume (b). Given $a, b \in S$, there is a common multiple, viz. ab . Let m be a common multiple of a, b for which mS is maximal; if m' is another common multiple of a, b , we claim that $m | m'$, for otherwise the HCF, d say, of m and m' is again a common multiple and has the property $mS \subset dS$. Thus m is in fact the *least* common multiple as claimed.

To prove (c) $\Rightarrow$ (d), let p be an atom such that $p | ab$. Denote the LCM of p and a by m , then $m = ap_1 = pa_1$, say, and since ap is a common multiple, $ap = md = ap_1d$, say. Thus $p = p_1d$, but p is an atom, so either (i) $d = 1$ and $p = p_1$ or (ii) $p_1 = 1$ and $d = p$ (because 1 is the only unit). Case (i): $m = ap$ is an LCM; since $p | ab$, $a | ab$, we have $ab = me = ape$, therefore $b = pe$, $p | b$. Case (ii): $m = a$ is an LCM, hence $p | a$. Thus p is prime, and S is atomic by ACC_1 , so (d) holds.

Finally, to prove (d) $\Rightarrow$ (a), take distinct atoms a, b and let $ab = ba_1$, say, then $a | ba_1$ but $a \nmid b$. Since a is prime by (d), it follows that $a | a_1$, say $a_1 = au$. Thus $ab = bau$; by symmetry $ba = abv = bauv$, hence $uv = 1$, i.e. $u = v = 1$ and $ab = ba$. This shows that the monoid generated by the atoms is commutative. In fact it is the whole of S , for if not, let aS be an ideal which is maximal subject to a not being a product of atoms, then $a = bc$, where $aS \subset bS, cS$. Hence b, c are products of atoms and so a is too, which contradicts our choice of a .

The uniqueness proof follows a well-known pattern. If

$$c = p_1^{\alpha_1} \dots p_n^{\alpha_n} = p_1^{\beta_1} \dots p_n^{\beta_n}, \quad (3)$$

where the p_i are distinct atoms and $0 \leq \alpha_i < \beta_i$ say, then $p_1 | c$ but $p_1 \nmid p_2^{\alpha_2} \dots p_n^{\alpha_n}$, hence $p_1 | p_1^{\alpha_1}$, therefore $\alpha_1 > 0$. Now cancel p_1 in (3) and use induction on $\sum \alpha_i$ to obtain (a). ■

In any invariant cancellation monoid S , if a, b have an LCM, m say, then

$$aS \cap bS = mS. \quad (4)$$

Conversely, when (4) holds, then m is an LCM of a and b . Thus we obtain

COROLLARY 1.2. *An invariant cancellation monoid is a UF-monoid if and only if it satisfies ACC_1 and the intersection of any two principal ideals is principal.* ■

This result cannot be dualized, replacing $aS \cap bS$ by $aS \cup bS$, or even by the ideal generated by aS and bS .

It is clear how Th. 1.1 may be used to study factorization in rings whose non-zero elements form an invariant cancellation monoid. More generally, let R be any ring; an element $c \in R$ is said to be *invariant* in R if it is regular and $cR = Rc$. The set $I = I(R)^\dagger$ of all invariant elements of R is closed under multiplication and contains all units of R , in fact it is always an invariant cancellation monoid. This follows from the fact that if any two elements of the equation

$$c = ab$$

are invariant in R , then so is the third.

For an invariant element the sets of left and right multiples coincide; we can therefore define a *prime* in R as an invariant non-unit p satisfying

$$p \mid ab \text{ implies } p \mid a \text{ or } p \mid b.$$

This is to be distinguished from the notion of a prime in I , which will be called an *I-prime* for contrast. Thus every prime lies in I and hence is an *I-prime*, but not in general conversely. Similarly, an atom in I will be called an *I-atom*, so any invariant atom is an *I-atom*, but not always conversely. The ring R is said to have *unique factorization of invariant elements* if $I(R)$ is a UF-monoid. By applying Th. 1.1 we thus obtain

THEOREM 1.3. *Let R be any ring and $I = I(R)$ the set of its invariant elements. Then the following conditions are equivalent:*

- (a) *R is a ring with unique factorization of invariant elements,*
- (b) *R satisfies ACC on invariant principal ideals and any two invariant elements have an HCF in I ,*
- (c) *R satisfies ACC on invariant principal ideals and any two invariant elements have an LCM in I ,*
- (d) *$I(R)$ is atomic and every I -atom is I -prime. ■*

We note that e.g. (b) certainly holds if any two invariant elements have an HCF in R and this HCF is invariant. If we merely know that $a, b \in I$ have an HCF d in R , we cannot assert that d is invariant, though there is an important case in which this holds, namely when R is a 2-fir (cf. 6.1).

Frequently one needs to know the effect of localizing on unique factorization. Again we begin by setting out the problem in terms of monoids. If S is an invariant cancellation monoid and T is a suitable submonoid, then T is a

† The context will usually prevent its confusion with the idealizer.

right Ore set and we can form the localization S_T . The following criterion for S to be a UF-monoid generalizes a theorem of Nagata [57]:

THEOREM 1.4. *Let S be an invariant cancellation monoid and T a submonoid generated by certain primes of S and containing all units of S . Moreover, assume that S satisfies ACC_1 or that S is atomic, and that the localization S_T is a UF-monoid. Then S is itself a UF-monoid.*

Proof. By Th. 1.1 we need only verify that every atom of S is prime. Denote the canonical map $S \rightarrow S_T$ by $x \mapsto x'$ and let p be an atom of S . Either p' is an atom of S_T and hence a prime, or it is a unit; we treat these two cases separately.

(i) p' is a prime. If $p \mid ab$, say $pc = ab$, then $p'c' = a'b'$, hence $p' \mid a'$ or $p' \mid b'$, say the former. This means that

$$pe = ar, \quad \text{where } e \in S, r \in T.$$

No prime factor of r can be an associate of p , for otherwise p' would be a unit. Hence the prime factors of r divide e , and cancelling them one by one, we obtain an equation $pe_1 = a$, so $p \mid a$.

(ii) p' is a unit in S_T . Then $pq = r$, where $r \in T$ and $q \in S$. Again we can cancel the prime factors of r one by one, but this time we find that one of them is associated to p , for otherwise p would be a unit. Hence p is a prime. ■

Applied to rings, Th. 1.4 yields

THEOREM 1.5. *Let R be any ring, $I = \mathbf{I}(R)$ the set of invariant elements and T a submonoid of I generated by I -primes. If $\mathbf{I}(R)$ is atomic and the image of I in the localization R_T is a UF-monoid, then I is itself a UF-monoid. ■*

Taking R to be commutative, we obtain the following slight generalization of Nagata's theorem:

COROLLARY 1.6. *Let R be a commutative atomic integral domain. If T is a submonoid of $R^\times$ generated by certain primes and the localization R_T is a UFD, then R is a UFD. ■*

Exercises 3.1

1. Show that every Noetherian integral domain is atomic.
2. Let k be a commutative field and $R = k[x, y, z, t]$, $a = (xt - yz)R$. Show that the

ring R/a (the coordinate ring of a quadric) is an atomic integral domain, but not a UFD.

3. If an invariant cancellation monoid S satisfies ACC_1 and the join of any two principal ideals is principal, show that S is a UF-monoid. Show that the converse is false, by considering the multiplicative monoid of a suitable UFD.

4. If S is an invariant monoid and Q a normal submonoid, i.e. $aQ = Qa$ for all $a \in S$, define the quotient monoid S/Q and show that it is again invariant.

5. Let S be a cancellation monoid. If $a \in S$ is an invariant element, show that $b \in S$ is a left factor of a if and only if it is a right factor.

6. Let S be a cancellation monoid and T a submonoid generated by certain primes and containing all units of S . Show that T , as monoid, is invariant.

7. Let R be a commutative integral domain. If $a, b \in R$ have an LCM, show that they have an HCF. Show that the converse is false, by considering the subring of $\mathbb{Z}[x]$ consisting of all polynomials with even coefficient of x .

8. Show that in a commutative integral domain the LCM of any pair of elements exists if and only if the HCF of any pair of elements exists.

9*. (Amitsur [48]) Show that a commutative UFD with DCC on the ideals containing a fixed non-zero ideal is a principal ideal domain.

10*. In the ring of integral quaternions, show that

$$i(1 + ai + bj + ck) = (1 + ai + bj + ck)(qi + rj + sk),$$

where $q = 1 - (b^2 + c^2)u$, $r = (ab - c)u$, $s = (ac + b)u$ and $u = 2(1 + a^2 + b^2 + c^2)^{-1}$. Hence find all invariant elements.

11⁰. Let S be an invariant cancellation monoid and S/U the associated conical monoid. Is S/U necessarily commutative? (S/U is commutative whenever it is generated by primes, cf. the proof of Th. 1.1.)

12. Show that any invariant element in a simple ring is a unit.

3.2 Similarity in semifirs

In the further study of factorizations it is necessary to consider modules of the form R/aR ; for the moment we look more generally at cyclic modules R/a , where a is a right ideal, and examine when two such modules are isomorphic. The condition is closely related to the similarity of matrices (cf. 0.6) and it will be convenient to use the same term; the connexion between the two uses will soon become clear.

Thus we define two right ideals a, a' of a ring R to be *similar*, $a \sim a'$, if

$$R/a \cong R/a', \quad (1)$$

as right R -modules. It is clear from the definition that similar right ideals have isomorphic eigenrings. A criterion for similarity is given by

PROPOSITION 2.1. *Let a, a' be two right ideals in a ring R . Then a is similar to a' if and only if there exists $b \in R$ such that*

- (i) $a + bR = R$ and
- (ii) $a' = \{x \in R \mid bx \in a\}$.

In an abbreviated notation, (ii) may be expressed as $a' = b^{-1}a$.

Proof. Suppose that a is similar to a' and let $1(\bmod a')$ correspond to $b(\bmod a)$ in the isomorphism (1). Then b generates $R(\bmod a)$, hence (i) holds. Further, $x \mapsto bx$ and so $bx \in a$ if and only if $x \in a'$, i.e. (ii). Conversely, if (i) and (ii) hold, then $R/a \cong (a + bR)/a \cong bR/(a \cap bR) \cong R/a'$. ■

Note that whereas the relation of similarity is clearly symmetric, the criterion of Prop. 2.1 is not, so there are two ways of applying the result, once as it stands and once with a and a' interchanged. We also see from (ii) that a' is determined in terms of a and b . For example, if a is a maximal right ideal, then (i) holds provided that $b \notin a$, and it follows that the right ideal a' determined by (ii) is also maximal, because of (1). We state this as

COROLLARY 2.2. *If a is a maximal (proper) right ideal of a ring R and $b \in R \setminus a$, then the set $b^{-1}a = \{x \in R \mid bx \in a\}$ is a right ideal similar to a and hence also maximal.* ■

If b is right regular, (ii) takes on the form

$$a \cap bR = ba'. \quad (2)$$

In particular, when R is an integral domain and a is a proper right ideal, then $b \neq 0$ by (i) and so (ii) can be put in the form (2) in this case.

We now consider the case of principal ideals. In this case the similarity of aR and $a'R$ (for regular elements a and a') just corresponds to the similarity of a and a' , as defined in 0.6, Th. 0.6.2.

In a semifir it is possible to simplify this condition still further. Let R be any ring and $A \in {}^m R^n$; the matrix A is said to be *left full* if

$$A = PQ, \quad P \in {}^m R^r, Q \in {}^r R^n \text{ implies } r \geq m.$$

Right full matrices are defined similarly and a left and right full matrix is called a *full* matrix. Thus a matrix is full if and only if it is square, say $n \times n$ and cannot be written as a product of an $n \times r$ by an $r \times n$ matrix, where $r < n$.

We note that every non-zero element (in any ring) is full, as 1×1 matrix and the unit matrix of every size is full precisely if the ring has UGN. Over a field the full matrices are just the non-singular matrices (cf. 5.4), but in general there is no relation between full and regular matrices. However, over a semifir R , every right full matrix is right regular. For if $A \in {}^m R^n$ is right full and $AB = 0$ for some $B \neq 0$, then there exists $T \in \text{GL}_n(R)$ such that $AT = (A', 0)$, where $A' \in {}^n R^{n-1}$. But then $A = A'(I_{n-1}, 0)T^{-1}$, which contradicts the fact that A is right full. We note that this argument holds in any n -fir. We also note

PROPOSITION 2.3. *An $m \times n$ matrix (over any ring) with an $r \times s$ block of zeros cannot be left full unless $r + s \leq n$.*

Proof. Let A be $m \times n$, with an $r \times s$ block of zeros in the top right hand corner, say, then

$$A = \begin{pmatrix} A_1 & 0 \\ A_3 & A_4 \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ A_3 & A_4 \end{pmatrix},$$

where A_1 is $r \times (n-s)$ and A_4 is $(m-r) \times s$. This expresses A as a product of an $m \times (m+n-r-s)$ by an $(m+n-r-s) \times n$ matrix, so if A is left full, we have $m+n-r-s \geq m$, i.e. $r+s \leq n$. ■

It is clear that the conclusion still holds if the zeros do not occur in a single block, but are distributed over r rows and s columns, not necessarily consecutive. An $n \times n$ matrix with an $r \times s$ block of zeros, where $r+s > n$, is sometimes called *hollow*.

An $m \times n$ matrix A over any ring R is called *left prime* if in any equation

$$A = PQ, \quad \text{where } P \in R_m, Q \in {}^m R^n, \quad (3)$$

P necessarily has a right inverse; right prime matrices are defined correspondingly. Any right invertible matrix A is left prime, for if $A = PQ$ and $AB = I$, then $P \cdot QB = I$, so P is right invertible. For a square matrix the converse holds: a square matrix which is left prime is right invertible, as we see by taking $Q = I$ in (3). Thus, over a weakly finite ring any square matrix which is left prime is a unit. Any left prime matrix over a weakly finite ring is left full, for if $A \in {}^m R^n$ and $A = P_1 Q_1$, where $P_1 \in {}^m R^r$, $r < m$, then we can write $A = PQ$, where $P = (P_1, 0) \in R_m$ and since P has a column of zeros, it is not a unit in R and so has no right inverse; this shows that A cannot be left prime. So for a weakly finite ring we have the implications

$$\text{right invertible} \Rightarrow \text{left prime} \Rightarrow \text{left full.}$$

A matrix relation

$$PQ = 0, \quad \text{where } P \in {}^r R^n, Q \in {}^n R^s, \quad (4)$$

is called *comaximal* if P has a right inverse and Q has a left inverse; it is called *coprime* if P is left prime and Q is right prime; it is called *full* if P is left full and Q is right full. This terminology also applies to relations of the form $AB' = BA'$, since they may be brought to the form (4) by writing them as

$$(A, B) \begin{pmatrix} -B' \\ A' \end{pmatrix} = 0.$$

PROPOSITION 2.4. *Let R be an n -fir and $P \in {}^rR^n, Q \in {}^nR^s$ such that*

$$PQ = 0 \quad (5)$$

is a full relation. Then

$$r + s \leq n, \quad (6)$$

and when equality holds in (6), then by cancelling full matrices from the left and right in (5) we get a coprime relation, which is also comaximal.

The relation (6) is a special case of the law of nullity, which we shall meet again in Ch. 5.

Proof. Since R is an n -fir, we can choose $T \in \text{GL}_n(R)$ so that the first t columns of PT^{-1} are linearly independent, while the rest are 0. Then $PT^{-1} = (P', 0)$, $TQ = (Q', Q'')^T$ and (5) becomes $P'Q' = 0$. By the linear independence of the columns in P' we have $Q' = 0$. Moreover, PT^{-1} , like P is left full, hence so is P' . It follows that $r \leq t$; similarly Q'' is right full, so $s \leq n - t$, and (6) follows. When equality holds in (6), then $r = t$, $s = n - t$, hence P', Q'' are then square and so are full. Cancelling them, we obtain the relation $(I, 0)(0, I)^T = 0$, which is clearly comaximal. Hence the original relation (5) becomes comaximal after cancelling full factors on the left and right. ■

A relation (4) is said to be *proper* if $r + s = n$. We note that in any relation $AB' = BA'$ we have $i(A) - i(A') = i(B) - i(B')$, and the relation is proper if and only if either A, A' or equivalently B, B' have the same index. As we shall see later (5.10), if R satisfies left ACC, and right ACC, then we can cancel full matrices on the left and right of (5) so as to obtain a coprime relation; Prop. 2.4 shows that for proper relations (over an n -fir) we do not need ACC. The last part of Prop. 2.4 may be restated as

COROLLARY 2.5. *Any comaximal relation*

$$AB' = BA' \quad (7)$$

between matrices over a ring R is coprime. Conversely, over a semifir R , any

proper coprime relation is comaximal, and here (7) is proper if and only if A, A' have the same index. ■

Together with Prop. 0.6.6 this shows that two matrices A, A' over a semifir are stably associated if and only if they have the same index and occur in a coprime relation (7). The condition on the index is necessary since over $k\langle x, y \rangle$ we have

$$(x, y)(x, y)^T = (x, y)(x, y)^T,$$

a coprime relation which is not comaximal, and of course (x, y) is not stably associated to $(x, y)^T$.

Sometimes two elements a and a' are called 'similar' if the right ideals aR and $a'R$ are similar; as we have seen, for regular elements this condition is left-right symmetric. We then have three names: stably associated, similar and GL-related, which for elements of a ring mean the same thing. Nevertheless it is convenient to use them all, depending on the context. Thus 'stably associated' is mainly used for matrices, 'similar' for regular elements, while 'GL-related' is used in the context of relations in $\text{GL}_2(R)$, as in 2.7.

Another consequence of Prop. 2.4 is the fact that the product of full matrices over a semifir is full. We shall prove a slightly more general result:

PROPOSITION 2.6. *Over a semifir R any product of left full matrices is left full.*

Proof. Let $A \in {}^rR^m$ and $B \in {}^mR^n$ be left full and suppose that

$$AB = PC, \quad \text{where } P \in {}^rR^s, C \in {}^sR^n,$$

and $s < r$. We write $(B, C)^T = (B_1, C_1)^T Q$, where $B_1 \in {}^mR^t$, $C_1 \in {}^sR^t$, $Q \in {}^tR^n$; if t is chosen to be minimal, then $(B_1, C_1)^T$ is right full, hence the relation $AB_1 = PC_1$ is full, because A is left full. Hence by Prop. 2.4, $r+t \leq m+s$, so $m-t \geq r-s > 0$, i.e. $t < m$, but this contradicts the fact that B is left full. ■

Finally, we describe right Ore sets in 2-firs, using similarity:

PROPOSITION 2.7. *Let R be a 2-fir and S a submonoid of $R^\times$ such that (i) each element of S is right large, (ii) $ab \in S$ implies $b \in S$ and (iii) if $a \in S$, any element similar to a is in S . Then S is a right Ore set in R .*

Proof. By definition (0.8) we need only show that for any $a \in R$, $u \in S$, $aS \cap uR \neq \emptyset$. When $a = 0$, this is clear, so let $a \neq 0$, then $aR \cap uR \neq 0$, by (i), hence $aR + uR = dR$, say $a = da_1$, $u = du_1$. Here $u_1 \in S$ by (ii) and there is a comaximal relation $a_1u' = u_1a'$ with u' similar to u_1 , hence $u' \in S$ by (iii). Thus $au' = ua'$ is the required common right multiple. ■

Exercises 3.2

1. (a) Let a, b be two right ideals in a ring R such that $a + b = R$. Verify that the sequence

$$0 \longrightarrow a \cap b \xrightarrow{\lambda} a \oplus b \xrightarrow{\mu} R \longrightarrow 0$$

is split exact, where $\lambda(x) = (x, x)$, $\mu(x, y) = x - y$.

(b) Let a, a' be similar right ideals of R and assume that the criterion of Prop. 2.1 holds with a regular element b . Show that

$$a \oplus R \cong a' \oplus R.$$

(c) Give an example of a ring containing similar right ideals for which this isomorphism does not hold.

(d) Give examples of similar right ideals that are not isomorphic.

2. Show that in a 2-fir any right ideal similar to a principal right ideal is itself principal. More generally, show that this holds in any 2-Hermite domain. (Hint: Use Prop. 0.4.2.)

3. (Bowtell [67]) If $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\begin{pmatrix} d' & -b' \\ -c' & a' \end{pmatrix}$ are two mutually inverse matrices, show that $aR \cap bR = ad'bR + bc'aR$. Deduce that if R is an integral domain in which the sum of any two principal ideals with non-zero intersection is principal, then this intersection is again principal.

4. Show that any element GL-related to a zero-divisor is itself a zero-divisor, and that any element GL-related to a unit is a unit. Explicitly, if a, a' occur in mutually inverse matrices as in Ex. 3, and a is a unit, then $a'^{-1} = d - ca^{-1}b$.

5. In a weakly 1-finite ring, show that any element GL-related to 0 is 0. Is there a converse?

6. In a 2-fir let $c = ab$. Given c' similar to c , find a' similar to a and b' similar to b such that $c' = a'b'$. Does this hold in more general rings?

7. Let R be a ring and R_∞ the ring of infinite matrices over R which differ from a scalar matrix only in a finite square. Show that any two regular matrices in R_∞ that are similar are associated.

8*. (Fitting [36]) Let R be any ring and $a, b \in R$. Show that aR and bR are similar if and only if the matrices $A = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} b & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ are associated.

9. Let k be a commutative field of characteristic 2 and α the endomorphism which sends each element to its square. Show that in the skew polynomial ring $k[x; \alpha]$ there are just two similarity classes of elements linear in x . Generalize to the case of characteristic $p > 2$.

10⁰. Examine similarity classes of polynomials of higher degree in the ring of Ex. 9; also the case of more general skew polynomial rings.

11. Show that in any skew polynomial ring, similar elements have the same degree.
12. In a free algebra, show that if two homogeneous elements are similar, then they are associated.
- 13*. (Bergman) In the complex-skew polynomial ring, show that $x^4 + 1$ can be written as a product of two atomic factors in infinitely many different ways. By considering the factors, deduce the existence of a similarity class of elements which contains infinitely many elements that are pairwise non-associated. (See also Williams [68].)
14. In a 2-fir R , if $au - bv = c$ and a, c are right comaximal, show that $ua' - b'w = c'$, where a' is similar to a , b' to b and c' to c . (Hint: Let $ac' = ca'$ be a comaximal relation; multiply the given equation by a' on the right and use the fact that $aR \cap bR$ is principal as a right ideal.)
15. Show that over a commutative ring the determinant of any non-full matrix is zero.
16. Show that in any ring the intersection of all maximal right ideals similar to a given one is a 2-sided ideal. (Hint: Use Cor. 2.2.)

3.3 Factorization in matrix rings over semifirs

In order to study factorization in a principal ideal domain R it is convenient, in considering an element $c \in R^\times$, to look at the module R/cR ; this is a module of finite composition length and the Jordan-Hölder theorem immediately yields the fact that R is a UFD, at least in the commutative case. In the non-commutative case this continues to hold, if we define a non-commutative *unique factorization domain* as an integral domain which is atomic and is such that any two complete factorizations (into atoms) of an element c are *isomorphic*, in the sense that they have the same length, say $c = a_1 \dots a_r = b_1 \dots b_r$, and there is a permutation $i \mapsto i'$ of $1, \dots, r$ such that a_i is similar to $b_{i'}$.

For principal ideal domains this also covers the notion of matrices, because every matrix is then associated to a diagonal matrix, as is well known (cf. 8.1), but this is no longer so for firs. We therefore need to consider, together with the ring R , the $n \times n$ matrix ring R_n .

Let R be an n -fir; an R -module M is said to be an n -torsion module if it is finitely presented by a full $n \times n$ matrix. By a *torsion module* we understand an n -torsion module for some n ; in this case the ground ring is assumed to be an n -fir, to ensure that full $n \times n$ matrices are regular. It follows that every torsion module M has a free presentation of length 2 and $\chi(M) = 0$. Moreover, there is no submodule M' satisfying $\chi(M') < 0$, for such a submodule would

correspond to a factorization $A = PQ$ of the presenting matrix of M , where P is a matrix of fewer than n columns.

Any n -torsion module M has a presentation

$$0 \rightarrow R^n \rightarrow R^n \rightarrow M \rightarrow 0.$$

Given any other presentation of M (by a free module of rank n)

$$0 \rightarrow Q \rightarrow R^n \rightarrow M \rightarrow 0,$$

we have, by Schanuel's lemma, $R^n \oplus Q \cong R^{2n}$, so if R is a $2n$ -fir, then $Q \cong R^n$.

Let us denote by $\mathcal{T}_R$ the category of all right torsion modules over a semifir R and all homomorphisms between them, and the full subcategory of all n -torsion modules by $\mathcal{T}_R^n$; similarly we write ${}_R\mathcal{T}$ (${}_R\mathcal{T}^n$) for the category of all left (n -)torsion R -modules. Of course $\mathcal{T}_R^n$ and ${}_R\mathcal{T}^n$ can be defined more generally over an n -fir. We begin by noting a duality between them.

THEOREM 3.1. *Let $n \geq 1$. For any n -fir R , the categories $\mathcal{T}_R^n$ and ${}_R\mathcal{T}^n$ are dual, i.e. there exist contravariant functors $C: \mathcal{T}_R^n \rightarrow {}_R\mathcal{T}^n$, $D: {}_R\mathcal{T}^n \rightarrow \mathcal{T}_R^n$ such that CD and DC are each naturally isomorphic to the identity functor. Hence, over a semifir R the categories $\mathcal{T}_R$ and ${}_R\mathcal{T}$ are also dual.*

Proof. By passing to the $n \times n$ matrix ring (cf. 0.1) we may take $n = 1$. Given a regular element a of R , we define $C(R/aR) = R/Ra$; for any R -homomorphism

$$f: R/aR \rightarrow R/bR \quad (1)$$

we recall that f is determined by an element $c \in R$ such that for some $c' \in R$,

$$ca = bc'. \quad (2)$$

Conversely, any pair (c, c') satisfying (2) defines a homomorphism (1) by the rule

$$x(\text{mod } aR) \mapsto cx(\text{mod } bR).$$

Now c is determined by f up to an element of bR ; thus if $c_1 = c + bz$, then by (2), $c_1a = ca + bza = b(c' + za)$. This shows c' to be determined by f up to an element of Ra , and by (2) it defines a unique R -homomorphism

$$C(f): R/Rb \rightarrow R/Ra.$$

The correspondence $f \mapsto C(f)$ is clearly a contravariant functor. In particular, if f is an isomorphism, then so is $C(f)$, so $C(R/aR)$ depends only on the isomorphism class of R/aR , not on a itself. Now $D: {}_R\mathcal{T} \rightarrow \mathcal{T}_R$ is defined by symmetry and a routine verification shows that $CD \cong 1$, $DC \cong 1$. ■

The duality described here will reappear in Ch. 5 in another form.

Th. 3.1 may be interpreted as follows: given a factorization of a full matrix c in R , where R is the $n \times n$ matrix ring over an n -fir,

$$c = a_1 \dots a_r, \quad (3)$$

we associate with (3) the series of right ideals of R ,

$$R \supseteq a_1 R \supseteq a_1 a_2 R \supseteq \dots \supseteq a_1 \dots a_r R = cR,$$

and the corresponding series of torsion quotient modules

$$R/a_1 R, a_1 R/a_1 a_2 R \cong R/a_2 R, \dots, R/a_r R.$$

We also have the series of left ideals

$$R \supseteq Ra_r \supseteq Ra_{r-1}a_r \supseteq \dots \supseteq Ra_1 \dots a_r = Rc,$$

and torsion quotients

$$R/Ra_r, \dots, R/Ra_1.$$

Clearly it does not matter which we choose and Th. 3.1 is a general way of expressing this symmetry. It will be described as the *factorial duality*. Taking $n = 1$, we see that it holds for non-zero elements in any integral domain (=1-fir).

In general neither a submodule nor a quotient of a torsion module need be torsion, even over a semifir. However, if

$$0 \rightarrow M' \rightarrow M \rightarrow M'' \rightarrow 0 \quad (4)$$

is a short exact sequence of n -generator modules over a $2n$ -fir and any two of M, M', M'' are torsion, then so is the third. For it has zero characteristic by Prop. 0.6.1, so it only remains to show that submodules have non-negative characteristic.

(i) M' and M'' are torsion modules. Let N be a submodule of M , then

$$\chi(N) = \chi(N \cap M') + \chi(N/(N \cap M')). \quad (5)$$

Here the first term is ≥ 0 because $N \cap M' \subseteq M'$, and similarly for the second term, because

$$N/(N \cap M') \cong (N + M')/M',$$

and the right-hand side is a submodule of $M/M' \cong M''$.

(ii) M is a torsion module. Any submodule N of M' is also a submodule of M , hence $\chi(N) \geq 0$, and this shows M' to be torsion. Now any quotient module Q of M'' is also a quotient of M , hence $\chi(Q) \leq 0$, and this shows M'' to be a torsion module.

Furthermore, in any homomorphism between torsion modules,

$$f: M \rightarrow N,$$

$\ker f$ and $\operatorname{coker} f$ are torsion. To show this we observe that $\operatorname{im} f \cong \operatorname{coim} f$; since N is torsion, $\chi(\operatorname{im} f) \geq 0$, and since M is torsion, $\chi(\operatorname{coim} f) \leq 0$, hence $\chi(\operatorname{im} f) = 0$. Now any submodule of $\operatorname{im} f$ is a submodule of N and so has non-negative characteristic. Hence $\operatorname{im} f$ is a torsion module; by what has been shown it follows that $\operatorname{coker} f = N/\operatorname{im} f$ is torsion, and likewise for $\ker f$. Thus we have proved

PROPOSITION 3.2. *The category $\mathcal{T}_R^n$ of n -torsion modules over a $2n$ -fir R is a category with kernels, images and cokernels; more precisely, given $f: M \rightarrow N$ in $\mathcal{T}_R^n$, then $\ker f$, $\operatorname{im} f$ and $\operatorname{coker} f$ lie in $\mathcal{T}_R^n$.*

When R is a semifir, this result also applies to $\mathcal{T}_R$; moreover, $\mathcal{T}_R$ admits finite direct sums, hence (by Prop. A.3) we find

THEOREM 3.3. *Over a semifir R the category $\mathcal{T}_R$ of torsion modules is an abelian category. ■*

This result has numerous applications. In the first place we can derive an analogue of Schur's lemma, using the notion of a simple object in $\mathcal{T}_R$. A torsion module M is said to be $\mathcal{T}$ -simple if it is a simple object in $\mathcal{T}_R$, i.e. if $M \neq 0$ and no submodule of M (apart from 0 and M) lies in $\mathcal{T}_R$. For example, if $R = k\langle x, y \rangle$ is a free algebra, then R/xR is $\mathcal{T}$ -simple, though of course far from simple as R -module. If we write down the factorization corresponding to a chain of submodules in $\mathcal{T}_R$, we see that over a 2-fir R , a 1-torsion module R/cR is $\mathcal{T}$ -simple if and only if c is an atom.

PROPOSITION 3.4. *Let R be a $2n$ -fir and A an atom in R_n , i.e. an $n \times n$ matrix which is full, a non-unit and not a product of non-units, then the eigenring of A is a field.*

For A is an atom precisely if $R^n/R^n A$ is $\mathcal{T}$ -simple. By Prop. 3.2, any non-zero endomorphism must have kernel and cokernel 0, and hence it is an automorphism. ■

In some cases a more precise statement is possible. Thus, let R be a $2n$ -fir which is an algebra over a commutative field k . Then the eigenring of any matrix over R is again an algebra over k . Let A be an atom, so that its eigenring E is a field over k . If λ is any element of E which is transcendental

over k , then the elements $(\lambda - \beta)^{-1}$ for $\beta \in k$ all lie in E and are linearly independent over k . This shows that the dimension of E over k is at least $\text{card}(k)$. If the dimension of R over k is less than $\text{card}(k)$, this cannot happen (at least when k is infinite), so we have proved

PROPOSITION 3.5. *If R is a $2n$ -fir which is an algebra over a commutative field k , of dimension less than $\text{card}(k)$, then the eigenring of any full $n \times n$ matrix which is an atom is algebraic over k .*

When k is finite, the conclusion follows directly. ■

As an immediate consequence we have

COROLLARY 3.6. *If R is a $2n$ -fir which is an algebra of countable dimension over an uncountable commutative field k which is algebraically closed, then the eigenring of any full $n \times n$ matrix which is an atom is k itself. ■*

Over a $2n$ -fir R with left and right ACC_n every ascending chain of v -generator submodules of R^n , for any fixed $v \leq n$, becomes stationary, by Cor. 1.2.4, and the same is true for descending chains, by duality. Hence every torsion module in $\mathcal{T}_R^n$ has finite length, and by the Jordan-Hölder theorem, any two composition series are isomorphic. In terms of matrices the result may be stated as

THEOREM 3.7. *Over a $2n$ -fir with left and right ACC_n every full $n \times n$ matrix admits a complete factorization into atomic factors, and any two such complete factorizations are isomorphic. ■*

In view of Th. 1.2.3 the result holds for any $\aleph_0$ -firs, in particular for any firs. The first assertion of the theorem may also be expressed by saying that a $2n$ -fir with left and right ACC_n is n -atomic. Whether, conversely, every n -atomic $2n$ -fir (or even semifir) satisfies left and right ACC_n is not known for $n > 1$, but the answer is unlikely to be 'yes' (cf. Ex. 3 and 4).

A semifir which is n -atomic for all n is said to be *fully atomic*.

Exercises 3.3

1. Show that in any integral domain the following are equivalent: (i) right ACC_1 , (ii) ACC on principal right ideals generated by left factors of a fixed non-zero element and (iii) DCC on principal left ideals generated by right factors of a fixed non-zero element.

2. Show that in a 2-fir with left ACC_1 any family of elements, not all 0, have a highest common left factor.

3. Show that an integral domain with left and right ACC_1 is atomic, and conversely, that a UFD (in particular an atomic 2-fir) has left and right ACC_1 .

4°. Let R be a k -algebra with generators (in matrix form) $c_i = (c_{i1}, \dots, c_{in}) \in R^n$ and $A_i = (a_{rs}^{(i)}) \in R_n$ and defining relations $c_{i-1} = c_i A_i$ ($i = 1, 2, \dots$). Show that R is a semifir; if a_i denotes the right ideal generated by the components of c_i , show that $a_1 \subset a_2 \subset \dots$ and deduce that R does not satisfy right ACC_n . Is R n -atomic?

5. Let R be an integral domain which is a k -algebra, $a \in R^\times$ and let c be an element of the eigenring of a which is algebraic over k , thus $ca = ac'$, $f(c) = ab$ for some monic polynomial f over k and $b \in R$. Prove that $f(c') = ba$ and $c'b = bc$.

6. Let R be a 2-fir and $M = R/cR$ a 1-torsion module. Show that every finitely generated submodule of M is a torsion module if and only if c is right large. Further, the qualifier 'finitely generated' can be omitted if R is atomic.

7. Show that any element stably associated to a right large element of a 2-fir is again right large.

8. Over a 2-fir R , show that R/cR is simple as R -module if and only if c is a right large atom.

9. Let R be weakly 1-finite and $a, b \in R$. Show that there is an automorphism of R as right R -module which maps aR to bR if and only if a is associated to b .

10°. Show that if R is a simple principal ideal domain, then $\mathcal{T}_R^1$ admits direct sums; for which 2-firs does this hold? Interpret left or right ACC_n as a property of $\mathcal{T}_R$.

11. Let R be a 2-fir. Show that $\text{Hom}_R(R/aR, R/bR) \neq 0$ if and only if there is a non-unit right factor of a similar to a left factor of b .

12*. Let A be the group algebra of the free group on x, y, z over a commutative field k and let R be the subalgebra generated by $z^{-n}x, z^{-n}y, z$ ($n = 1, 2, \dots$). Show that R is a right fir, but that the family $\{z^{-n}x, z^{-n}y\}$ has no highest common left factor.

13. Let R be an atomic 2-fir in which every principal right ideal is an intersection of maximal principal right ideals. Show that every 1-torsion module is a direct sum of $\mathcal{T}$ -simple 1-torsion modules. Show also that in general there are infinitely many pairwise non-associated elements stably associated to a given atom.

14. Show that an atomic right Bezout domain is a principal right ideal domain. Deduce that a ring is a principal ideal domain if and only if it is an atomic (left and right) Bezout domain.

15. In the Weyl algebra $A_1(k)$, where $\text{char}(k) = 0$, on x and y , verify that $1 + xy$ is an atom, and is stably associated to xy . Deduce that $xyx + x$ has two complete factorizations of different lengths.

16. Show that $\mathbb{Z}\langle x, y \rangle$ is not a UFD, by considering factorizations of $xyx + 2x$. What forms can the eigenring of an atom take?

3.4 Rigid factorizations

Let us return to the definition of a UFD and consider more closely in what respects it differs from the definition in the commutative case. As we saw, given two factorizations,

$$c = a_1 \dots a_r \quad (1)$$

and

$$c = b_1 \dots b_r \quad (2)$$

of an element c , necessarily of the same length, there is a permutation $i \mapsto i'$ of $1, \dots, r$ such that

$$R/a_i R \cong R/b_{i'} R. \quad (3)$$

However, we note that

- (i) a_i and $b_{i'}$ are not necessarily associated and
- (ii) in general we do not obtain c by writing the b 's in the order $b_1, b_2, \dots, b_r$.

To this extent, unique factorization in non-commutative domains, as here defined, is a more complicated phenomenon than in the commutative case. But there is a more restrictive definition which is sometimes useful.

Let R be an integral domain; we recall from 0.8 that $c \in R$ is called *rigid* if $c = ab' = ba'$ implies $aR \subseteq bR$ or $bR \subseteq aR$. In other words, c is rigid if the lattice $L(cR, R)$ is a chain. If every non-zero element of R is rigid, R is called a *rigid domain*. In that case $R^\times$ is a rigid monoid in the sense of 0.8 and by Th. 0.8.12 we have

PROPOSITION 4.1. *Let R be a rigid domain. Then $R^\times$ can be embedded in a group. ■*

We observe that a rigid domain R is necessarily a 2-fir, by Th. 1.5.1. Now a rigid UFD is defined as an atomic rigid domain. For example, in the commutative case a rigid domain is just a valuation ring and a rigid UFD is a principal valuation ring. More generally, a rigid domain is right Ore if and only if it is a *right chain ring*, i.e. the principal right ideals form a chain. However, a non-commutative rigid domain can be much more general than a chain ring.

Our main source of rigid UFD's is the following result.

THEOREM 4.2. *Any complete inversely filtered ring with 2-term inverse weak algorithm is a rigid UFD.*

Proof. Let R be a ring satisfying the hypothesis. By Th. 2.9.4, R is a 2-fir; if $ab' = ba' \neq 0$ and $v(a) \leq v(b)$ say, then a, b are right v -dependent and so, by

the same theorem, $aR \supseteq bR$, hence R is rigid. By Prop. 2.9.5 it has left and right ACC_1 and so is atomic. Thus R is a rigid UFD. ■

An obvious example of a rigid UFD (other than a valuation ring) is a free power series ring $k\langle\langle X \rangle\rangle$.

We go on to describe rigid domains in more detail; for this we need the following lemma, which is also useful elsewhere. Over any ring R , a pair of matrices A, B (in that order) is said to be *comaximally transposable* if there exist matrices A', B' over R such that $AB = B'A'$ is a proper comaximal relation.

LEMMA 4.3. *Let R be an Hermite ring. Two matrices A, B over R are comaximally transposable if and only if there exist matrices X, Y over R such that*

$$XA - BY = I. \quad (4)$$

Proof. Suppose that we have a proper comaximal relation

$$AB = B'A'. \quad (5)$$

By Prop. 0.6.5 there exist mutually inverse matrices $P = \begin{pmatrix} A & B' \\ Y & T \end{pmatrix}$, $P^{-1} = \begin{pmatrix} X & -B \\ Z & A' \end{pmatrix}$, hence $XA - BY = I$, i.e. (4). Conversely, given (4), the matrix $(X, -B)$ is completable, so we obtain $Q = \begin{pmatrix} X & -B \\ * & * \end{pmatrix}$ with inverse $Q^{-1} = \begin{pmatrix} A & * \\ * & * \end{pmatrix}$, and equating the (1,2)-entries in the equation $Q^{-1}Q = I$ we obtain (5). ■

As a first consequence we characterize the Jacobson radical of a semifir, or more generally, of a matrix ring over a semifir.

THEOREM 4.4. *Let R be a full matrix ring over a semifir. Then $J(R)$ consists of those elements of R that cannot be comaximally transposed with any non-unit.*

Proof. Suppose that $c \in R$ cannot be comaximally transposed with any non-unit. For any $x \in R$ we have the proper comaximal relation

$$c(cx + 1) = (cx + 1)c, \quad (6)$$

hence $cx + 1$ is a unit for all $x \in R$ and this shows that $c \in J(R)$. Conversely, assume that $c \in J(R)$ and suppose that b, c can be comaximally transposed.

Then by Lemma 4.3, $cx + yb = 1$ for some $x, y \in R$ and by the choice of c , $yb = 1 - cx$ is a unit, hence b is a unit. A similar argument holds if c, b can be comaximally transposed. ■

COROLLARY 4.5. *Let R be a right fir. Then a right large element of R lies in $J(R)$ if and only if it has every atom of R as a left factor.*

Proof. If every atom is a left factor of c , then in any equation (6), $cx + 1$ must be a unit, for if it were a non-unit, it would have an atomic left factor p ; by hypothesis this is also a left factor of c , and hence of 1, which is a contradiction. Thus we see as before that $c \in J(R)$. Conversely, if $c \in J(R)$, let p be an atom which is not a left factor of c . Since c is right large, we have $cR \cap pR \neq 0$; by Th. 1.5.1, this intersection is principal, so we have a coprime relation $cp' = pc'$. By Cor. 2.5 this is comaximal, so p' is a non-unit and c, p' are comaximally transposable, which contradicts Th. 4.4. ■

We remark that in a principal right ideal domain this result holds for every non-zero element.

Secondly we shall obtain sufficient conditions for an element (or matrix) in a semifir to be rigid.

PROPOSITION 4.6. *Let R be a full matrix ring over a semifir and let c be a full matrix in R such that any two neighbouring non-unit factors occurring in a factorization of c generate a proper ideal of R . Then c is rigid. If c has an atomic factorization, then it is enough to check a single complete factorization of c .*

Proof. If c is not rigid, then $c = ab' = ba'$, where neither of a, b is a left factor of the other, say $a = da_1, b = db_1$, where a_1, b_1 are non-units. Let $a_1b'_1 = b_1a'_1$ be an LCRM, then $a' = a'_1e, b' = b'_1e$ and $a_1b'_1 = b_1a'_1$ is a comaximal transposition occurring in $c = da_1b'_1e$. By Lemma 4.3, $xa_1 - b'_1y = 1$ for suitable $x, y \in R$, so the ideal generated by a_1 and b'_1 is improper, a contradiction. In the atomic case we can pass from any atomic factorization to any other by a series of comaximal transpositions, so it is enough to test a single such factorization. ■

Now rigid domains are described by

THEOREM 4.7. *An integral domain is rigid if and only if it is a 2-fir and a local ring.*

Proof. We have seen that any rigid domain is a 2-fir, and clearly no two non-units can be comaximally transposable, hence by Th. 4.4, $J(R)$ includes

all non-units, i.e. R is a local ring. Conversely, in a 2-fir which is a local ring, any two non-units generate a proper ideal, hence by Prop. 4.6, every non-zero element is rigid. ■

Adding atomicity, we obtain

COROLLARY 4.8. *A ring is a rigid UFD if and only if it is an atomic 2-fir and a local ring. ■*

The description of commutative rigid UFD's, namely as principal valuation rings, can be extended to right Ore domains. It is easily verified that any right principal valuation ring is a rigid UFD and conversely, a rigid UFD is a right principal valuation ring if and only if any two atoms are right associated. The next theorem gives conditions for this to happen.

THEOREM 4.9. *Let R be a rigid UFD. Then R is a right principal valuation ring if and only if it contains a non-unit c such that $cR \cap pR \neq 0$ for every atom p ; in particular, this holds when R contains a right large non-unit.*

Proof. Suppose that the rigid UFD R contains an element c satisfying the given conditions. For any atom p of R we have $cR \cap pR \neq 0$, hence $cR \subseteq pR$ or $pR \subseteq cR$. Since c is a non-unit, the second alternative would mean that $cR = pR$, so in any case $cR \subseteq pR$, i.e. c has every atom as left factor. Let p, q be atoms, then $pR \cap qR \supseteq pR \cap cR \neq 0$, and the same reasoning as before shows that $pR = qR$, thus all atoms are right associated. Let p be an atom, then any non-zero element a of R can be written as $a = p^k u$ and by choosing k maximal we ensure that u is a unit. Therefore R is a right principal valuation ring; conversely in a right principal valuation ring the unique atom is a right large non-unit. ■

From the normal form of the elements it is clear that in a right principal valuation ring every right ideal is two-sided. By symmetry we obtain

COROLLARY 4.10. *If R is a rigid UFD with a left and right large non-unit, then R has a unique atom p (up to unit factors) and every left or right ideal of R is two-sided, of the form $p^n R = R p^n$. ■*

For matrices over local semifirs the factorization is no longer rigid, but we can say a little more than in the general case. We begin with a general lemma.

LEMMA 4.11. (i) *Let R be a ring and α a proper ideal of R such that R/α is weakly finite. If a matrix A over R has all its entries in α , then A is not stably associated to a matrix of smaller size.*

(ii) If R is a weakly finite ring and

$$AB' = BA' \quad (7)$$

is a comaximal relation, where $A, A' \in {}^sR^m$, and A has entries in $J(R)$, then B, B' are invertible and hence A, A' are associated.

Proof. (i) Let $A \in {}^sR^m$, $A' \in {}^sR^n$ and denote the images in $\bar{R} = R/\mathfrak{a}$ by bars. If A, A' are stably associated, then so are $\bar{A}, \bar{A}'$. Interpreting these matrices as homomorphisms between free modules, we have $\bar{A} = 0$, hence $\text{coker } \bar{A} \cong \bar{R}^m$, therefore $\text{coker } \bar{A}' \cong \bar{R}^m$ and so $\bar{R}^n$ splits over $\text{coker } \bar{A}' : \bar{R}^n \cong \bar{R}^m \oplus \text{im } \bar{A}'$. This shows that $m \leq n$ and a dual argument shows that $r \leq s$.

(ii) Clearly $\bar{R} = R/J(R)$ is again weakly finite. Over $\bar{R}$ we have a comaximal relation $\bar{A}\bar{B}' = \bar{B}\bar{A}'$ and $\bar{A} = 0$, so by comaximality $\bar{B}$ is a unit, hence $\bar{A}' = 0$ and $\bar{B}'$ is also a unit. But any matrix invertible mod J is invertible over R , hence B, B' are invertible and so A, A' are associated. ■

Let R be a local ring; it is clear that R is weakly finite, hence the index of a matrix is preserved by stable association. In particular, any matrix stably associated to a square matrix is itself square. For a square matrix A we define the level of A as the least value of the order of any matrix stably associated to A . For example, A has level 0 if and only if it is a unit. We shall also say that a matrix A is in $\mathfrak{a}$ if all its entries lie in $\mathfrak{a}$.

PROPOSITION 4.12. *Let R be a local ring with maximal ideal $\mathfrak{m}$, and let A be an $m \times n$ matrix over R with image $\bar{A}$ over $R/\mathfrak{m}$ of rank r . Then A is stably associated to an $(m-r) \times (n-r)$ matrix A' in $\mathfrak{m}$, but to no matrix of smaller size than A' . In particular, if A is $n \times n$, then its level is $n-r$.*

Proof. Since $\bar{A}$ has rank r , it is stably associated to the $(m-r) \times (n-r)$ zero matrix, but to no matrix of smaller size. Write

$$U \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} V \equiv \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \pmod{\mathfrak{m}},$$

where U, V are invertible $\pmod{\mathfrak{m}}$. It follows that U, V are invertible over R and we have

$$U \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} V = \begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix},$$

where P_1, P_2, P_3 are in $\mathfrak{m}$ and $P_4 \equiv I \pmod{\mathfrak{m}}$, hence P_4 is invertible over R . By row and column operations we can reduce P_4 to I and P_2, P_3 to 0 and thus find that $A \oplus I$ is associated to $A' \oplus I$, where A' is in $\mathfrak{m}$. Thus A is stably

associated to an $(m-r) \times (n-r)$ matrix A' in $\mathfrak{m}$, but clearly to no matrix of smaller size, by Lemma 4.11.

In particular, if A is $n \times n$, the level of A is $n - \text{rk } \bar{A}$. ■

We conclude this section with another example of a one-sided fir; this depends on a lemma of Chase [62] on projective modules, which is of independent interest.

LEMMA 4.13. *Let R be an integral domain with an element $p \neq 0$ such that $\cap p^n R = 0$. Then R^N , as left R -module, is not projective.*

Proof. Let us write $A^{(i)} \cong R$, $A = \prod_{i=1}^{\infty} A^{(i)}$, $A_r = \prod_{i=r+1}^{\infty} A^{(i)}$, so that $R^N = A \cong A_r$ for all r . If A were projective, there would be a left R -module B such that

$$A \oplus B \cong \bigoplus_I C_{\alpha}, \quad \text{where } C_{\alpha} \cong R. \quad (8)$$

Let us denote by $f_{\alpha}: A \rightarrow C_{\alpha}$ the projection from A on the summand C_{α} . We first show that there exists $r \geq 1$ such that

$$A_r f_{\alpha} = 0 \quad \text{for almost all } \alpha. \quad (9)$$

For suppose that this is false, then for each r , $A_r f_{\alpha} \neq 0$ for infinitely many α . We shall construct an increasing sequence $\{n_i\}$ of integers and sequences $\{\alpha_i\}, \{x_i\}$ such that $\alpha_i \in I$, $x_i \in p^{n_i-1} A_i$ and

$$x_k f_{\alpha_i} = 0 \quad \text{for } k < i, \quad x_i f_{\alpha_i} \notin p^{n_i} C_{\alpha_i}. \quad (10)$$

Choose $\alpha_1 \in I$ such that $A_1 f_{\alpha_1} \neq 0$ and set $n_0 = 0$, then $A_1 f_{\alpha_1} \notin p^{n_1} C_{\alpha_1}$ for some $n_1 \geq 0$, hence there exists $x_1 \in p^{n_0} A_1$ such that $x_1 f_{\alpha_1} \notin p^{n_1} C_{\alpha_1}$ and (10) holds for $i = 1$.

Suppose now that x_k, α_k, n_k have been constructed for $k < i$ to satisfy (10). Since $x_1, \dots, x_{i-1}$ lie in the sum of a finite subfamily of the C_{α} , there exist $\beta_1, \dots, \beta_r \in I$ such that $x_k f_{\alpha} = 0$ for $\alpha \neq \beta_1, \dots, \beta_r$ and for all $k < i$. So we may choose $\alpha_i \neq \beta_1, \dots, \beta_r$ such that $A_i f_{\alpha_i} \neq 0$, hence $A_i f_{\alpha_i} \notin p^m C_{\alpha_i}$ for some m . We put $n_i = n_{i-1} + m$, then there exists $x_i \in p^{n_i-1} A_i$ such that $x_i f_{\alpha_i} \notin p^{n_i} C_{\alpha_i}$. Now the sequences x_k, α_k, n_k for $k \leq i$ satisfy (10), so the construction is complete.

Let us write $x_k = (x_k^{(i)})$, where $x_k^{(i)} \in A^{(i)}$. Since $x_k \in p^{n_k-1} A_k$, it follows that $x_k^{(i)} = 0$ for $k > i$, so $x^{(i)} = \sum_{k=1}^{\infty} x_k^{(i)}$ is a well-defined element of $A^{(i)}$. Further, since $n_0 \leq n_1 \leq \dots$, there exists $y_r^{(i)} \in A^{(i)}$ such that $x^{(i)} = x_1^{(i)} + \dots + x_r^{(i)} + p^{n_r} y_r^{(i)}$, so on writing $x = (x^{(i)})$, $y_r = (y_r^{(i)})$, we have

$$x = x_1 + x_2 + \dots + x_r + p^{n_r} y_r \quad \text{for all } r \geq 1. \quad (11)$$

From (10) it is clear that the α_i are all distinct, hence there exists α_r such that $x f_{\alpha_r} = 0$. Writing x in the form (11) and applying f_{α_r} , we find that $x_r f_{\alpha_r} = -p^{n_r}(y_r f_{\alpha_r}) \in p^{n_r} C_{\alpha_r}$, but this contradicts (10), and so (9) is established.

Suppose now that $A_r f_{\alpha} = 0$ except for $\alpha = \beta_1, \dots, \beta_n$, then on writing $C' = C_{\beta_1} \oplus \dots \oplus C_{\beta_n}$, $C'' = \bigoplus C_{\alpha}$, we have $A_r \subseteq C'$. Now (8) may be written as

$$A^{(1)} \oplus \dots \oplus A^{(n)} \oplus A_r \oplus B \cong C' \oplus C''.$$

Since $A_r \subseteq C'$, it follows that A_r is complemented in C' , i.e. there exists D such that $A_r \oplus D \cong C'$. Now $C' \cong R^n$ is finitely generated, whereas $A_r \cong R^N$ clearly is not, and so we have reached a contradiction. This shows that R^N cannot be projective. ■

We shall use this result to construct a one-sided fir as follows. Let K be a field with an endomorphism α which is not surjective and consider the skew power series ring $R = K[[x; \alpha]]$. This is a right principal valuation ring with maximal ideal xR . Moreover, R is atomic and so it is a rigid UFD.

To show that R is not a left fir it will be enough to find a left ideal which is not free. In fact we shall find a left ideal isomorphic to R^N ; the result then follows by Lemma 4.13 (with $p = x$). To obtain such a left ideal we only need a sequence (u_n) tending to 0, which is left R -linearly independent. For then $\sum f_n u_n \in R$ for all $(f_n) \in R^N$, and by the linear independence the map

$$(f_n) \mapsto \sum f_n u_n$$

is an isomorphism of left R -modules. Take $c \in K \setminus K^\alpha$, then the sequence (xcx^n) has the required properties. Clearly it tends to 0, and if $\sum f_n xcx^n = 0$, then by cancelling a power of x on the right, if possible, we may assume that $f_0 \neq 0$. We can extend α to an endomorphism of R by putting $x^\alpha = x$, then

$$\sum x^{n+1} f_n^{\alpha^{n+1}} c^{\alpha^n} = 0,$$

and so $x f_0^\alpha c \in R^\alpha$, whereas not all the coefficients of $x f_0^\alpha c$ lie in K^α . This contradiction shows that the xcx^n are left R -linearly independent, and we have proved

THEOREM 4.14. *Let K be a field with a non-surjective endomorphism α . Then the skew power series ring $K[[x; \alpha]]$ is a right fir (in fact a right principal valuation ring), but not a left fir. ■*

Exercises 3.4

1. Show that a direct limit of rigid UFD's is again rigid, but not necessarily a UFD.

2*. Let $R = k[x; \alpha]$, where k is a commutative field of characteristic p and α is the p th power map. Find all elements of low degree that are rigid.

3. Let R be a rigid domain with right ACC_1 and a left and right large non-unit. Show that R is a principal valuation ring.

4. Let R be a semifir and consider a matrix relation $PA' = AP'$, where P, A are left coprime and A, A' are square. Show that if A' is full and P is in $\mathbf{J}(R)$, then A is invertible. Let $ab = pc$, where $b \neq 0$ and p is an atom of R contained in $\mathbf{J}(R)$. Show that if a is a non-unit, then $a \in pR$.

5. (Koshevoi [66]) An ideal $\mathfrak{p}$ in a ring R is called *strongly prime* if $R/\mathfrak{p}$ is an integral domain. Let R be an atomic 2-fir and $\mathfrak{p}$ a strong prime ideal in R . Show that $\mathfrak{p}$ contains with any atom a , all atoms similar to a . Let $c \in R$ have the complete factorizations $c = a_1 \dots a_r = b_1 \dots b_r$; show that if factors not in $\mathfrak{p}$ are omitted, then the same number of factors remain in each product and corresponding terms are similar.

6. Let R be a right hereditary local ring but not a valuation ring. Show that its centre is a field (Cohn [66'''], cf. 6.3).

7. Show that when $|X| > 1$, then $R = k\langle\langle X \rangle\rangle$ has left ideals isomorphic to R^N and hence is not a fir.

8°. In a rigid UFD, is the intersection of the powers of the maximal ideal necessarily zero? (cf. 5.11, 8.8).

9°. Can every UFD be embedded in a group?

10°. Investigate firs with finitely many atoms, and those with finitely many matrix atoms.

3.5 Factorization in semifirs: a closer look

We shall now examine the relation between different factorizations of an element in a 2-fir, or more generally, a full $n \times n$ matrix over a 2n-fir. For ease of notation we shall take R to be the $n \times n$ matrix ring over a 2n-fir and consider a full matrix c in R . A study of the factorizations of c is essentially a study of the lattice $\mathbf{L}(cR, R)$, but we shall usually express the result directly in terms of factorizations. In speaking of the 'left factors' of an element we shall tacitly understand the equivalence classes under right multiplication by units. In this way the *left* factors of an element c correspond to the principal *right* ideals containing cR :

$$c = ab \quad \text{if and only if } cR \subseteq aR.$$

Similarly, a chain of principal right ideals from cR to R corresponds to a factorization of c which is determined up to unit factors; thus we shall not

distinguish between the factorizations

$$c = a_1 a_2 \dots a_r \text{ and } c = b_1 b_2 \dots b_r,$$

if $b_i = u_{i-1}^{-1} a_i u_i$, where u_i is a unit and $u_0 = u_r = 1$.

With these conventions a rigid UFD may be described as an integral domain in which each element has essentially only one atomic factorization. In a general UFD the atomic factorizations are, of course, by no means unique, but neither can the factors be interchanged at will. Let us compare the different factorizations of a full $n \times n$ matrix over a $2n$ -fir. A factorization

$$c = a_1 \dots a_r \quad (1)$$

is said to be a *refinement* of another,

$$c = b_1 \dots b_s, \quad (2)$$

if (2) can be obtained from (1) by bracketing some of the a 's together, or in other words, if (1) arises from (2) by factorizing the b 's further. The factors of c correspond to torsion submodules of R/cR and by Prop. 3.2 admit $+$ and $\cap$ and so form a modular lattice. By the Schreier refinement theorem for modular lattices (cf. Appendix 1) we obtain

THEOREM 5.1. *Over a $2n$ -fir, any two factorizations of a full $n \times n$ matrix have isomorphic refinements. ■*

Looking at the proof of the lattice-theoretic result quoted here, we find that we can go from the refinement of one chain to that of the other by a series of steps, which each change the chain at a single point, from

$$\dots \geq x \vee y \geq x \geq x \wedge y \geq \dots \text{ to } \dots \geq x \vee y \geq y \geq x \wedge y \geq \dots$$

This corresponds to a change in the factorizations of the form

$$a_1 \dots a_i a_{i+1} \dots a_r \rightarrow a_1 \dots a'_i a'_{i+1} \dots a_r,$$

where $a_i a_{i+1} = a'_i a'_{i+1}$ is a comaximal relation. Calling such a change of factorizations a *comaximal transposition* of terms, we get the following more precise form of Th. 5.1.

THEOREM 5.2. *Over a $2n$ -fir, any two factorizations of a full $n \times n$ matrix have refinements which can be obtained from one another by a series of comaximal transpositions of terms. ■*

As a simple illustration, let $c = pq = rs$ be two factorizations in R and write $pR + rR = uR$, $p = ua$, $r = ub$, so that $aR + bR = R$, and similarly $Rq + Rs = Rv$,

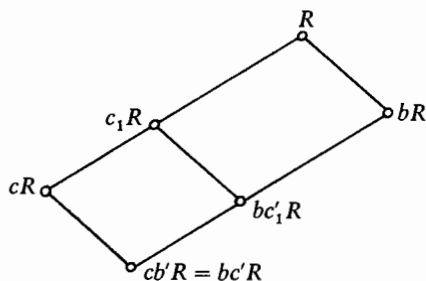
$q = b'v$, $s = a'v$. Then our factorizations become $(ua)(b'v) = (ub)(a'v)$, where $ab' = ba'$ is a comaximal relation.

In a general ring R a factorization $c = ab$ of a regular element need not induce a corresponding factorization in a similar element c' —the submodule of R/cR corresponding to aR/cR in R/cR will again be cyclic, but need not have a cyclic inverse image in R . However, in the $n \times n$ matrix ring over a 2n-fir, the principal right ideals between cR and R are characterized by the fact that they give rise to n -torsion submodules of R/cR ; now an application of the parallelogram law for modular lattices gives

PROPOSITION 5.3. *In the $n \times n$ matrix ring R over a 2n-fir, let c and c' be full matrices which are stably associated. Then the lattices $L(cR, R)$ and $L(c'R, R)$ are isomorphic, and the right ideals corresponding to each other under this isomorphism are similar.*

The last statement follows because corresponding right ideals are end-points of perspective intervals. ■

The isomorphism in Prop. 5.3 can be described explicitly as follows: if $cb' = bc'$ is a comaximal relation for c and c' , then to each left factor c_1 of c corresponds the left factor c'_1 of c' given by $c_1R \cap bR = bc'_1R$ and to each left factor c'_1 of c' corresponds the left factor c_1 of c given by $c_1R = cR + bc'_1R$. These maps are inverse to one another and induce an isomorphism between the lattices of left factors of c and c' .



However, we note that the actual lattice-isomorphism we get may depend on our choice of comaximal relation—or, equivalently, on our choice of the isomorphism between R/cR and $R/c'R$. For example, take the ring $C[x, -]$ of complex-skew polynomials; this is a principal ideal domain and hence a fir. The automorphism of the lattice of factors of $x^2 - 1$ induced by the comaximal relation

$$(x^2 - 1)x = x(x^2 - 1)$$

interchanges the factorizations $(x+i)(x+i)$ and $(x-i)(x-i)$ and leaves the factorization $(x+1)(x-1)$ fixed. The automorphism induced by the relation

$$(x^2-1)1 = 1(x^2-1)$$

is of course the identity, while the automorphism induced by

$$(x^2-1)i = i(x^2-1)$$

interchanges the factorizations $(x+1)(x-1)$ and $(x-1)(x+1)$.

Over any semifir the factorizations of a given full matrix are closely related to those of its factors; this is best understood by looking first at the situation in lattices. In any lattice L , a *link* or *minimal interval* is an interval $[a, b]$ in L consisting of just two elements, namely its end-points, and no others, thus $a < b$ and no $x \in L$ satisfies $a < x < b$.

In any modular lattice L of finite length, there are only finitely many projectivity classes of links, and the homomorphic images of L that are subdirectly irreducible are obtained by collapsing all but one equivalence class of links; these are, in fact, the simple homomorphic images of L . Here we count two homomorphic images as the same if and only if they are induced by the same homomorphism. If the distinct simple images are $L_1, \dots, L_r$, we have a representation of L as a subdirect product of $L_1, \dots, L_r$. In a distributive lattice of finite length, no two links in any chain are projective, as we shall see in 4.4, so the only simple homomorphic image is the 2-element lattice $[0, 1]$, also written 2. Hence, by the Birkhoff representation theorem (cf. Appendix 1) we obtain

THEOREM 5.4. (i) *Any modular lattice L of finite length can be expressed as a subdirect product of a finite number of simple modular lattices, viz. the simple homomorphic images of L .*

(ii) *Any distributive lattice L of finite length is a subdirect power of 2; more precisely, if L has length n , it is a sublattice of 2^n .* ■

There is another representation of modular lattices, to some extent dual to that of Th. 5.4, which is of use in studying factorizations.

THEOREM 5.5. *Let L be a modular lattice. Given $a, b \in L$, there is a lattice embedding $[a \wedge b, a] \times [a, a \vee b] \rightarrow L$ given by*

$$(x, y) \mapsto x \vee (b \wedge y) = (x \vee b) \wedge y. \quad (3)$$

Proof. Since $y \mapsto b \wedge y$ is an isomorphism, it follows that $(x, y) \mapsto x \vee (b \wedge y)$ preserves joins and similarly $(x, y) \mapsto (x \vee b) \wedge y$ preserves meets. Thus (3) is

a lattice homomorphism. It is injective because we can recover x and y from the right-hand side: $x \vee (b \wedge y) \vee a = (b \wedge y) \vee a = y$, $(x \vee b) \wedge y \wedge a = x \vee (b \wedge a) = x$. ■

Let L be a lattice and for any $a \in L$ put $[a] = \{x \in L \mid x \leq a\}$, $[a] = \{y \in L \mid a \leq y\}$. The proof of Th. 5.5 suggests that we consider the map $L \rightarrow [a] \times [a]$ defined by

$$z \mapsto (z \wedge a, z \vee a). \quad (4)$$

Of course, in general this is not a lattice homomorphism, but it clearly is one when L is distributive; further it is then injective, because two elements with the same image are relative complements of a in the same interval, and so must coincide. If moreover, a has a complement, b say, then $[a] = [a \wedge b, a]$, $[a] = [a, a \vee b]$ and if we now invoke Th. 5.5, we obtain

PROPOSITION 5.6. *Let L be a distributive lattice and $a \in L$. Then there is an embedding $L \rightarrow [a] \times [a]$ given by (4), and when a has a complement in L , this is an isomorphism. ■*

The translation of these results into factorizations reads as follows:

PROPOSITION 5.7. *Let R be the $n \times n$ matrix ring over a 2n-fir, and let c be a full matrix in R . Then any comaximal relation*

$$c = ab' = ba' \quad (5)$$

gives rise to an embedding of $\mathbf{L}(aR, R) \times \mathbf{L}(bR, R)$ into $\mathbf{L}(cR, R)$. If $\mathbf{L}(cR, R)$ is distributive, this is an isomorphism. ■

This result gives us a powerful tool relating the factorizations of a and b to those of c . For example suppose that we have a comaximal relation (5) in which a has a factorization xyz and b has a factorization uv . This gives us chains of lengths 3 and 2 in the lattice of left factors of a, b , respectively, and by applying Prop. 5.7 we find that the lattice of left factors of c will have a sublattice of the form shown in Fig. 1. Here intervals are marked with the factor of c to which they correspond.

Factorizations are given by paths from the top to the bottom of this diagram; every parallelogram corresponds to a comaximal relation, not only minimal parallelograms, giving relations such as $yu_2 = u_1y'$, but also larger ones, such as $(y'z')v_3 = v_1(y''z'')$. Thus, in these various factorizations of c , any factor from a and any factor from b are comaximally transposed.

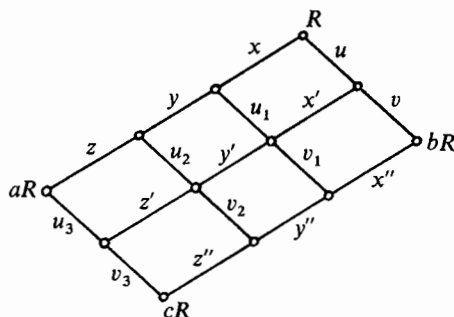


FIGURE 1.

Further, if, say, x and y are comaximally transposable: $xy = \bar{y}\bar{x}$, giving the subdiagram of left factors shown in Fig. 2, then from Prop. 5.7 we get a corresponding expanded diagram of factors of c , including comaximal parallelograms $x'y' = \bar{y}'\bar{x}'$, etc.

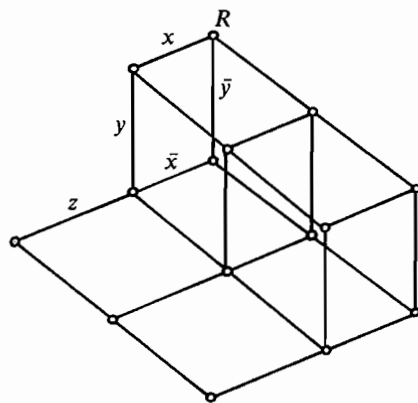


FIGURE 2.

However, when $L(cR, R)$ is not distributive, the embedding of Prop. 5.7 need not be an isomorphism. In terms of factorizations this means that some factors of a and b may be comaximally transposable in more ways than the one induced by the comaximal relation (5). For example, in Fig. 1 some of the parallelograms may be replaced by the diagram of Fig. 3. An example of such behaviour occurs in the complex-skew polynomial ring. Here $x^2 - 1 = (x + 1)(x - 1) = (x - 1)(x + 1)$ is a comaximal relation in which each factor is an atom, yet its full diagram of factorizations is of the form of Fig. 4, because $x^2 - 1 = (x + u)(x - \bar{u})$ for any u on the complex unit circle.

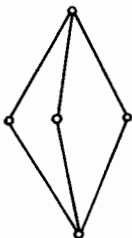


FIGURE 3.

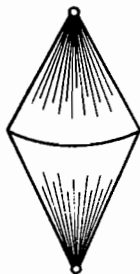


FIGURE 4.

Of course this cannot happen when $L(cR, R)$ is distributive; as we have seen, it is then a subdirect power of 2. This case will be studied in more detail in 4.4.

Exercises 3.5

1. Show that the group algebra (over k) of the additive group of rational numbers (written multiplicatively) is a non-atomic Bezout domain. (Hint: Write the elements as $\sum c_\alpha x^\alpha$, $\alpha \in \mathbb{Q}$, and express the ring as a directed union of polynomial rings.)

2. Show that the ring of power series $\sum c_\alpha x^\alpha$ ($\alpha \in \mathbb{Q}$, $\alpha \geq 0$) with well-ordered support is an atomless Bezout domain.

3. (Schofield) Let $X = \{x_1, x_2, \dots\}$ and consider the free power series ring $P = k\langle\langle X \rangle\rangle$, where each x_i has degree 1. Verify that the endomorphism $\theta: x_i \mapsto x_1 x_{i+1}$ on P maps non-units to non-units and that every non-unit in $\text{im } \theta$ has x_1 as proper left factor. Deduce that the direct limit of the system (P, θ^n) is an atomless semifir, which is neither left nor right Ore.

4. In a 2-fir, if an element c can be written as a product of two atoms in at least three different ways, show that all the atomic factors of c are similar. Generalize the result to n factors.

5. Over a 2-fir R consider an equation $p \oplus q = AB$, where p, q are dissimilar atoms and A, B non-invertible matrices. Show that this factorization is equivalent to either $(p \oplus 1)(1 \oplus q)$ or $(1 \oplus q)(p \oplus 1)$. What are the possibilities when p is similar to q ?

6°. Develop a theory of UFD's that are not necessarily 2-firs (cf. Brungs [69], Cohn [70, 73°]).

7. Over a fully atomic semifir, if A is a full matrix which is stably associated to AU , show that U is a unit (without atomicity this need not hold, cf. 8.8).

8°. Investigate non-commutative 2-firs in which any two elements with no common left or right factor can be comaximally transposed.

3.6 The primary decomposition

Besides the multiplicative decomposition of elements there is the primary decomposition of ideals which plays a role for commutative Noetherian rings. Much of this can be formulated in terms of lattices and by applying it to $\mathcal{T}_R$ we obtain various types of decomposition of full matrices over semifirs.

In any weakly 1-finite ring R , an element c is said to be *decomposable* if it has two proper factorizations (i.e. factorizations into non-unit factors)

$$c = ab' = ba', \quad (1)$$

which are left and right coprime; if c is not decomposable and a non-unit, it is said to be *indecomposable*. If c has two proper factorizations (1) which are right (left) coprime, c is said to be *right (left) decomposable*; if not, and c is a non-unit, it is *right (left) indecomposable*. Clearly any decomposable element is both right and left decomposable; hence any element which is either right or left indecomposable is indecomposable. For example, in an atomic 2-fir an element c is right indecomposable if and only if every complete factorization of c ends in the same right factor (cf. Prop. 6.9, below). We remark that a left and right decomposable element need not be decomposable (cf. Ex. 4); on the other hand, a 2-fir which is a local ring is rigid by Th. 4.7, hence every element not 0 or a unit is indecomposable.

If R is the $n \times n$ matrix ring over a 2n-fir, the definitions may be rephrased as follows:

(i) a full matrix c of R is *right decomposable* if and only if there exist $a, b \in R$ such that

$$cR = aR \cap bR, \quad \text{where } cR \neq aR, bR, \quad (2)$$

(ii) a full matrix c of R is *decomposable* if and only if there exist $a, b \in R$ such that

$$cR = aR \cap bR \text{ and } aR + bR = R, \quad \text{where } cR \neq aR, bR. \quad (3)$$

We observe that the definitions depend only on the module type of R/cR , as the following result shows:

PROPOSITION 6.1. *Let R be the $n \times n$ matrix ring over a 2n-fir. Then a full matrix c of R is right decomposable if and only if R/cR is an irredundant subdirect sum of two torsion modules, and c is decomposable if and only if R/cR is a direct sum of two non-zero torsion modules (or itself 0). In particular, if c is (left, right) decomposable, then so is any element similar to it.*

Proof. Let c be right decomposable, say (2) holds. Then there is a monomorphism

$$R/cR \rightarrow R/aR \oplus R/bR. \quad (4)$$

The result of composing this with the projection onto either of the summands is surjective, but neither is injective, hence R/cR is an irredundant subdirect sum of R/aR and R/bR . Conversely, given any subdirect sum representation of R/cR , the kernels of the projections are again torsion modules, say R/aR , R/bR , which will satisfy (2).

Next let c be decomposable, as in (3). Then

$$R/cR = (aR + bR)/cR \cong aR/cR \oplus bR/cR,$$

where the last sum is direct, because $aR \cap bR = cR$; moreover, the summands are non-zero torsion modules. The converse is clear. ■

We now consider the representations of an element corresponding to subdirect sum decompositions and direct decompositions of a module. Let R be the $n \times n$ matrix ring over a $2n$ -fir and c a full matrix of R ; suppose that we have an irredundant representation

$$cR = a_1R \cap \dots \cap a_rR, \quad (5)$$

where each a_i is right indecomposable. This means that the natural map

$$R/cR \rightarrow \bigoplus R/a_iR \quad (6)$$

obtained by composing the maps $R/cR \rightarrow R/a_iR$ ($i = 1, \dots, r$) is injective. Thus we have a subdirect sum representation of R/cR . Such a representation certainly exists if R is atomic. The irredundancy of (5) means that no term on the right of (6) can be omitted, while the right indecomposability of the a_i shows that each module R/a_iR is subdirectly irreducible, by Prop. 6.1. If we now apply the Kurosh–Ore theorem for modular lattices (cf. Appendix 1), we obtain

THEOREM 6.2. *Let R be the $n \times n$ matrix ring over an n -atomic $2n$ -fir. Then for each full matrix c in R , cR has an irredundant representation (5), where each a_i is right indecomposable, and if a second such decomposition of cR is given, $cR = b_1R \cap \dots \cap b_sR$, then $r = s$ and the b_jR may be exchanged against the a_iR , i.e. after suitably renumbering the b 's we have for $i = 1, \dots, r$,*

$$cR = a_1R \cap \dots \cap a_iR \cap b_{i+1}R \cap \dots \cap b_rR. \quad \blacksquare$$

Similarly, cR has a representation (5), where each a_i is indecomposable and

such that

$$a_i R \neq R, a_i R + \left(\bigcap_{j \neq i} a_j R \right) = R, \quad (i = 1, \dots, r). \quad (7)$$

Let us call a representation (5) satisfying these conditions a *complete direct decomposition* of cR . When (7) holds, the map (6) is an isomorphism, so that a complete direct decomposition of cR corresponds to a direct sum representation

$$R/cR \cong R/a_1 R \oplus \dots \oplus R/a_r R. \quad (8)$$

The first condition in (7) shows that each term on the right is non-zero, while the fact that each a_i is indecomposable means (by Prop. 6.1) that each module $R/a_i R$ is indecomposable in the category $\mathcal{T}_R$. We shall express this by saying that each $R/a_i R$ is $\mathcal{T}$ -indecomposable; thus R/aR is $\mathcal{T}$ -indecomposable (by Prop. 6.1) precisely when a is indecomposable. If we now apply the Krull-Schmidt theorem for modular lattices (cf. Appendix 1) to the lattice of principal right ideals containing cR , we obtain

THEOREM 6.3. *Let R be the $n \times n$ matrix ring over an n -atomic $2n$ -fir. Then for each full matrix c of R , cR has a complete direct decomposition (5) and if a second such decomposition $cR = b_1 R \cap \dots \cap b_s R$ is given, then $r = s$ and the b 's may be exchanged against the a 's. Moreover, the a 's and b 's are similar in pairs, in fact, after suitably renumbering the b 's, we have a set of coprime relations*

$$ua_i = b_i u_i, \quad (i = 1, \dots, r), \quad (9)$$

where u corresponds to a unit in the eigenring $E(cR)$.

Only the last part still needs proof. Let $u \in I(cR)$ correspond to the automorphism of R/cR transforming the isomorphism (8) into the corresponding relation with the b 's; then there is a coprime relation $uc = cu'$, say, and if $c = b_i b'_i = a_i a'_i$, then we have $ua_i a'_i = b_i b'_i u'$; since u, c are right comaximal, so are u, b_i for $i = 1, \dots, r$. If the b 's are now renumbered so that $R/a_i R \rightarrow R/b_i R$ in the automorphism, then we have the coprime relations (9), for some $u_i \in R$. ■

Since any left or right indecomposable element is indecomposable, it follows that any complete direct decomposition (5) can be refined to an irredundant decomposition. Therefore the decomposition of Th. 6.3 can have at most as many terms as that of Th. 6.2.

The Krull-Schmidt theorem for modules can be proved either by lattice-theory, as above, or by Fitting's lemma, which states that for an indecom-

posable R -module M of finite length the endomorphism ring $\text{End}_R(M)$ is *completely primary*, i.e. a local ring in which the maximal ideal is nilpotent (cf. e.g. Cohn [77], p. 355). Using Prop. 0.7.1, we can restate Fitting's lemma in the following form:

PROPOSITION 6.4. *Let R be the $n \times n$ matrix ring over an n -atomic $2n$ -fir. Then the eigenring of an indecomposable full matrix is completely primary. ■*

Thus if c is indecomposable and $a, a' \in R$ are such that $ac = ca'$, then either a, c are right comaximal or there exists $b \in R$ and $r \geq 0$ such that $a' = cb$. In particular, assuming a to be an atom, we find

COROLLARY 6.5. *Let R be as in Prop. 6.4 and c be an indecomposable full matrix, with idealizer $I(cR)$. If $I(cR)$ contains an atom p , then either p is right comaximal with c or there exists $b \in R$ and $r \geq 0$ such that $cb = p^r$. In this case all atomic factors of c are similar. ■*

Of course, in general the atomic factors of an indecomposable element need not all be similar, e.g. take xy in the free algebra $k\langle x, y \rangle$; any atom in the idealizer of xy , such as $1 + xy$, is comaximal with xy .

We also note the special case of a rigid UFD:

COROLLARY 6.6. *Let R be a rigid UFD. Then the eigenring of any non-zero element is completely primary.*

For, by Cor. 4.8, such a ring R is an atomic 2-fir and a local ring. By the remark at the beginning of this section, any non-unit in R is indecomposable, so the result follows from Prop. 6.4. ■

For cyclic modules there is a criterion for direct decomposability which actually holds in quite general rings.

LEMMA 6.7. *Let R be any ring and $a, b \in R$, where a is right regular. Then*

$$R/bR \cong aR/abR, \quad (10)$$

and this is a direct summand of R/abR if and only if there exist $c, d \in R$ such that

$$da - bc = 1. \quad (11)$$

Moreover, in that case we have

$$R/abR \cong R/bR \oplus R/aR. \quad (12)$$

Proof. Write $M = R/abR$ and suppose that

$$M = N_1 \oplus N_2, \quad \text{where } N_1 = aR/abR.$$

Denote by u the image of 1 in M , then $u = v_1 + v_2$, where $v_i \in N_i$. Further, $N_1 \cong R/bR$ by (10), which follows because a is right regular. Now $N_i = v_iR$ and we have $v_2x = 0 \Leftrightarrow ux \in N_1 \Leftrightarrow x \in aR$, hence $N_2 \cong R/aR$ and (12) follows.

To establish (11), put $v_2 = ur$, then $u = v_1 + ur$ and $v_1 \in uaR$, say $v_1 = uad$; hence $u = uad + ur$ and so $1 - ad - r = abx$ for some $x \in R$. Replacing r by $r - abx$, we have $r = 1 - ad$. Now $u(1 - ad)a = 0$, so $(1 - ad)a = -abc$ for some $c \in R$, and hence $a(1 - da + bc) = 0$. Since a was right regular, we have $da - bc = 1$, i.e. (11).

Conversely, given (11), we put $r = 1 - ad$, $u = uad + u(1 - ad)$, and so $M = uadR + urR$. If $uax = ury$, then $ax - ry = abz$, hence $ax - y + ady = abz$, so $y \in aR$. Now $ra = (1 - ad)a = a(1 - da) = -abc$. It follows that $ura = 0$, hence $ury = 0$ and so $uaR \cap urR = 0$, as claimed. ■

To apply the result to $2n$ -firs, we note that if S is a $2n$ -fir, then an n -torsion module over S corresponds under the Morita equivalence (cf. 0.1) to a cyclic S_n -module. We thus have

PROPOSITION 6.8. *Let R be the $n \times n$ matrix ring over a $2n$ -fir and let M, N be cyclic right R -modules defined by full matrices. Then $M \oplus N$ is cyclic if and only if there is a left comaximal pair of full matrices a, b in R such that $M \cong R/aR$, $N \cong R/bR$.*

Proof. If $M \oplus N$ is cyclic, where $M = R/aR$, $N = R/bR$, then (11) holds, hence a, b are comaximally transposable, say $ab = b'a'$. But then a is similar to a' and a', b are left comaximal.

Conversely, if a, b are left comaximal, let $a_1b = b_1a$ be a comaximal relation, then a_1 is similar to a and a_1, b are comaximally transposable, hence by the lemma, $R/a_1R \oplus R/bR$ is cyclic. ■

We also note the following criterion for one-sided decomposability:

PROPOSITION 6.9. *Let R be the $n \times n$ matrix ring over a $2n$ -fir and c a full matrix in R . Then the following conditions are equivalent:*

- c is right indecomposable, with an atomic right factor,*
- the lattice $L(cR, R)$ has a unique minimal element covering cR ,*
- the lattice $L(Rc, R)$ has a unique maximal element covered by R ,*
- c has an atomic right factor, unique up to left associates.*

The proof may be left to the reader. ■

Using (c), we obtain

COROLLARY 6.10. *If R, c are as in Prop. 6.9 and c is right indecomposable, then c is any right factor of c . ■*

Another type of decomposition, sometimes of interest, exists for certain elements, namely those whose associated torsion module is completely reducible (or semisimple). A full matrix c in the $n \times n$ matrix ring R over a $2n$ -fir is said to be *fully reducible* if

$$cR = \bigcap p_i R, \quad (13)$$

where the p_i are atoms. By passing to the corresponding torsion module we see that c is fully reducible if and only if R/cR is a subdirect product of $\mathcal{T}$ -simple modules. Thus being fully reducible is again a property of the module type R/cR . If, moreover, R is atomic, then it is enough to take finitely many atoms on the right of (13). Taking the number of terms to be minimal, we obtain R/cR as a subdirect sum of the $R/p_i R$, and by minimality the sum is actually direct:

$$R/cR \cong R/p_1 R \oplus \dots \oplus R/p_r R. \quad (14)$$

By the factorial duality, being fully reducible in R is a left-right symmetric property. Further, any factor of a fully reducible element is again fully reducible, as the representation (14) shows.

In order to relate full reducibility to factorizations we need another concept. A factorization

$$c = a_1 \dots a_r \quad (15)$$

of a fully regular element c is said to be a *cleavage* if every atomic factorization of c can be obtained by refining (12). A non-unit with no cleavage into more than one non-unit is said to be *uncleft*. For example, take $R = \mathbb{Z}$, or more generally, any commutative principal ideal domain. If p is an atom (i.e. a prime number in the case $R = \mathbb{Z}$), then the element $p^2(p+1)$ is uncleft, but it has a cleft factor p^2 . Let us call an element *totally uncleft* if all of its factors are uncleft; thus in a commutative PID the totally uncleft elements are just the 'squarefree' elements. The connexion with fully reducible elements is given by

PROPOSITION 6.11. *Let R be the $n \times n$ matrix ring over a $2n$ -fir. Given a full*

matrix c in R of finite length (i.e. number of factors in a complete factorization), the following conditions on c are equivalent:

- (a) c is totally uncleft,
- (b) c is fully reducible,
- (c) any two neighbouring factors in a complete factorization of c can be comaximally transposed.

Proof. (a) $\Rightarrow$ (b). Let c be totally uncleft, then not every complete factorization of c ends in the same atom, say $c = c_1 p_1 = c_2 p_2$, where $R p_1 \neq R p_2$. Then $cR = c_1 R \cap c_2 R$; now c_1, c_2 as factors of c are again totally uncleft and by induction on the length of c , each c_i is fully reducible, say $c_i R = \cap_j p_{ij} R$. Hence $cR = \cap_{ij} p_{ij} R$, which shows c to be fully reducible.

(b) $\Rightarrow$ (c). Assume c fully reducible, take a complete factorization $c = a_1 \dots a_r$ and fix i in the range $2 \leq i \leq r$. Then $a_{i-1} a_i$ is fully reducible, as factor of c , hence there is a comaximal relation $a_{i-1} a_i = a'_i a'_{i-1}$.

(c) $\Rightarrow$ (a). This follows because any cleavage of any factor of c , on being refined to a complete factorization, leads to a pair of neighbouring atomic factors which cannot be comaximally transposed. ■

In a rigid UFD any factorization is clearly a cleavage. Generally, let us say that a factorization of a certain type, e.g. into maximal uncleft factors, is *rigidly unique* if it is the only one of this type (up to unit factors).

THEOREM 6.12. *Let R be the $n \times n$ matrix ring over a $2n$ -fir. Then every full matrix c in R of finite length has a factorization*

$$c = a_1 \dots a_r \quad (16)$$

into maximal uncleft factors, and this is rigidly unique.

Proof. Let a_1 be an uncleft left factor of c of maximal length. If a'_1 is another maximal uncleft left factor of c , say $c = a_1 b_1 = a'_1 b'$, let

$$a_1 d = a'_1 d' \quad (17)$$

be an LCRM of a_1 and a'_1 . If d is not a unit, then $a_1 d$ is cleft, say $a_1 d = uv$, and every complete factorization is obtained by refining uv , whence either $a_1 = ue$, $v = ed$, or $u = a_1 e$, $d = ev$. In the first case $a_1 = ue$ is a cleavage, which contradicts the fact that a_1 is uncleft, hence this case is ruled out and $u = a_1 e$, $d = ev$. Similarly, $u = a'_1 e'$, $d' = e'v$, but by (17), d, d' are right co-prime, so v is a unit, again a contradiction. This shows d to be a unit, likewise d' is a unit and $a_1 R = a'_1 R$. Thus a_1 in (16) is determined up to a right unit factor, and an induction on the length of c shows the uniqueness of (16). ■

In general (16) need not be a cleavage, but this is the case when the factors on the right of (16) are atoms. By Prop. 6.11 this happens precisely when no two factors in (16) are comaximally transposable. In terms of lattices this means that $\mathbf{L}(cR, R)$ is a chain. Thus we have

COROLLARY 6.13. *A full matrix c of finite length in the $n \times n$ matrix ring R over a $2n$ -fir has a cleavage into atomic factors if and only if $\mathbf{L}(cR, R)$ is a chain, or equivalently, if c is rigid. ■*

Exercises 3.6

1. In the complex-skew polynomial ring, obtain all possible irredundant representations of $(x^2 - 1)R$.
2. Show that in any commutative principal ideal domain that is not a local ring there are elements which have no rigidly unique factorization into maximal totally unleft elements.
3. Show that in a commutative integral domain every factor of an indecomposable element is again indecomposable. Show that in $k\langle x, y, z \rangle$ the element $zx(yx + 1)z$ is indecomposable (even left and right indecomposable) but has a decomposable factor.
4. Find an element in $k\langle x, y, z \rangle$ which is indecomposable but not left or right indecomposable.
5. Show that in an atomic 2-fir any factor of a fully reducible element is fully reducible.
6. An element c of a ring R is said to be *primary* if $cR = \bigcap p_i R$, where the p_i are pairwise similar atoms. Show that in the $n \times n$ matrix ring over an n -atomic $2n$ -fir every full matrix which is fully reducible can be written uniquely as LCRM of primary elements.
7. Let R be the $n \times n$ matrix ring over a $2n$ -fir and let $a = a_1 \dots a_k$ be a decomposition into maximal unleft factors. If b is a left factor of a , with decomposition $b = b_1 \dots b_h$ into maximal unleft factors, show that $b_1 \dots b_i$ is a left factor of $a_1 \dots a_i$ for $i = 1, \dots, h$.
- 8⁰. Investigate atomic 2-firs (other than principal valuation rings) in which any two atoms are similar (cf. Faith [73], p. 362).

Notes and comments on Chapter 3

The first non-commutative UFD to have been studied is the ring $K[D; 1, \cdot]$ of linear differential operators. It is discussed at some length by Schlesinger [1897], who proves that it is an integral domain. Landau [02] shows that all complete factorizations of a given operator have the same length, and corresponding irreducible factors have the same order (=degree in D). Loewy [03] shows that corresponding factors are

'equivalent' operators, in a sense introduced by Poincaré, and this turns out to correspond to the notion of similarity. A large number of papers on the subject appeared at this time. The first abstract account of this ring was given by Ore [32], who also introduced the notion of 'eigenring'. A further generalization, to principal ideal domains, is undertaken by Asano [38]. This and much other work is summarized in Chapter 3 of Jacobson [43], where the criterion of Prop. 2.1 is proved for principal ideal domains. The general notion of non-commutative UFD is defined in Cohn [63], and it is shown there that free algebras have this property. Rigid UFD's are also defined there *a propos* of the example of free power series rings.

Nagata [57] proved Th. 1.4 for commutative Noetherian rings; the other results of 3.1 are mostly well known. The notion of full matrix (3.2) was introduced by Cohn [71]; for left full, left prime, etc., see Cohn [82], where 2.4–6 are also proved. Much of 3.3 follows Cohn [63, 69, 70], though the strictly cyclic (=cyclic torsion) modules of the first edition have been replaced by torsion modules. This corresponds to taking factorizations of full matrices rather than elements, and most of the subsequent results are stated in this more general form, which is usually no harder to prove. Many of the results are just consequences of the fact (Th. 3.3) that the torsion modules over a semifir form an abelian category. Factorizations of unbounded length (in Bezout domains, and more generally in 2-firs) have been studied by Beauregard [69], Beauregard and Johnson [70], Brungs [78] and Paul [73]. Prop. 3.5, based on the linear independence of the $(\lambda - \beta)^{-1}$, where λ is transcendental, is Amitsur's well-known trick (which he apparently noticed while lecturing on complex function theory).

An interesting generalization of UFD, to include the case of $\mathbb{Z}\langle X \rangle$, has been proposed by Brungs [69]; for another approach see Cohn [70] and for a survey, Cohn [73']. A detailed study of similarity, foreshadowing a form of Schanuel's lemma, was undertaken by Fitting [36]; the parts relevant to us are contained in 0.6 and 3.2 (*cf.* also Cohn [82]). For 3.4 see Cohn [62] and Bowtell [67]. Prop. 4.6 generalizes a result by Koshevoi [66] on free algebras, and Prop. 4.12 is taken from Cohn [85]. Lemma 4.13 is the special case for integral domains of a result of Chase [62]; for the application made here, see Cohn [66''].

Section 3.5 is due to Bergman [67] (*cf.* also an unpublished manuscript by Bergman, dating from 1969), while 3.6 generalizes earlier results of Ore [33], *cf.* also Feller [60], Johnson [65] and Cohn [69, 70, 73']. For another approach to primary decomposition in non-commutative rings, see Barbilian [56]. Lemma 6.7 generalizes a result proved in the first edition and is taken from Dieudonné [73], p. 164.

Cozzens and Faith [75] define a V-ring as a ring for which every simple right module is injective (e.g. the commutative V-rings are just the v. Neumann regular rings). It can be shown that a right PID is a V-ring if and only if every non-zero element is fully reducible. Every V-ring is a TC-ring, defined as follows: A *test module* for a ring R is a module T such that $\text{Hom}_R(M, T) = 0$ implies $M = 0$; if every test module is a cogenerator, R is called a TC-ring. Let R be a TC-ring and S a simple R -module; if $E(S)$ is its injective hull, then $\text{End}_R(E(S))$ is a local ring and a semifir (*cf.* Vámos [76]).

4. Rings with a distributive factor lattice

This chapter examines more closely those 2-firs in which the lattice of factors of any non-zero element is distributive. It is shown in 4.3 that this holds for free algebras and the consequences are traced out in 4.2, 4 and 5, while 4.6 describes the form taken by eigenrings in this case.

4.1 Distributive modules

In this section $\mathcal{A}$ will be a full subcategory of ${}_R\mathcal{M}$, the category of all left R -modules over a given ring R ; moreover, any kernel or cokernel (taken in ${}_R\mathcal{M}$) of a map of $\mathcal{A}$ is again in $\mathcal{A}$. In general $\mathcal{A}$ may not admit direct sums (if it does, we have an abelian category, by Prop. A.3 of Appendix 2). Any notions such as submodule, quotient, simple, etc., will be understood as referring to $\mathcal{A}$ without explicit mention.

Since $\mathcal{A}$ admits kernels and cokernels (and hence images and coimages), the set $\text{Lat}_{\mathcal{A}}(M)$ of submodules of M is a lattice, necessarily modular. If $\text{Lat}_{\mathcal{A}}(M)$ is distributive, the module M is said to be *distributive*. The following are some examples of occurrences of distributive modules:

(i) $\mathcal{A} = {}_R\mathcal{M}$, where R is a commutative Bezout domain (or more generally, a Prüfer domain). Any cyclic R -module is distributive (Jensen [63]).

(ii) $\mathcal{A} = {}_R\mathcal{M}$, where R is semisimple Artinian. An R -module M is distributive if and only if each simple module occurs at most once in M . If each simple module occurs exactly once, then M is faithful. For this reason R is called *distributively representable* by Behrens [65].

(iii) In every right Artinian algebra of finite representation type the lattice of all 2-sided ideals is distributive, cf. Pierce [82], p. 104.

(iv) In 4.3 we shall see that any 1-torsion module M over a free algebra $R = k\langle X \rangle$ is distributive. Here it is essential to consider M as object of ${}_R\mathcal{T}^1$ rather than ${}_R\mathcal{M}$.

In the rest of this section we shall examine the structure of a distributive module, in preparation for what follows. If M is a distributive module, then clearly any submodule and any quotient of M are again distributive.

For the moment let M be any module, A, B submodules of M and α a

homomorphism into M , then we clearly have

$$A\alpha^{-1} + B\alpha^{-1} \subseteq (A+B)\alpha^{-1},$$

but equality need not hold. It is easily seen that we have equality when A, B lie in the image of α . Dually, we have, for a homomorphism β from M ,

$$(A \cap B)\beta \subseteq A\beta \cap B\beta,$$

and here we have equality if A or B contains $\ker \beta$, but not generally. In fact, equality (in either case) characterizes the distributivity of M :

PROPOSITION 1.1. *For any module M (over any ring) the following conditions are equivalent:*

- (a) M is distributive,
- (b) for any module P and $\alpha: P \rightarrow M$,

$$(A+B)\alpha^{-1} = A\alpha^{-1} + B\alpha^{-1} \quad \text{for all } A, B \in \text{Lat}_{\mathcal{A}}(M),$$

- (c) for any module Q and any $\beta: M \rightarrow Q$,

$$(A \cap B)\beta = A\beta \cap B\beta \quad \text{for all } A, B \in \text{Lat}_{\mathcal{A}}(M).$$

Proof. (a) $\Rightarrow$ (b). We have $X \cap P\alpha = X\alpha^{-1}\alpha$ for any $X \subseteq M$, hence

$$(A+B)\alpha^{-1}\alpha = (A+B) \cap P\alpha = A \cap P\alpha + B \cap P\alpha = A\alpha^{-1}\alpha + B\alpha^{-1}\alpha,$$

and so $(A+B)\alpha^{-1} = (A+B)\alpha^{-1}\alpha\alpha^{-1} = (A\alpha^{-1}\alpha + B\alpha^{-1}\alpha)\alpha^{-1}$; any element of this last module has the form $x + y$, where $x\alpha \in A$, $y\alpha \in B$, hence it is contained in $A\alpha^{-1} + B\alpha^{-1}$, as we had to show.

- (b) $\Rightarrow$ (a). If $A, B, C \subseteq M$, let $\gamma: C \rightarrow M$ be the inclusion map, then

$$(A+B) \cap C = (A+B)\gamma^{-1} = A\gamma^{-1} + B\gamma^{-1} = A \cap C + B \cap C.$$

Now (a) $\Leftrightarrow$ (c) follows by duality. ■

We shall need another technical result on homomorphisms between distributive modules.

LEMMA 1.2. *Let P, Q be any modules and $\alpha, \beta \in \text{Hom}(P, Q)$.*

- (i) *If Q is distributive and X is a submodule of P , then*

$$(a) \quad P = (\text{im } \alpha)\beta^{-1} + (\text{im } \beta)\alpha^{-1}, \quad (b) \quad X = (X \cap X\alpha\beta^{-1}) + (X \cap X\beta\alpha^{-1}),$$

- (ii) *if P is distributive and Y is a submodule of Q , then*

$$(a) \quad 0 = (\ker \alpha)\beta \cap (\ker \beta)\alpha, \quad (b) \quad Y = (Y + Y\alpha^{-1}\beta) \cap (Y + Y\beta^{-1}\alpha).$$

Proof. (i) We have $\text{im}(\alpha + \beta) \subseteq \text{im } \alpha + \text{im } \beta$, hence by Prop. 1.1,

$$\begin{aligned} P &= (\text{im } \alpha + \text{im } \beta)(\alpha + \beta)^{-1} = (\text{im } \alpha)(\alpha + \beta)^{-1} + (\text{im } \beta)(\alpha + \beta)^{-1} \\ &= (\text{im } \alpha)\beta^{-1} + (\text{im } \beta)\alpha^{-1}. \end{aligned}$$

Now (b) follows by applying (a) to $1_X \cdot \alpha, 1_X \cdot \beta$, thus:

$$X = X\alpha(1_X \cdot \beta)^{-1} + X\beta(1_X \cdot \alpha)^{-1} = (X \cap X\alpha\beta^{-1}) + (X \cap X\beta\alpha^{-1}).$$

Again, (ii) follows by duality. ■

A module M is said to be *semi-Artinian* if every non-zero quotient of M has a simple submodule; dually, if every non-zero submodule of M has a simple quotient, M is called *semi-Noetherian*. For modules satisfying these hypotheses Lemma 1.2 can be extended as follows.

PROPOSITION 1.3. *Let P, Q be modules and $\alpha, \beta \in \text{Hom}(P, Q)$.*

(i) *If P is semi-Artinian and Q is distributive, then $\ker \alpha \subseteq \ker \beta$ implies $\text{im } \beta \subseteq \text{im } \alpha$.*

(ii) *If Q is semi-Noetherian and P is distributive, then $\text{im } \beta \subseteq \text{im } \alpha$ implies $\ker \alpha \subseteq \ker \beta$.*

Proof. Suppose that $\ker \alpha \subseteq \ker \beta$ but $\text{im } \beta \not\subseteq \text{im } \alpha$, and define

$$A = \Sigma\{B \subseteq P \mid B\beta \subseteq B\alpha\}.$$

Clearly A is the largest submodule of P such that $A\beta \subseteq A\alpha$. In particular it follows that $\ker \alpha \subseteq \ker \beta \subseteq A$. Since $\text{im } \beta \not\subseteq \text{im } \alpha$, we have $A \neq P$. By hypothesis there exists $X \subseteq P$ such that $A \subset X$ and X/A is simple. But $A \subseteq A\alpha\beta^{-1} \subseteq X\alpha\beta^{-1}$ and so $A \subseteq X \cap X\alpha\beta^{-1} \subseteq X$. Since $X\beta \not\subseteq X\alpha$ and X/A is simple, we find that $A = X \cap X\alpha\beta^{-1}$. By Lemma 1.2 (i) (b),

$$X = (X \cap X\alpha\beta^{-1}) + (X \cap X\beta\alpha^{-1}) = A + X \cap X\beta\alpha^{-1}$$

and so

$$\begin{aligned} X\alpha &= A\alpha + (X \cap X\beta\alpha^{-1})\alpha = A\alpha + (X\alpha \cap X\beta) = A\alpha + (X\alpha\beta^{-1} \cap X)\beta \\ &= A\alpha + A\beta = A\alpha. \end{aligned}$$

Since $\ker \alpha \subseteq A$, we find that $X = A$, a contradiction; hence $\text{im } \beta \subseteq \text{im } \alpha$. This proves (i); now (ii) follows by duality. ■

We recall that a submodule of a module M is said to be *fully invariant* if it is mapped into itself by all endomorphisms of M . We shall see that under

suitable finiteness conditions distributive modules have an even stronger property:

COROLLARY 1.4. (i) *Let M be a semi-Artinian distributive module and N a submodule; then any homomorphism $\alpha: N \rightarrow M$ is such that $N\alpha \subseteq N$. In particular, any two isomorphic submodules of M are equal.*

(ii) *If M is a semi-Noetherian distributive module and N a submodule of M , then for any $\beta: M \rightarrow M/N$, we have $\ker \beta \supseteq N$, so that β is induced by an endomorphism of M/N . In particular, distinct submodules of M determine non-isomorphic quotients.*

The proof is immediate from Prop. 1.3, taking α to be the inclusion map and $\beta = \alpha$, and similarly for (ii). ■

COROLLARY 1.5. *Let M be a distributive module. If M is either semi-Artinian or semi-Noetherian, then every submodule of M is fully invariant.* ■

There is another characterization of distributive modules which is sometimes useful. Let M be a module and suppose that $M = M_1 \oplus M_2$. With any map $\alpha: M_1 \rightarrow M_2$ we associate a submodule of M , its *graph*

$$\Gamma(\alpha) = \{(x, x\alpha) \mid x \in M_1\}.$$

It is clear that $\Gamma(\alpha) \cap M_2 = 0$, $\Gamma(\alpha) + M_2 = M$; thus $\Gamma(\alpha)$ is a complement of M_2 in M , and it is easily seen that any complement of M_2 in M defines a map $M_1 \rightarrow M_2$ in this way. Since in a distributive module, complements are unique, we see that when M is distributive, $\text{Hom}(M_1, M_2) = 0$. In particular, it follows that M cannot be isomorphic to a module of the form N^2 , i.e. the square of a module. Let us call M *square-free* if M has no factor isomorphic to a square. Then we have

PROPOSITION 1.6. *For any module M the following are equivalent:*

- (a) *M is distributive,*
- (b) *$\text{Hom}(P/P \cap Q, Q/P \cap Q) = 0$ for all submodules P, Q of M ,*
- (c) *M is square-free.*

Proof. (a) $\Rightarrow$ (b). If M is distributive, then so is $(P+Q)/(P \cap Q)$, and now (b) follows from the preceding remarks.

(b) $\Rightarrow$ (c). Let $M \supseteq A \supset B$, where A/B is a square; this means that there exist $P, Q \supset B$ such that $P+Q = A$, $P \cap Q = B$ and $P/B \cong Q/B$, but this contradicts (b).

(c) $\Rightarrow$ (a). Suppose that M is not distributive. Then $\text{Lat}_{\mathcal{A}}(M)$, being modular, has a sublattice of the form given in Prop. A.2, i.e. there exist P_i ($i = 1, 2, 3$) with $P_i + P_j = A$, $P_i \cap P_j = B$ ($i \neq j$), hence $A/P_i \cong P_j/B$, and so A/B is a square. ■

We apply this result to analyse the endomorphism ring of a distributive module which is a direct sum.

THEOREM 1.7. *Let M be a distributive module such that*

$$M = M_1 \oplus \dots \oplus M_n. \quad (1)$$

Then

$$\text{End}(M) \cong \prod_{i=1}^n \text{End}(M_i). \quad (2)$$

For any endomorphism α of M is represented by a matrix (α_{ij}) where $\alpha_{ij} \in \text{Hom}(M_i, M_j)$ and by Prop. 1.6, $\alpha_{ij} = 0$ for $i \neq j$. ■

If M is indecomposable of finite length, then $\text{End}(M)$ is completely primary, by Fitting's lemma, and Th. 1.7 may be sharpened as follows:

COROLLARY 1.8. *Let M be a distributive module of finite length and let (1) be a direct decomposition into indecomposable modules. Then each $\text{End}(M_i)$ in (2) is completely primary and $\mathbf{J}(\text{End}(M))$ is nilpotent, given by*

$$\mathbf{J}(\text{End}(M)) = \prod_{i=1}^n \mathbf{J}(\text{End}(M_i)). \quad (3)$$

For it is clear from (2) that $\mathbf{J}(\text{End}(M))$ has the form (3); since each $\mathbf{J}(\text{End}(M_i))$ is nilpotent, the same holds of $\mathbf{J}(\text{End}(M))$. ■

In 4.6, below we shall need a further result; we begin with a lemma.

LEMMA 1.9. *Given a module M , let U, V be submodules of M , neither contained in the other, and suppose that U, V each have a unique maximal submodule U', V' , respectively. Then*

$$(U + V)/(U' + V') \cong U/U' \oplus V/V'. \quad (4)$$

Proof. We have

$$\begin{aligned} (U + V)/(U + V') &= ((U + V') + V)/(U + V') \cong V/V \cap (U + V') \\ &= V/(V \cap U) + V' = V/V', \end{aligned}$$

where the last step follows because $U \cap V$ is a proper submodule of V and hence is contained in V' . Similarly, $(U + V)/(U' + V) \cong U/U'$, hence we can represent $U + V$ as a subdirect sum of U/U' and V/V' . The kernel is

$$\begin{aligned}(U + V') \cap (U' + V) &= V' + (U \cap (U' + V)) \\ &= V' + ((U \cap V) + U') = U' + V'.\end{aligned}$$

This shows that the left-hand side of (4) is isomorphic to a submodule of the right-hand side. Now the module on the right has length 2; on the other hand we have seen that $(U + V)/(U + V')$ is simple and

$$(U + V')/(U' + V') \cong U/(U \cap (U' + V')) \cong U/((U \cap V) + U') \cong U/U',$$

so the left-hand side of (4) also has length 2 and the isomorphism follows. ■

PROPOSITION 1.10. *Let M be a module with a unique maximal submodule M' and let N be a distributive module. Then for any two homomorphisms $\alpha, \beta: M \rightarrow N$, either $M\alpha \subseteq M\beta$ or $M\beta \subseteq M\alpha$.*

Proof. We may assume that $\alpha, \beta \neq 0$; then $\ker \alpha$ is a proper submodule of M , hence $\ker \alpha \subseteq M'$, $M\alpha \cong M/\ker \alpha$, and it follows that $M\alpha$ has a unique maximal submodule, namely $M'\alpha$. Similarly, $M\beta$ has a unique maximal submodule $M'\beta$, and if $M\alpha, M\beta$ are incomparable, then by Lemma 1.9,

$$(M\alpha + M\beta)/(M'\alpha + M'\beta) \cong M\alpha/M'\alpha \oplus M\beta/M'\beta \cong (M/M')^2,$$

but this contradicts the fact that N is square-free. ■

Exercises 4.1

1. Show that a finitely generated $\mathbf{Z}$ -module is distributive if and only if it is cyclic.
2. (Jensen [63]) Let M be a distributive R -module. If $a, b \in M$, show that either $Ra \subseteq Rb$ or $Rxb \subseteq Ra$ for some $x \in R$ such that $1 - x$ is a non-unit. Deduce that if R is local ring, $\text{Lat}(M)$ is totally ordered.
3. Let M be a distributive module. Show that if a perspectivity (between two chains of modules) interchanges two simple factors P, Q , then P, Q cannot be isomorphic.
4. Let M be a distributive module of finite length. By the Jordan-Hölder theorem, any two composition series are projective, in the sense that we can pass from one to the other by a series of perspectivities. Show that any such projectivity preserves the order of the factors of a given isomorphism type.
5. Let M be a module. If any homomorphism α from a submodule N of M into M satisfies $N\alpha \subseteq N$, show that M is distributive.

6. Let M be a distributive module of finite length. If $A_1, \dots, A_n$ are the simple factors (with their multiplicities) occurring in a composition series of M , show that every endomorphism of M maps each term of a composition series into itself. Deduce that there is a homomorphism $\text{End}(M) \rightarrow \prod_{i=1}^n \text{End}(A_i)$, whose kernel is the radical of $\text{End}(M)$, and that this radical is nilpotent.

7*. (Vámos [78]) (i) If $M = Ra + Rb$ is distributive, show that $Ra \cap Rb + R(a+b)$ contains Ra and Rb ; deduce that $M = Ra \cap Rb + R(a+b)$. (ii) Show that any finitely generated Artinian distributive module is cyclic. (Hint: Let A be minimal finitely generated non-cyclic, and B minimal of the form $Ra \cap Rb$, where $A = Ra + Rb$. Use (i) to show that $A = R(a+b) + B_0$, where B_0 is a cyclic submodule of B , and that $B_0 \cap R(a+b) \subseteq B$; deduce that $B_0 = B \subseteq R(a+b)$ and obtain a contradiction.)

8. (Stephenson [74]) Let $A = k[x, y]$, $K = k(x, y)$, denote by R, S the subrings of K obtained from A by localizing at all elements prime to x, y , respectively, and let $\alpha: R \rightarrow S$ be the k -isomorphism interchanging x and y . Verify that neither of R, S is contained in the other. Show that $T = \left\{ \begin{pmatrix} a & u \\ 0 & a^\alpha \end{pmatrix} \mid a \in R, u \in K \right\}$ is a ring whose left as well as right ideals are totally ordered, hence T is distributive as left or right T -module, but T is neither left nor right invariant.

9. (Camillo [75]) Let k be a commutative field and α an endomorphism such that $1 < [k:k^\alpha] < \infty$. Show that the ring $R = \left\{ \begin{pmatrix} x & y \\ 0 & x^\alpha \end{pmatrix} \mid x, y \in k \right\}$ has only $R, \mathbf{J}(R)$ and 0 as left ideals, hence $\text{Lat}_R(R)$ is distributive, but R is not left or right invariant.

4.2 Distributive factor lattices

From Th. 1.5.1 we see that a 2-fir may be defined as an integral domain R such that for any $c \in R^\times$ the set $\mathbf{L}(cR, R)$ is a sublattice of the lattice of all right ideals of R . In the commutative case this condition simply states that the principal ideals form a sublattice of the lattice of all ideals. In that case we can go over to the field of fractions and consider the principal fractional ideals; by what has been said they form a modular lattice with respect to the ordering by inclusion. Clearly they also form a group under multiplication, and the group operations respect the ordering. Thus we have a lattice-ordered group; such a group is always distributive, as a lattice (Birkhoff [67], p. 292). This suggests that we single out 2-firs with the corresponding property, and we make the following

Definition. An integral domain R is said to have a *distributive factor lattice*, DFL for short, if for each $c \in R^\times$, the set $\mathbf{L}(cR, R)$ is a distributive sublattice of the lattice of all right ideals in R .

From the definition it is clear that a ring with distributive factor lattice is a 2-fir. Moreover, since $L(cR, R)$ is anti-isomorphic to $L(Rc, R)$ (Th. 3.3.1), the definition is left-right symmetric. We shall reformulate this condition below in a number of ways, in terms of 1-torsion modules. We begin with a technical lemma.

LEMMA 2.1. *Let R be a 2-fir and $a \in R^\times$. Then a is right comaximal with an element similar to a if and only if a is right comaximal with ba , for some $b \in R$. Secondly, there exist two elements similar to a and right comaximal if and only if there is an equation $xa_1y + ua_2v = 1$, where a_1, a_2 are similar to a .*

Proof. The similarity of a and a' can be expressed by the existence of two mutually inverse matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} d' & -b' \\ -c' & a' \end{pmatrix}.$$

Let us replace these matrices by $TA, A^{-1}T^{-1}$, where

$$T = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix},$$

then the equation of comaximality obtained by equating the (1, 1)-entry in $TA \cdot A^{-1}T^{-1}$ is

$$a(d' + b't) - b(c' + a't) = 1. \quad (1)$$

By hypothesis a and a' are right comaximal, say

$$au - a'v = 1, \quad (2)$$

hence $a'vc' + c' = auc'$, and taking $t = vc'$ in (1), we find

$$a(d' + b'vc') - bauc' = 1, \quad (3)$$

which shows a and ba to be right comaximal. Conversely, if a and ba are right comaximal, say $ad' - bac' = 1$, then by Lemma 3.4.3, taking this relation in the form $a \cdot d' - b \cdot ac' = 1$, we have a comaximal relation $ac' \cdot a = a_1v$, where $a_1 \sim a$. Hence a is right comaximal with a_1 as claimed.

For the last part we remark that for any $p, q \in R$, $1 + pq \sim 1 + qp$, as the comaximal relation $(1 + pq)p = p(1 + qp)$ shows. Now if $xa_1y + ua_2v = 1$, then $1 - xa_1y = ua_2v$, hence $1 - a_1yx \sim ua_2v$, say $1 - a_1yx = u'a_3v'$ (cf. Ex. 3.2.6), where $a_3 \sim a_2 \sim a$, thus $1 - u'a_3v' = a_1yx$, and repeating the process, we have $1 - a_3v'u' = a_4z$, where $a_4 \sim a$. Thus the elements a_3, a_4 similar to a are right comaximal. Conversely, if a_1, a_2 are similar to a and right comaximal, then $a_1R + a_2R = R$. ■

We now list some conditions for the distributivity of $L(cR, R)$; it turns out to be more convenient to list the negations, i.e. conditions for non-distributivity:

PROPOSITION 2.2. *Let R be a 2-fir and $c \in R^\times$. Then the following conditions are equivalent:*

- (a) $L(cR, R)$ is not distributive,
 - (b) $c = amb$, where $R/mR \cong M^2 \neq 0$ for some 1-torsion module M ,
 - (c) $c = amb$, where $m = a_1a_2 = a_3a_4$ is a comaximal relation in which $a_1, \dots, a_4$ are all similar non-units.
- (a⁰)–(c⁰). The left–right analogues of (a)–(c).

Proof. By Prop. 1.6, $L(cR, R)$ fails to be distributive precisely if it is not square-free, (b) or contains a 5-element sublattice of length 2, (c), and the symmetry holds by the form of (c). ■

Next we have the following conditions for the global distributivity:

THEOREM 2.3. *Let R be a 2-fir. Then the condition*

- (a) *for any similar a, a' and any $b \in R$, $baR \cap a'R \neq 0 \Rightarrow ba \in a'R$, implies*
- (b) *for any similar $a, a' \in R$, $aR \cap a'R \neq 0 \Rightarrow aR = a'R$,*
and this implies the following conditions, which are equivalent among themselves and to their left–right analogues:
- (c) *for each $c \in R^\times$, the lattice $L(cR, R)$ is distributive,*
- (d) *there is no comaximal relation $a_1a_2 = a_3a_4$, where $a_1, \dots, a_4$ are all similar non-units,*
- (e) *if a is a non-unit, $(R/aR)^2$ is not cyclic,*
- (f) *if a, a' are similar non-units, there is no equation $xay + ua'v = 1$,*
- (g) *for any non-unit a , there is no equation $ax + uav = 1$.*

Moreover, when R satisfies right ACC_1 , then all of the conditions (a)–(g) are equivalent.

We see that DFL is the condition (c); a 2-fir satisfying (a) is sometimes said to possess the *strong DFL-property*.

Proof. Taking $b = 1$ in (a) we find that when $aR \cap a'R \neq 0$, then $aR \subseteq a'R$, and by symmetry, $aR = a'R$, i.e. (b). Now (b) asserts that isomorphic quotient modules of a 1-torsion module have the same kernel; if $L(cR, R)$ is not distributive, it has a 5-element sublattice of length 2, hence we can find

isomorphic quotient modules with distinct kernels, i.e. $\text{non-(c)} \Rightarrow \text{non-(b)}$. In the presence of right ACC_1 we have $(c) \Rightarrow (a)$ by Prop. 1.3, since (a) asserts essentially that any homomorphism $R/acR \rightarrow R/aR$ is induced by an endomorphism of R/aR .

Now $(c) \Leftrightarrow (d) \Leftrightarrow (e)$ by Prop. 2.2 and $(d) \Leftrightarrow (f) \Leftrightarrow (g)$ by Lemma 2.1. ■

We observe that every commutative 2-fir, and in particular, every commutative principal ideal domain, satisfies (a) and hence also (b)–(g). The implication $(c) \Rightarrow (a)$ does not hold without chain conditions (cf. Ex. 16). Some consequences of (a) are listed in the exercises; here we shall consider 2-firs for which (a) fails to hold and show that whenever an element has two similar atomic right factors that are not left associated, then it has (generally) infinitely many. More precisely, we give a lower bound to the number of similarity classes.

Let R be a 2-fir, $c \in R^\times$ and let $c = ab$. Then $M = R/Rc$ has the submodule $N = Rb/Rab \cong R/Ra$. If $c = a'b'$ is another factorization in which $a' \sim a$, then $N' = Rb'/Ra'b' \cong R/Ra'$ is a submodule of M which is isomorphic to N , and $N = N'$ if and only if $Rb = Rb'$, or equivalently, $aR = a'R$, i.e. a, a' are right associated. Suppose now that $N \neq N'$, say $N' \not\subseteq N$. Since N, N' are isomorphic cyclic modules, they have generators u, u' respectively, which correspond under this isomorphism. Given any $\alpha \in \text{End}(N)$, $u\alpha + u'$ generates a submodule N_α of M which is a homomorphic image of N , for the map $f_\alpha: xu \mapsto x(u\alpha + u')$ clearly defines a homomorphism. We note that $f_\alpha \neq 0$, for if $f_\alpha = 0$, then $u\alpha = -u'$, hence $N' = N\alpha \subseteq N$, which is not the case.

Assume now that $N \cap N' = 0$; this means that $Rb \cap Rb' = Rc$ or, equivalently, $aR + a'R = R$, i.e. a and a' are right comaximal. Then the submodules N_α defined by the different endomorphisms of N are all distinct; for if $N_\alpha = N_\beta$, then $u\alpha + u' = x(u\beta + u')$, i.e. $(1-x)u' \in N \cap N' = 0$, so $u' = xu'$, hence $u = xu$ and $u\alpha = xu\beta = u\beta$, therefore $\alpha = \beta$. Thus there are at least as many different submodules N_α as there are elements in $\text{End}(N)$, and each corresponds to a left factor of c similar to a right factor of a . The result may be stated as

THEOREM 2.4. *Let R be a 2-fir and suppose that $c \in R^\times$ has two left factors a, a' that are similar and right comaximal. Then the number of non-right-associated left factors of c that are similar to a right factor of a is at least $\text{card}(\mathbf{E}(Ra))$. In particular, if a is an atom in R , then the number of non-right-associated left factors of c similar to a is 0, 1 or at least $\text{card}(\mathbf{E}(R))$. ■*

If R is an algebra over an infinite commutative field k , any eigenring is an algebra over k and hence infinite. Thus we obtain

COROLLARY 2.5. *Let R be a 2-fir which is also a k -algebra, where k is an infinite field. Then the number of non-right-associated left factors of $c \in R^\times$ similar to a given atom is 0, 1 or infinite. ■*

Finally, we specialize to the case of principal ideal domains. In this case we can give a simple criterion for the distributivity of factor lattices. In any ring R , an element c is called *right invariant* if c is left regular and $Rc \subseteq cR$.

LEMMA 2.6. *In any ring, two similar elements that are right invariant are right associated.*

Proof. Let c be right invariant, then $Rc \subseteq cR$, hence cR annihilates the module R/cR . In fact, cR is the precise annihilator, for if a annihilates R/cR , then $Ra \subseteq cR$ and so $a \in cR$.

Now let $c \sim c'$ and assume that both c, c' are right invariant. Then $R/cR \cong R/c'R$ and equating annihilators, we find that $cR = c'R$, hence $c = c'a$, $c' = cb$, so $c = c'a = cba$. Since c is left regular, $ba = 1$ and similarly $ab = 1$, therefore c and c' are right associated. ■

THEOREM 2.7. *A principal right ideal domain has a distributive factor lattice if and only if every non-zero element is right invariant.*

Proof. Let R be a right PID whose non-zero elements are right invariant. Then any two similar elements are right associated, by Lemma 2.6, and so by Th. 2.3, R has a distributive factor lattice.

Conversely, if R has DFL, then R_R is distributive and any homomorphism $\alpha: R \rightarrow R/cR$ is such that $\ker \alpha \supseteq cR$, by Cor. 1.4; taking α to be the map $x \mapsto \overline{ax}$, we therefore have $ac \in cR$. This holds for all $a \in R$, hence $Rc \subseteq cR$ and c is right invariant. ■

Exercises 4.2

1. Let R be an atomic 2-fir in which no two similar non-units are right comaximal. Show that R has DFL.
2. Show that a right invariant UFD has strong DFL.
3. Show that an atomic 2-fir in which any two atoms are either right associated or comaximally transposable is right invariant and hence has DFL.
4. If c is a right invariant element of finite length in a 2-fir, show that any element similar to c is right associated to c and hence again right invariant. Can the condition on the length of c be omitted?

5. Show that a skew polynomial ring over a field has DFL if and only if it is commutative.

6. Let R be a 2-fir with DFL. If S is a subring which is a 2-fir such that $U(S) = U(R) \cap S$, show that S has DFL.

7. Let R be an atomic 2-fir with DFL. Show that if all atoms in R are similar, then R is a local ring. Show that the 'atomic' hypothesis can be omitted if we assume instead that 'any two non-units have a similar non-unit factor' (i.e. no two non-units are totally coprime).

8*. In a ring R with DFL, show that if $a, b, c \in R$ are such that any two of aR, bR, cR have non-zero intersection, then $aR \cap bR \cap cR \neq 0$. Show that $aR \cap bR \neq 0$, $aR \cap cR \neq 0$ is not enough. (Hint: Take $b = au + 1$, $c = av + 1$ for suitable u, v .) Find a generalization to n terms.

9. Let R be a 2-fir with strong DFL. Show that for any $a, b \in R^*$, $I(abR) \subseteq I(aR)$.

10*. Show that a UFD with strong DFL and with a right large element has a right invariant element and hence is a right Ore domain. Show that the conclusion does not hold for every 2-fir with DFL.

11⁰. In a 2-fir, if $xay + ua'v = 1$, where $a \sim a'$, is a necessarily right comaximal with an element similar to a ?

12⁰. When the conditions of Prop. 2.2 hold, does it follow that $xc - cy = 1$ for some $x, y \in R$?

13⁰. Show that every UFD with strong DFL has elements that are not fully reducible.

14⁰. Determine the structure of non-commutative invariant principal ideal domains.

15. (Bergman [67]) Let K/k be a Galois extension with group $G = \text{Gal}(K/k)$ and let M be any K -bimodule satisfying $\lambda x = x\lambda$ ($x \in M$, $\lambda \in k$). Show that $M = \bigoplus_{\sigma \in G} M_{\sigma}$ ($\sigma \in G$), where $M_{\sigma} = \{x \in M \mid x\alpha = \alpha^{\sigma}x\}$. Let R be a UFD with strong DFL and a k -algebra, containing K but not in its centre. Define R_{σ} as above; show that for any $x \in R_{\sigma}$ $x(x - \alpha) = (x - \alpha^{\sigma})x$ is a comaximal relation, and deduce that for any $x \in R_{\sigma}$ ($\sigma \neq 1$), $1 - x$ is a unit. (Hint: For the last part, replace x by $x\alpha$.)

16. (Brungs and Törner [81]) Let k be a commutative field and let σ be the automorphism of the rational function field $k(x, y)$ interchanging x and y . Denoting by (x) the localization at the set of elements prime to x , show that the ring $R = k[x, y]_{(x)} + tk(x, y)[[t; \sigma]]$ is a principal ideal domain and a chain ring, and hence has DFL. Verify that $y \in U(R)$ and $ytR \subset tR$, so R satisfies (c), but neither (a) nor (b) of Th. 2.3.

4.3 Conditions for a distributive factor lattice

In order to find general conditions for an atomic 2-fir R to have a distributive factor lattice we recall from Th. 2.4 that this is equivalent to

requiring that if an element c has two factorizations

$$c = ab = a_1 b_1 \quad (1)$$

with similar right factors b, b_1 , then these factors must be left associated. Moreover, if an equation (1) holds with b, b_1 similar atoms which are not left associated, then in case R is an algebra over an infinite field, there must be infinitely many right factors similar to b , but pairwise not left associated, by Cor. 2.5. Thus in a sense we have a 1-parameter family of factorizations of c ; this idea may be formalized by adjoining an indeterminate to k . In general this will lead to factorizations of c in $R \otimes k(t)$ which cannot be pulled back to R . Suppose, however, that R is such that every factorization in $R \otimes k(t)$ can be pulled back to R ; this holds, for example, if R is a free algebra. Then this situation cannot arise, and we conclude that similar right factors are necessarily left associated. We shall see how this property can be used to provide us with many examples of rings with a distributive factor lattice. Throughout this section, k is a commutative field, and all tensor products are understood to be over k .

Definition. Let R be a k -algebra; any property X of R is said to be *absolute* if it holds for $R \otimes E$, where E is any algebraic field extension of k ; X is said to be *persistent* if it holds for $R \otimes k(t)$, where t is a central indeterminate.

For example, the free k -algebra $k\langle X \rangle$ is a persistent and absolute fir over k ; if E is a commutative field extension of k , then $E_k\langle X \rangle$ is an absolute fir if and only if E is a regular field extension of k , and a persistent fir if and only if E is algebraic over k .

PROPOSITION 3.1. *Let R be a k -algebra; if R is an absolute integral domain over k , then R is 1-inert in $R \otimes k(t)$. If R is an absolute integral domain and a persistent 2-Hermite ring over k , then $R[t]$ is 2-Hermite.*

Proof. Consider an equation

$$c = ab, \quad (2)$$

where $c \in R$, $a, b \in R[t]$. Since R is an integral domain, it follows that $a, b \in R$, so that R is 1-inert in $R[t]$. Now the first assertion will follow if we show that $R[t]$ is 1-inert in $R \otimes k(t)$, so suppose that $c \in R[t]$ has the factorization (2), where now $a, b \in R \otimes k(t)$. If we multiply (2) by the denominators of a and b , we obtain an equation of the form

$$fc = a'b', \quad f \in k[t]^\times, a', b' \in R[t]. \quad (3)$$

If f has degree zero, it lies in k and may be absorbed in a' . Otherwise we

specialize t to a zero t_0 of f in k_a , the algebraic closure of k , then in $R \otimes k_a$ we have $\alpha'(t_0)b'(t_0) = 0$, hence either α' or b' is divisible by $t - t_0$ and hence by an irreducible factor f_1 of f , say $\alpha' = \alpha'f_1$. Then we can cancel the factor f_1 in (3) and by induction on the degree of f obtain a factorization of c in $R[t]$, which is clearly obtained by pulling the original factorization in $R \otimes k(t)$ down to $R[t]$. This proves the first part.

Suppose now that R is an absolute integral domain and a persistent 2-Hermite ring. Then $R \otimes k(t)$ is 2-Hermite, hence weakly 2-finite, and so is $R[t]$ as a subring. It remains to show that every 1×2 matrix over $R[t]$ with a ring inverse is completable. Let $a \in R[t]^2$ be right invertible over $R[t]$, say $ab = 1$ for some $b \in {}^2R[t]$. Then we can complete a, b to a pair of mutually inverse matrices over $R \otimes k(t)$; clearing denominators, we obtain a relation

$$\begin{pmatrix} a \\ \alpha' \end{pmatrix} (b \quad b') = \begin{pmatrix} 1 & 0 \\ 0 & f \end{pmatrix}, \quad f \in k[t]^\times,$$

$\alpha' \in R[t]^2$, $b' \in {}^2R[t]$. If f has degree 0, this shows a to be completable over $R[t]$; otherwise we have by weak 2-finiteness,

$$(b \quad b'f^{-1}) \begin{pmatrix} a \\ \alpha' \end{pmatrix} = I,$$

and so

$$fba + b'\alpha' = f \cdot I.$$

Therefore, if t_0 is any zero of f in k_a , we have $b'(t_0)\alpha'(t_0) = 0$. Here the column b' is multiplied by the row α' , so we have a set of four equations; since $R \otimes k_a$ is an integral domain, either $\alpha'(t_0) = 0$ or $b'(t_0) = 0$. If $b'(t_0) = 0$, then b' is divisible by an irreducible factor of f and so we can decrease the degree of f . If $\alpha'(t_0) = 0$, then α' is divisible by an irreducible factor f_1 say of f , and we can replace α', b' by $\alpha'f_1^{-1}, b'f_1$. Now we can again reduce the degree of f and complete the proof by induction on the degree of f . ■

Before we come to the main result we need another lemma. For any element $a \in R[t]$ we shall indicate the value of a obtained by specializing t to 0 by a suffix: $a_0 = a(0)$.

LEMMA 3.2. *Let R be a 2-fir which is a k -algebra and*

$$ab = cd \tag{4}$$

an equation holding in $R[t]$ such that b, d are left comaximal in $R \otimes k(t)$, b_0, d_0 are not both 0 and a_0, c_0 are right comaximal in R . Then b_0, d_0 are left comaximal in R .

Proof. Since a_0, c_0 are right comaximal and $a_0 b_0 = c_0 d_0$, this product is the least common multiple of b_0 and d_0 in R (it was only to get this conclusion that we had to assume R to be a 2-fir).

Now b, d are left comaximal in $R \otimes k(t)$, therefore we have an equation

$$pb - qd = f, \quad f \in k[t]^*, \quad (5)$$

in $R[t]$. If $f_0 = 0$, then $p_0 b_0 = q_0 d_0$ and this is a left multiple of $a_0 b_0 = c_0 d_0$. Hence by subtracting a suitable left multiple of (4) from (5) we can modify p, q so that both become divisible by t . We can then divide p, q, f all by t and obtain an equation of the same form as (5), but with f of lower degree. In this way we reduce the degree of f , and continuing this process we eventually reach a case where $f_0 \neq 0$. Taking the constant term of f as 1, we then find that $p_0 b_0 - q_0 d_0 = 1$. ■

We now come to the main result of this section, giving two sets of conditions for DFL.

THEOREM 3.3. *Let R be a k -algebra which is a 2-fir and an absolute integral domain over k .*

- (i) *If R is a persistent 2-fir over k , then it has the strong DFL property.*
- (ii) *If R is a persistent 2-Hermite ring over k , then it has DFL.*

Proof. In case (i) we have to verify condition (a) of Th. 2.3. Suppose that R is a persistent 2-fir and that $b\alpha'R \cap aR \neq 0$, where $a \sim \alpha'$. Say

$$b\alpha'c = ad, \quad c \neq 0$$

and let

$$a\alpha' = u\alpha'$$

be a comaximal relation between a and α' . Then in $R[t]$ we have

$$a(dt + u'c) = (bt + u)\alpha'c. \quad (6)$$

Now any common right factor (in $R \otimes k(t)$) of $dt + u'c$ and $\alpha'c$ can by 1-inertia be taken in R (Prop. 3.1). Hence it must divide $d, u'c$ and $\alpha'c$, i.e. generate a left ideal containing $Rd + Ru'c + R\alpha'c = Rd + Rc = R$. Thus (6) is in fact left comaximal in $R \otimes k(t)$; further, the constant terms of the right factors are not both 0 and those of the left factors are right comaximal in R . Hence by Lemma 3.2, $u'c$ and $\alpha'c$ are left comaximal, i.e. c is a unit and $b\alpha' \in aR$, which is condition (a) of Th. 2.3.

Now assume that R is a persistent 2-Hermite ring. Then by Prop. 3.1, $R[t]$ is 2-Hermite. We shall verify (g) of Th. 2.3, thus consider an equation

$ax - yaz = 1$. In $R[t]$ we have

$$a(tz + x) - (t + y)az = 1. \quad (7)$$

By hypothesis we have mutually inverse matrices

$$P = \begin{pmatrix} a & t+y \\ p & g \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} tz+x & -f \\ -az & q \end{pmatrix}, \quad \text{where } f, g, p, q \in R[t].$$

Since $aza = qp$, it follows that p, q have degree 0 in t . Further,

$$af = (t+y)q, \quad p(tz+x) = gaz. \quad (8)$$

A comparison of degrees shows f and g to be linear in t . We put

$$f = f_1 t + f_0, \quad g = g_1 t + g_0, \quad f_i, g_i \in R,$$

then (8) yields $af_1 = q$, $p = g_1 a$, hence

$$\begin{pmatrix} a & t+y \\ p & g \end{pmatrix} = \begin{pmatrix} 1 & t+y \\ g_1 & g \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix},$$

hence a is a unit, and (g) follows. ■

The free k -algebra $k\langle X \rangle$ is clearly an absolute and persistent fir over k , hence we have

COROLLARY 3.4. *The free k -algebra $k\langle X \rangle$ has DFL.* ■

For non-commutative fields this no longer holds, for even when $|X| = 1$, we have in the principal ideal domain $K[x]$ an equation

$$(x-a)bc - b(x-a)c = 1, \quad \text{where } c = -(ab-ba)^{-1},$$

so by Th. 2.3 (g), $K[x]$ does not have DFL. In fact, by combining the above result with Th. 2.7, we obtain

COROLLARY 3.5. *If a principal ideal domain is an absolute integral domain and a persistent 2-fir over a field, then it is invariant.* ■

Of course, the converse does not hold, since a commutative PID need not be absolute or persistent (cf. 4.6). But there is a partial converse to Th. 3.3.

PROPOSITION 3.6. *Let R be a 2-fir with DFL and assume that R is an algebra over an infinite field k . Then R is 1-inert in $R \otimes k(t)$.*

Proof. Let $c \in R$ have a factorization over $R \otimes k(t)$, then by clearing denomi-

nators we can write this in the form

$$fc = ab, \quad a, b \in R[t], f \in k[t]. \quad (9)$$

Denote the degrees of a, b, f in t by p, q, r , respectively.

Let a be the right ideal generated by the coefficients of $a = a(t)$, considered as a polynomial in t . Then for any $p+1$ distinct element $\alpha_0, \dots, \alpha_p \in k$ we can express all the coefficients of $a(t)$ as k -linear combinations of $a(\alpha_0), \dots, a(\alpha_p)$, hence these $p+1$ elements generate a as right ideal. If we choose $p+1$ such values so as to avoid the zeros of f (at most r in number), then the $a(\alpha_i)$ are left factors of c , hence the $a(\alpha_i)R$ contain cR , and so does their sum: $a = dR$ for some $d \in R$. Similarly, the intersection of $q+1$ of these ideals is a fixed member of $L(cR, R)$, say $eR \subseteq dR$; we assert that equality holds.

If $eR \neq dR$, then the principal ideals between eR and dR form a distributive lattice $\neq 0$, and hence (Appendix 1) there is a homomorphism of this lattice onto the 2-element lattice $2 = [0, 1]$ such that $eR \mapsto 0, dR \mapsto 1$. Since the sum of the ideals $a(\alpha)R$, as α ranges over $p+1$ values of k , avoiding the zeros of f , is the maximal element dR , which maps to 1, it follows that $a(\alpha)R$ can map to 0 for at most $p+r$ values of α . But the intersection of the $a(\alpha)R$, as α ranges over any $q+1$ values as before, is eR , which maps to 0, so $a(\alpha)R$ maps to 1 for at most $q+r$ values. Since k is infinite, we have a contradiction, which shows that $eR = dR$. Now d is a left factor of $a(t)$ in $R \otimes k(t)$ and e a right multiple, hence $a(t)$ is associated to d in $R \otimes k(t)$, so our factorization of c can be pulled down to R , as claimed. ■

We note, however, that in the situation of this proposition R need not be a persistent 2-fir, e.g. take $R = E[x]$, where E is a commutative field extension of k which is not algebraic over k .

Finally we show how the weak algorithm can be used to obtain a good supply of persistent 2-firs.

PROPOSITION 3.7. *Let R be a filtered k -algebra such that every element of degree 0 lies in k . If R satisfies the 2-term weak algorithm, then it is a persistent 2-fir over k .*

Proof. Let $A = \text{gr } R$ be the graded ring associated with R ; then A satisfies the 2-term weak algorithm as a graded ring. Moreover,

$$\text{gr}[R \otimes k(t)] = A \otimes k(t),$$

where $A \otimes k(t)$ is the graded ring with i th component $A_i \otimes k(t)$. So we have to check the 2-term weak algorithm for $A \otimes k(t)$. Let

$$ab' = ba' \quad (10)$$

be a non-trivial homogeneous relation and suppose that $v(a) \geq v(b)$, say; we have to find $q \in A \otimes k(t)$ such that $a = bq$. On clearing denominators in (10) (and changing the notation appropriately) we may assume that $a, b, \alpha', b' \in A[t]$, and that no two of these elements that occur on opposite sides of (10) have a common factor in $k[t]$. On equating terms independent of t in (10) we find

$$a_0 b'_0 = b_0 \alpha'_0. \quad (11)$$

If one of these elements is 0, say $a_0 = 0$, then $\alpha'_0 = 0$ or $b_0 = 0$, so either a, α' or a, b have a factor t in common, contrary to hypothesis. Thus no term in (11) is 0. By the 2-term weak algorithm in A , there exists $q_0 \in A$ such that $a_0 = b_0 q_0$. If we replace a by $a - b q_0$ and α' by $\alpha' - q_0 b'$ in (10), we obtain an equation of the same form as (10), with a, α' divisible by t , and the degrees are unchanged, so we still have $v(a) \geq v(b)$. On dividing by t and using induction on the degree in t , we find that $a - b q_0 = b q_1$, whence $a = b(q_0 + q_1)$, as claimed. ■

Exercises 4.3

1. Let R be a k -algebra which is an absolute integral domain, a persistent 2-fir and a right Ore domain. Show that any two similar elements are right associated.

2. Let R be a k -algebra which is an absolute integral domain and a persistent 2-fir. Show that if two elements a, α' of R are similar in $R \otimes k(t)$, then they are similar in R .

3. Let R be a k -algebra which is an integral domain. Show that an atom in R remains an atom in $R \otimes k(t)$.

4. Let R be an n -fir over an algebraically closed field k . Show that every $n \times n$ matrix over $R[t]$ is inert in $R \otimes k(t)$. (Hint: The transforming matrix can be taken to lie in the subgroup generated by $\text{GL}_n(R[t])$ and the diagonal matrices over $k(t)$.)

5*. Let R be a k -algebra which is an absolute n -fir and a persistent $2n$ -fir. Given a full matrix c in R_n such that any factor of c in $R[t]_n$ can be reduced to one in R_n on multiplying by an element of $\text{GL}_n(R[t])$, show that the corresponding module R_n/cR_n is distributive. Taking $n = 2$ and p, q dissimilar atoms in R , show that $c = p \oplus q$ satisfies the above hypotheses, but not $p \oplus p$. Verify that when $c = p \oplus p$, R_2/cR_2 is not distributive.

6. In the complex-skew polynomial ring $R = \mathbb{C}[x; -]$ show that $x^2 - 1$ has the factorizations $x^2 - 1 = (x - u)(x + \bar{u})$, where u ranges over the unit circle. Obtain the corresponding factorizations over $\mathbb{C}(t)[x; -]$ with $u = (t + i)(t - i)^{-1}$ and show that these cannot be pulled back to R .

7. Let R be a k -algebra, M an R -module, P a submodule of M with inclusion map $i: P \rightarrow M$ and $f: P \rightarrow M$ a homomorphism. Assume that the image of the homomor-

phism $i + tf$ of $P \otimes k(t)$ into $M \otimes k(t)$ is of the form $N \otimes k(t)$ for some $N \subseteq M$. Show that $Pf \subseteq P$ and hence obtain another proof of Th. 3.3(i).

8°. Investigate Prop. 3.6 when the field k is finite.

9. (Beauregard [80]) Let E be a commutative field with an automorphism α and F a subfield mapped into a proper subfield of itself by α . Write $R = E[[x; \alpha]]$, $P = \{f \in R \mid f(0) \in F\}$. Verify that R is a local ring and a principal ideal domain; hence show that P is an Ore domain which is right invariant but not left invariant.

10. In a free algebra, if $ab = bc^2$, show that a is a square (cf. Ex. 2.7.12).

11°. Show that in a 2-fir with DFL, for any atom c and any $n \geq 1$, c^n is indecomposable. Investigate atomic 2-firs in which the length of indecomposable elements is bounded.

12. Let R be an atomic 2-fir with DFL. Show that if R is a matrix local ring, then its capacity must be 1, i.e. it is a scalar local ring.

4.4 Finite distributive lattices

In an atomic 2-fir with distributive factor lattice, the left factors of a given non-zero element form a distributive lattice of finite length. For a closer study of this lattice we shall in this section describe its structure in terms of partially ordered sets.

Let us denote by Pos the category of finite partially ordered sets, with isotone (i.e. order-preserving) maps as morphisms. By DL we shall denote the category of all distributive lattices of finite length, with lattice-homomorphisms, i.e. maps preserving meet, join, 0 and 1, as morphisms. In each of these categories $2 = \{0, 1\}$ denotes the chain of length 1. We begin by defining two functors between these categories.

Take $P \in \text{Pos}$ and consider $P^* = \text{Hom}_{\text{Pos}}(P, 2)$; this set P^* may be regarded as a finite distributive lattice, namely a subset of 2^P . The elements of P^* —isotone maps from P to 2 —may also be described by the subsets of P mapped to 0. They are precisely the *lower segments* of P , i.e. subsets X with the property

$$a \in X, \quad b \leq a \quad \text{implies } b \in X.$$

A lower segment of the form $\lambda_a = \{x \in P \mid x \leq a\}$ is said to be *principal*. We observe that the partially ordered set of all principal lower segments of P is isomorphic to P , as member of Pos .

Clearly each $\alpha \in P^*$ is completely determined by the lower segment mapped to 0, and every lower segment defines such a map. Hence P^* may also be identified with the set of lower segments of P .

Next take $L \in \text{DL}$ and write $L^* = \text{Hom}_{\text{DL}}(L, 2)$. Here we can regard L^* as a partially ordered set, writing $f \leq g$ if and only if $xf \leq xg$ for all $x \in L$. The elements of L^* may be characterized by the subsets they map to 1. Given $f \in L^*$, let $a \in L$ be the meet of all x satisfying $xf = 1$. Then $af = 1$ and a is the unique minimal element with this property. Clearly $a > 0$ and if $a = x \vee y$, then $1 = xf \vee yf$, hence $xf = 1$ or $yf = 1$, i.e. $x \geq a$ or $y \geq a$; thus a is *join-irreducible*:

$$a \neq 0 \text{ and } a = x \vee y \text{ implies } x = a \text{ or } y = a.$$

Conversely, any join-irreducible element a gives rise to an $f \in L^*$, defined by the rule $xf = 1$ if and only if $x \geq a$. We may thus identify L^* with the partially ordered set of its join-irreducible elements.

Our object will be to show that $*$ establishes a duality between DL and Pos. We note that the finite chains are objects in both DL and Pos, and $L^* \cong L$ holds precisely if L is a finite chain. Before coming to the main result we need a lemma on objects in DL, but this is just as easily proved in a more general setting:

LEMMA 4.1. *In any lattice with minimum condition, each element is the join of the join-irreducible elements below it.*

Proof. Let L be a lattice with minimum condition which does not satisfy the conclusion; then we can find $a \in L$ such that a is not the join of join-irreducible elements. We take a to be minimal with this property, then a cannot be join-irreducible and $a \neq 0$, because 0 is the join of the empty family. Hence $a = b \vee c$, where $b < a, c < a$. By the minimality of a , both b and c are joins of join-irreducible elements, hence so is $a = b \vee c$. ■

THEOREM 4.2. *The categories Pos and DL are dual to each other, via the contravariant functors*

$$\begin{aligned} P &\mapsto P^* = \text{lattice of lower segments of } P \text{ and} \\ L &\mapsto L^* = \text{set of join-irreducible elements of } L. \end{aligned}$$

Moreover, if P and L correspond, then $\cap, \cup$ in P correspond to $\wedge, \vee$ in L and the length of L equals $|P|$.

Proof. It is clear that two contravariant functors are defined between these categories by means of $\text{Hom}(-, 2)$; it only remains to show that $P^{**} \cong P, L^{**} \cong L$.

Let $P \in \text{Pos}$, then P^* consists of all lower segments of P . If $\alpha \in P^*$ and $a_1, \dots, a_r$ are the different maximal elements of the lower segment α , then $x \in \alpha$

if and only if $x \leq a_1$ or ... or $x \leq a_r$. Hence $\alpha = \lambda_{a_1} \vee \dots \vee \lambda_{a_r}$, where λ_c is the principal lower segment defined by c . This shows α to be join-irreducible if and only if it is principal, and so P^{**} , the set of join-irreducible lower segments, is just the set of principal lower segments, which we saw is isomorphic to P .

Next, given $L \in \text{DL}$, consider L^* , the partially ordered set of its join-irreducible elements. This set determines L , by Lemma 4.1: each $a \in L$ can be represented by the set of join-irreducible elements $\leq a$, and the set of join-irreducible elements occurring are just the lower segments, thus $L^{**} \cong L$.

Let L and P correspond under this duality, and suppose that P has n elements. Then we can form a chain in L by picking a minimal element $a_1 \in P$, next a minimal element a_2 in $P \setminus \{a_1\}$, etc. Therefore every chain has n elements. ■

It is clear that every P^* is finite, as subset of 2^P , hence we obtain

COROLLARY 4.3. *Any distributive lattice of finite length n is finite, with at most 2^n elements.* ■

The interest of the duality described in Th. 4.2 resides in the fact that for any $L \in \text{DL}$ and $P \in \text{Pos}$ that correspond under the duality, P is usually much

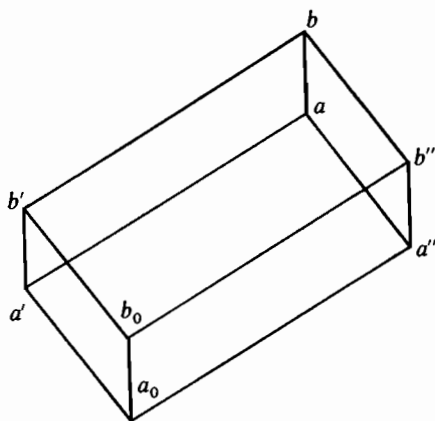
simpler than L . For example, the lattice  corresponds to the set 

and the free distributive lattice on three generators, a lattice of height 6 with 18 elements (cf. Grätzer [78], p. 38) corresponds to the three-cornered crown



There is another way of describing the correspondence of Th. 4.2 which is of importance for us in what follows.

In a distributive lattice of finite length, every link is projective to exactly one link with join-irreducible upper end-point. For let $[a, b]$ be a link in which b is join-reducible, say $b = x \vee y$, $x < b$, $y < b$. Then each of $x \vee a$, $y \vee a$ lies in $[a, b]$, and so equals a or b , but not both can be a , say $x \vee a = b$. Then $[a, b]$ is perspective to $[x \wedge a, x]$, a link which is lower in the sense that $l[0, x] < l[0, b]$. It follows that the lowest link in the projectivity class of $[a, b]$ has a join-irreducible upper end-point. But such a lowest link is unique in a distributive lattice: if $[a, b]$ is in perspective with lower links $[a', b']$ and $[a'', b'']$, then we have a cube diagram as shown:



$b = a \vee b' = a \vee b''$, $a \wedge b' = a'$, $a \wedge b'' = a''$; put $b_0 = b' \wedge b''$, then $b_0 \wedge a' = b' \wedge b'' \wedge a' = b' \wedge b'' \wedge a = a_0$ say, and by symmetry, $b_0 \wedge a'' = a_0$, while $a' \vee b_0 = (a \wedge b') \vee (b'' \wedge b') = (a \vee b'') \wedge b' = b \wedge b' = b'$; similarly $a'' \vee b_0 = b''$. This shows that in each projectivity class of links there is a unique lowest link, whose upper end-point is necessarily join-irreducible, after what has been said.

It follows that the join-irreducible elements correspond to the projectivity classes of links. If our lattice L has length n , then there are just n join-irreducible elements and hence n projectivity classes of links. Since each class has a representative in each chain (by the Jordan-Hölder theorem), there is exactly one representative from each class in each chain. Thus we have proved most of

PROPOSITION 4.4. *Let L be a distributive lattice of length n . Then the links of L fall into exactly n projectivity classes and each chain in L contains one link from each class. Moreover, each projectivity class contains a unique lowest link, and its upper end-point is join-irreducible. Given projectivity classes α and β , we have $\alpha < \beta$ if and only if the link from α occurs below that from β in every chain.*

It remains to prove the last assertion. Let $a, b \in P$ and let α, β be the projectivity classes of links corresponding to a, b , respectively. If a, b are incomparable in P , then we can form chains in P^* in which the representative of α lies lower than that of β , and chains in which it lies higher, depending on whether we choose a before b or b before a in forming the chain. But if $a < b$, in P , then we must choose a before b and hence in every chain the representative of α lies lower than that of β . ■

Exercises 4.4

1. Show that a modular lattice has finite length if and only if every chain in it is finite. Give examples of (i) an infinite modular lattice of finite length and (ii) a general lattice, all of whose chains are finite, but their lengths are unbounded.

2. Show that a finite distributive lattice is complemented if and only if the corresponding partially ordered set is totally unordered. (Hint: In a Boolean algebra, the join-irreducible elements are precisely the minimal non-zero elements.)

3. Show that a finite distributive lattice is indecomposable (as a direct product) if and only if the corresponding partially ordered set is connected.

4. Examine how the correspondence of Th. 4.2 is affected if we take (i) upper instead of lower segments, (ii) meet- instead of join-irreducible elements and (iii) make both these changes.

5. Let L be a modular lattice of finite length in which any two projective intervals are perspective. Show that L must be distributive. Determine all such lattices, using Th. 4.2.

4.5 The factor lattice

Let R be an atomic 2-fir with a distributive factor lattice. For each $c \in R^\times$, $L_c = L(cR, R)$ is a distributive lattice of finite length. We shall write P_c for the corresponding partially ordered set L_c^* . Each atomic factorization of c :

$$c = p_1 p_2 \cdots p_n \quad (1)$$

corresponds to a chain in L_c ; if $c = q_1 \cdots q_n$ is another atomic factorization of c , then p_i is said to be *equivalent* to q_j if we can pass from p_i to q_j by a series of comaximal transpositions. Here p_i refers not to an element of R but to its occurrence in the factorization (1) of c ; thus in xyx (in a free algebra) the two factors x are inequivalent. Since comaximal transpositions correspond to perspectivities in L_c , the equivalence classes of (occurrences of) atomic factors correspond to projectivity classes of links in L_c , and thus to elements of P_c . We shall refer to an equivalence class of atomic factors of c as an *abstract atomic factor* of c ; thus P_c may be thought of as the set of abstract atomic factors of c . By Prop. 4.4, each abstract atomic factor has just one representative in each complete factorization of c , and of two abstract atomic factors, p and q say, p precedes q if it occurs on the left of q in every complete factorization of c . On the other hand, when p, q are incomparable, then they may be comaximally transposed whenever they occur next to each other in a complete factorization. Every complete factorization is completely determined by the order in which the abstract factors occur; in particular, an element with n factors cannot have more than $n!$ complete factorizations.

Any expression of c as a product $c = ab$ corresponds to a decomposition of P_c into a lower and a complementary upper segment, which may be identified with P_a, P_b , respectively. Given two factorizations

$$c = ab' = ba', \quad (2)$$

we see that the highest common left factor and least common right multiple of a, b will correspond to the intersection and union, respectively, of P_a, P_b . In particular, a comaximal relation (2) for c corresponds to an expression of P_c as a union of two disjoint lower segments, which means a partition of its diagram into two disconnected components. We note also that in this case $L_c \cong L_a \times L_b$, in agreement with Prop. 3.5.6.

Recalling that projective links in L_c correspond to similar factors, we see that with every element of $P_c = L_c^*$ we can associate a similarity class of atoms in R . Abstract factors corresponding to the same similarity class must be comparable in P_c because similar atoms cannot be comaximally transposed in R (by Th. 2.3), hence every similarity class corresponds to part of a chain within P_c . It follows that the only automorphism of P_c preserving similarity classes is the identity; hence the same holds for L_c . Thus, for any similar elements c and c' the isomorphism between L_c and $L_{c'}$ (and between P_c and $P_{c'}$) is unique. We state this as

PROPOSITION 5.1. *Let R be an atomic 2-fir with distributive factor lattice. Then for any similar elements c, c' of R there is a unique isomorphism $L_c \rightarrow L_{c'}$ between the factor lattices. ■*

If $f: R \rightarrow R'$ is a homomorphism of atomic 2-firs with DFL, then for any $c \in R$ such that $c \notin \ker f$, we get a lattice homomorphism from L_c to L_{cf} : the obvious map preserves HCLF's because it preserves comaximality and it preserves LCRM's by the factorial duality. By Th. 4.2, a homomorphism in the opposite direction is induced from P_{cf} to P_c .

In a commutative principal ideal domain, or indeed in any commutative UFD, two atoms are coprimely transposable if and only if they are non-associated. It follows that the only possible structures for the sets P_c in this case are disjoint unions of finite chains. For example, in $\mathbb{Z}$, $720 = 2^4 \cdot 3^2 \cdot 5$, hence P_{720} consists of three chains, of lengths 3, 1 and 0. By contrast, in the non-commutative case all possible structures for P_c are realized:

THEOREM 5.2. *Let $A_n = k\langle x_1, \dots, x_n \rangle$ be the free k -algebra of rank n . Given any partially ordered set P of n elements, there exists $c \in A_n$ with $P_c \cong P$.*

Proof. The case $n = 0$ is clear, so assume that $n > 0$ and let α be any element

of P . Assume inductively that we have found $c_0 \in A_{n-1}$ such that $P_{c_0} \cong P' = P \setminus \{\alpha\}$.

Write $P' = U \cup V \cup W$, where U is the set of elements $< \alpha$ in P , V the set of elements incomparable with α and W the set of elements $> \alpha$. Clearly U , $U \cup V$ and $U \cup V \cup W$ are lower segments of P' ; they correspond to left factors a , ab and $abd = c_0$ of c_0 . Put $c = a(bx_n + 1)bd$, then we assert that $P_c \cong P$.

In the first place, $bx_n + 1$ is an atom, since it is linear in x_n and in any factorization the term independent of x_n must divide 1. We now identify P_c with P by letting the factors of a , b and d correspond as in the identification of P_{c_0} and P' , and letting the abstract atomic factor to which $bx_n + 1$ belongs correspond to α . It remains to check that the partial ordering of P_c agrees with that of P .

Since ab is a left factor of c and bd is a right factor, the orderings on the corresponding subsets of P_c will agree with those on P_{ab} and P_{bd} , as required. The new abstract factor is incomparable with the factors of b , because of the comaximal relation $(bx_n + 1)b = b(x_nb + 1)$. Now the partial ordering will be completely determined if we show that the factor corresponding to $bx_n + 1$ lies above all factors of a and below all factors of d . By symmetry it suffices to prove the first statement.

Suppose the contrary; then for some non-unit right factor e of a , we would have a comaximal relation $e(bx_n + 1) = fe'$. Now we obtain a ring homomorphism of A_n into A_{n-1} by putting $x_n = 0$; this will preserve comaximal relations and hence maps f to an element similar to $b \cdot 0 + 1 = 1$, i.e. a unit. However, f is similar to the non-unit $bx_n + 1$ in A_n and so must involve x_n . But then the product $e(bx_n + 1) = fe'$ will involve monomial terms in which x_n occurs, but is not the last factor (since e' is a non-unit). This is a contradiction, and it shows that every factor of a lies below $bx_n + 1$. ■

In fact all these partially ordered sets may already be realized in A_2 . We shall prove this by showing that A_n (for any n) can be embedded as a 1-inert subring in A_2 ; the next result shows more generally how the free algebra of countable rank can be embedded 1-inertly in A_2 .

THEOREM 5.3. *The free algebra of countable rank can be embedded 1-inertly in the free algebra of rank 2.*

Proof. Consider the free algebra $A_2 = k\langle x, y \rangle$, with the usual grading by total degree in x and y . We shall also $\mathbb{Z}$ -grade it by associating with each monomial p the function $i(p) = d_x(p) - d_y(p)$, the degree in x minus the degree in y , which we shall call the *imbalance* of p . The *balanced* elements of A_2 , i.e.

elements that are homogeneous of imbalance 0, form a subalgebra which we shall denote by B .

Let U be the class of monomials $u \neq 1$ such that $i(u) = 0$ but $i(u') > 0$ for every left factor $u' \neq 1, u$; note that every such u must begin with a factor x . Now A_2 has an automorphism $-:u \mapsto \bar{u}$ of order 2 which consists in interchanging x and y , and this automorphism maps B into itself. We claim that $U' = \{u, \bar{u} | u \in U\}$ is a free generating set of B , as free k -algebra. This is clear since every balanced monomial can be decomposed in a unique manner into factors in U' by writing it out as a product of x 's and y 's and putting a dividing mark at every point having the same number of x 's as y 's to the left of it.

It follows that the subalgebra C generated by the elements $u + \bar{u}$ ($u \in U$) is also free. Clearly it has infinite rank; we shall prove that it is 1-inert in A_2 . To do this we need only show that for any $c \in C$, $c = ab$ with $a, b \in A_2$ implies $a, b \in C$. This simplification arises because every unit of A_2 already lies in C .

We shall take the monomials of A_2 to be ordered by total degree and those of a given degree lexicographically; then the leading term in A_2 of any element $(u_1 + \bar{u}_1) \dots (u_n + \bar{u}_n)$ ($u_i \in U$) will be $u_1 \dots u_n$, because u is the leading term of $u + \bar{u}$. It follows that the leading term in every non-zero element of C is of this form, and that every element of $B \setminus C$ can be written as the sum of a term in C and an element whose leading term contains factors $\bar{u}$ ($u \in U$) in its decomposition.

Now let $c = ab$, where $c \in C$, $a, b \in A_2$. Since $c \in B$, it is i -homogeneous of imbalance 0, hence a and b are also i -homogeneous of imbalance m and $-m$, respectively. It follows that every monomial term of maximal degree occurring in c has the property that if we split it into a left factor of degree $d(a)$ and a right factor of degree $d(b)$, the former will have imbalance m and the latter imbalance $-m$. But combining this with the fact that $c = \bar{c}$, we see that m must be 0; hence a and b lie at least in B .

Note that the leading term of c is the product of the leading terms of a and b , hence neither of these leading terms can involve a $\bar{u}$ in its factorization into elements of U and $\bar{U}$. Let us write $a = a' + f$, $b = b' + g$, where $a', b' \in C$ and f, g are elements of B which are 0 or have leading terms involving at least one factor $\bar{u}$. Our earlier remark shows that f, g precede a', b' , respectively, in the lexicographic ordering. Now C will contain $c - a'b' = (a' + f)(b' + g) - a'b' = a'g + fb' + fg$. Here the leading term of $a'g$, if non-zero, will involve a $\bar{u}$ in its right subsegment of degree $d(g)$ but not to the left thereof, and we can make parallel observations about the leading terms of fb' and fg . We deduce that no two of these leading terms, insofar as they occur, are equal, and hence that the leading term of our sum is one of these, unless $f = g = 0$. But, being in C , the leading term of our sum cannot involve a $\bar{u}$, so we must have $f = g = 0$, i.e. $a, b \in C$. ■

Exercises 4.5

1. Let R be an atomic 2-fir with DFL. Show that a factorization $c = a_1 \dots a_n$ corresponds to an isotone map of P_c into the ordered set of n elements.

2. A subset X of a partially ordered set is called *convex* if $x, y \in X$, $x < a < y$ implies $a \in X$. Show that a subset X of P_c is convex if and only if c has a factorization $c = aub$, where $P_u = X$, and if $c = a'u'b'$ is another such factorization, then u' is obtainable from u by a series of comaximal transpositions.

3. Let R be an atomic 2-fir with DFL. Given two factorizations $c = a_1 b_1 = a_2 b_2$ of an element c of R , if each similarity class contributes at least as many terms to a factorization of a_1 as it does to a factorization of a_2 , show that $a_1 \in a_2 R$.

4. Find elements in $k\langle x, y \rangle$ with factor lattices corresponding to the following partially ordered sets:



5. In $k\langle x, y, z \rangle$ show that the factor lattice of

$$(x^2 + x)[(y(x^2 + x) + 1)z + 1](y(x^2 + x) + 1)$$

is the free distributive lattice on three free generators.

6. (M. L. Roberts) Let $*$ be the antiautomorphism of $k\langle x, y \rangle$ interchanging x and y . Show that if $P = \begin{pmatrix} y & xy^2 \\ x & yx^2 \end{pmatrix}$, then P^*P is an atom of C (in the notation of Th. 5.3, proof). Deduce that P^*P is not inert in $k\langle x, y \rangle$. Show also that the embedding of C in $k\langle x, y \rangle$ does not preserve full matrices (i.e. is not honest). (Hint: Use $Q = \begin{pmatrix} y & xy^2 & xyxy^2 \\ x & yx^2 & yxyx^2 \end{pmatrix}$.)

4.6 Eigenrings

We have seen in 0.7 that the eigenring of a matrix is just the endomorphism ring of the module defined by the matrix. In the case of a fir the eigenrings are as a rule small, only exceptionally are they large. So one can expect two kinds of results on eigenrings: (a) eigenrings are always small and (b) under suitable conditions eigenrings are large. Thirdly, in the case of a ring with DFL we have (c) the consequences of distributivity for the eigenring. Our results will mainly be under the headings (a) and (c). Thus when A is a regular matrix over R , we shall show that (i) when R is a persistent semifir over k , then the eigenring of A is algebraic over k (Th. 6.4) and (ii) when R is a free k -algebra, then the eigenring of A is finite-dimensional over k . This will be proved in a special case (column matrices) in Prop. 6.8, the general case being reserved for Ch. 5.

As before, all our rings will be algebras over a commutative field k ; hence the eigenring of an element or a matrix will also be a k -algebra. If the eigenring of C is k itself, we shall say that C has a scalar eigenring.

From Cor. 1.8 we obtain the following result on the structure of eigenrings.

PROPOSITION 6.1. *Let R be an atomic 2-fir with distributive factor lattice. Then the eigenring of any $c \in R^\times$ is a direct product of a finite number of completely primary rings. ■*

Beyond this rather general fact it seems difficult to apply the results on distributive modules to the study of arbitrary rings with DFL. In what follows we shall therefore put further restrictions on the ring; most of these will be satisfied by free algebras.

Let R be a k -algebra; we recall that R is said to be algebraic over k , if every element of R satisfies a polynomial equation over k . If the matrix ring R_n is algebraic over k , for all $n \geq 1$, then R is said to be matrix algebraic over k . When R is a field, this condition can be expressed in terms of the rational function field $R(t)$ (cf. 8.5), but no examples are known of algebraic fields that are not matrix algebraic. However, for eigenrings we shall be able to reduce one condition to the other by means of the following lemma.

LEMMA 6.2. *Let R be a ring and $A \in {}^m R^n$ any matrix with left eigenring E , then $A \oplus \dots \oplus A$ (r terms) has left eigenring E_r .*

Proof. E is the endomorphism ring of the left R -module $M = R^n/(R^m)A$, while $B = A \oplus \dots \oplus A$ defines the left R -module $R^{rn}/(R^{rm})B \cong M^r$, and clearly this has endomorphism ring E_r . ■

For our first main result we need a lemma on inertia.

LEMMA 6.3. *Let S be a ring containing a central regular element t such that $\bigcap t^n S = 0$, and such that $R = S/tS$ is a semifir. If the induced map $\text{GL}_n(S) \rightarrow \text{GL}_n(R)$ is surjective, for all $n \geq 1$, then every matrix over S is inert in $S[t^{-1}]$.*

Here $S[t^{-1}]$ denotes of course the ring obtained from S by adjoining an inverse of t ; since t is central and regular, S is embedded in $S[t^{-1}]$.

Proof. Write $x \mapsto \bar{x}$ for the natural homomorphism $S \rightarrow S/tS = R$; this amounts to putting $t = 0$. We take $A \in {}^m S^n$ and suppose that over $S[t^{-1}]$:

$$A = PQ, \quad \text{where } P \text{ is } m \times r, Q \text{ is } r \times n \text{ and } r = \min\{m, n\}. \quad (1)$$

If P or Q is 0, there is nothing to prove, so we may assume that $P, Q \neq 0$. Now every non-zero matrix B over $S[t^{-1}]$ can be written in the form $t^v B'$, where $v \in \mathbb{Z}$, B' has entries in S and $\bar{B}' \neq 0$. Hence on changing notation, we can rewrite (1) as

$$A = t^{-v} P Q, \quad \text{where } P \in {}^m S^r, Q \in {}^s S^n, \bar{P}, \bar{Q} \neq 0. \quad (2)$$

We have to show that v can be taken to be ≤ 0 , so suppose that $v > 0$; we shall show how to replace v by $v-1$ in (2). If $v > 0$, then $\bar{P}\bar{Q} = 0$; since R is a semifir, we can find a matrix $U_1 \in \text{GL}_r(R)$ trivializing this relation, and by hypothesis we can lift U_1 to $U \in \text{GL}_r(S)$. Hence on replacing P, Q by $PU, U^{-1}Q$ we find that all the columns in $\bar{P}$ after the first s are 0, while the first s rows of $\bar{Q}$ are 0. Now multiply P on the right by $V = tI_s \oplus I_{r-s}$ and Q on the left by V^{-1} ; then P becomes divisible by t , while Q still has entries in S . In this way we can, by cancelling t , replace v by $v-1$ in (2) and after v steps we obtain the same equation with $v = 0$; this shows A to be inert in $S[t^{-1}]$. ■

We note that the condition ' $\text{GL}_n(S) \rightarrow \text{GL}_n(R)$ is surjective' is satisfied under any of the following assumptions:

- (a) $R = S/tS$ is a retract of S , i.e. there is a homomorphism $R \rightarrow S$ such that the composition with the natural homomorphism $R \rightarrow S \rightarrow S/tS$ is the identity,
- (b) R is a GE_n -ring: $\text{GL}_n(R)$ is generated by elementary and diagonal matrices or
- (c) t lies in the Jacobson radical of S .

The verification is straightforward and may be left to the reader.

To illustrate the lemma, let R be any semifir, $R[[t]]$ the ring of formal power series in a central indeterminate and $R((t))$ the ring of formal Laurent series. Then, by the lemma, every matrix over $R[[t]]$ is inert in $R((t))$. Secondly, let R be a semifir which is a k -algebra, and denote by $R[t]_{(t)}$ the localization of $R[t]$ at the set of all polynomials in t over k with non-zero constant term. Then every matrix over $R[t]_{(t)}$ is inert in $R \otimes k(t)$.

We can now prove our first main result on algebraic eigenrings:

THEOREM 6.4. *Let R be a k -algebra which is a persistent semifir over k . Then the eigenring of any regular matrix over R is matrix algebraic over k .*

Proof. If $A \in {}^m R^n$ is regular, then $B = A \oplus \dots \oplus A$ (r terms) is again regular, and to show that A has a matrix algebraic eigenring we must show, by Lemma 6.2, that the eigenring of B is algebraic, for all $r \geq 1$. So it is enough to show that the eigenring of A is algebraic and then apply the result to B .

Take $P \in R_n$, $P' \in R_m$ such that $AP = P'A$, then in $R \otimes k(t)$ we have

$$A(I - tP) = (I - tP')A. \quad (3)$$

Let us show that A and $I - tP'$ are left coprime. If Q is a square common left factor, we have

$$(A, I - tP') = Q(S, T) \quad \text{over } R \otimes k(t). \quad (4)$$

By Lemma 6.3 and the remark following it we obtain such a factorization over $R[[t]]$, and by moving any denominators from (S, T) to Q we may assume that S has entries in $R[[t]]$. Now put $t = 0$ in (4), then we obtain $\bar{Q}\bar{T} = I$. Since R is weakly finite, $\bar{Q}$ is invertible over R , hence Q is invertible over $R[[t]]$. Over this ring we can therefore rewrite the equation $A = QS$ as $S = Q^{-1}A$. But S has entries in $R[[t]]$ and A is regular over R ; it follows that Q^{-1} involves only finitely many powers of t , and so has entries in $R[t] \subseteq R \otimes k(t)$. This shows that Q is invertible over $R \otimes k(t)$ and so $A, I - tP'$ are left coprime.

By symmetry A and $I - tP$ are right coprime, thus (3) is a coprime relation, hence it is comaximal (Cor. 3.2.5). Replacing t by $u = t^{-1}$, we obtain a relation

$$A(uI - P) = (uI - P')A,$$

still comaximal in $R \otimes k(t) = R \otimes k(u)$. Writing down a relation of comaximality and clearing denominators in u , we obtain

$$CA + D(uI - P) = fI, \quad C \in {}^n R[u]^m, D \in R[u]_n, f \in k[u]^\times. \quad (5)$$

We now write all powers of u on the right of the coefficients and substitute P for u . This is permissible since the substitution $u \mapsto P$ respects right multiplication by matrix polynomials whose coefficients are matrices commuting with P . If $C = \sum C_i u^i$, then the first term in (5) is $\sum C_i A P^i = \sum C_i P^i A$, while the second term vanishes. Thus (5) reduces to $EA = f(P)$, which means that P satisfies an equation mod $R_m A$, and this shows that $E(A)$ is algebraic over k . By the initial remark, $E(A)$ is also matrix algebraic, as we had to show. ■

COROLLARY 6.5. *Let R be a persistent semifir over an algebraically closed field. Then every matrix atom of R has a scalar eigenring.*

For we know that the eigenring of a matrix atom is a field, by Schur's lemma (Prop. 3.3.4), and the only algebraic skew field extension of k is k itself. ■

For a two-sided ideal α the eigenring is just the residue class ring modulo α ; thus we obtain

COROLLARY 6.6. *Let R be a persistent semifir over k and a a two-sided ideal of R , non-zero and finitely generated as left ideal. Then R/a is matrix algebraic over k .*

For let $u_1, \dots, u_r$ be a basis of a as free left R -module, then the column $u = (u_1, \dots, u_r)^T$ is regular, and by Th. 6.4 its eigenring is matrix algebraic over k . ■

For commutative Bezout domains that are persistent over k , this result holds even without assuming a to be finitely generated:

COROLLARY 6.7. *Let R be an invariant Bezout domain which is persistent over k . Then for any non-zero ideal a of R , R/a is algebraic over k .*

For if $a \neq 0$, then $a \supseteq Rc \neq 0$ and R/a is a homomorphic image of R/Rc , which is algebraic, by Th. 6.4. ■

Th. 6.4 can be applied to free algebras but, as already mentioned, there is a stronger result in this case. In the proof we shall need to compare modules over a ring with those over a retract. Let us briefly recall the facts we need.

Let R be a ring and K a subring which is a retract of R , say $\lambda: R \rightarrow K$ is a homomorphism which reduces to the identity on K . Then the functor

$$F: M \mapsto R \otimes_K M, \quad \text{where } M \text{ is any left } K\text{-module,}$$

provides a category-equivalence between all the left K -modules and all left R -modules of the form $M^F = R \otimes M$. For F has the inverse $G: P \mapsto K \otimes_R P$, where K is defined as right R -module by the rule $k \cdot r = k \cdot r \lambda$. It follows that for any left K -module M ,

$$\text{End}_R(R \otimes M) \cong \text{End}_K(M). \quad (6)$$

We can now prove

PROPOSITION 6.8. *Let $R = k\langle X \rangle$ be the free k -algebra on X . Then the eigenring of any finitely generated non-zero left ideal of R is finite-dimensional over k .*

Proof. Let a be the given left ideal and take a basis $u_1, \dots, u_r$ of a as left R -module. If we put all free generators of R not occurring in any u_i equal to 0, we obtain a retract K of R which is a free k -algebra of finite rank. Clearly $R/a \cong R \otimes M$, where $M = K/\sum Ku_i$, hence by (6) we may replace R by K , which is again free but of finite rank. Now $y \in K$ lies in $I(\sum Ku_i)$ precisely if

$$u_i y = \sum y_{ij} u_j, \quad \text{for some } y_{ij} \in K. \quad (7)$$

Let d be the maximum of the degrees of $u_1, \dots, u_r$ in a free generating set of K ; by (7) and the weak algorithm in K we have

$$y = \sum f_i u_i + y', \quad \deg y' < d. \quad (8)$$

Thus if F denotes the space of elements of degree $< d$ in K , then by (8) we have $I(a) \subseteq a + F$, hence

$$E(a) \cong I(a)/a \subseteq (a + F)/a \cong F/(F \cap a),$$

and this shows $E(a)$ to be finite-dimensional, because this is true of F . ■

For two-sided ideals this can be applied as in Cor. 6.6, to obtain the converse of a result in 2.5:

COROLLARY 6.9. *If R is a free k -algebra and a a two-sided ideal which is finitely generated as left ideal, then R/a is finite-dimensional over k . ■*

Prop. 6.8 shows in particular that the eigenring of any non-zero element in a free algebra is finite-dimensional; this will be proved for any regular matrices in 5.8.

As we saw in Ex. 1.8, the endomorphism ring of a distributive module need not be invariant in general, but in the special case of free algebras we can say more:

PROPOSITION 6.10. *Let R be a k -algebra which is an atomic 2-fir with a distributive factor lattice. If each atom of R has a scalar eigenring, then every non-zero element of R has a commutative eigenring.*

Proof. Let $M = R/Rc$; by induction on the length we may assume that every proper 1-torsion submodule of M has a commutative endomorphism ring. Every $\alpha \in E(Rc)$ maps each 1-torsion submodule of M into itself, so if M is the sum of its proper submodules, we can embed $E(Rc)$ into the direct product of the corresponding endomorphism rings, hence $E(Rc)$ is then commutative. The alternative is that the sum of all proper submodules of M is a unique maximal submodule, M' say. By hypothesis, every endomorphism of M/M' is induced by multiplication by an element of k , hence every endomorphism of M is of the form $\lambda + \alpha$, where $\lambda \in k$ and $M\alpha \subseteq M'$. It is therefore enough to show that any two non-surjective endomorphisms of M commute. Let us take such endomorphisms α, β of M ; by Prop. 1.10, one of $M\alpha, M\beta$ is contained in the other, say $M\alpha \subseteq M\beta \subseteq M'$.

Now α maps M into $M\beta \cong M/\ker \beta$, hence by Cor. 1.4(ii), α is induced by an endomorphism α' of $M\beta$, i.e. $\alpha = \beta\alpha'$. Let β_1 be the restriction of β to $M\beta$,

then we have

$$\beta\alpha = \beta\beta\alpha' = \beta\beta_1\alpha',$$

but on $M\beta$ all endomorphisms commute, by the induction hypothesis. Hence $\beta\beta_1\alpha' = \beta\alpha'\beta_1 = \alpha\beta_1 = \alpha\beta$, and so $\alpha\beta = \beta\alpha$. ■

COROLLARY 6.11. *Let R be a k -algebra which is an atomic 2-fir and remains one under arbitrary field extension. Then the eigenring of any non-zero element is commutative.*

For let k' be an algebraically closed field extension of cardinal greater than $\dim_k R$. Then all atoms in $R \otimes k'$ have scalar eigenrings (Cor. 3.3.6) and $R \otimes k'$ has DFL, by Th. 3.3, hence we can apply the result just proved (and the change of rings formula, Prop. 0.7.2) to reach the conclusion. ■

This corollary shows that in a free k -algebra all eigenrings of non-zero elements are commutative. Hence the result also holds for matrices that are stably associated to elements, but it does not extend to general matrices (cf. Ex. 8).

For examples of non-commutative eigenrings, consider the elements $x^2 + 1$ and $x^2 - 1$ in the complex-skew polynomial ring. Both are invariant (even central), hence their eigenrings are quotients of the whole ring by the ideals they generate. The element $x^2 + 1$ is an atom, so the quotient is a field (Prop. 3.3.4); this is easily seen to be the field of quaternions. The element $x^2 - 1$ is an I -atom, a product of two atoms, hence the eigenring is a 2×2 matrix ring over $\mathbb{C}$. For the algebra $\mathbb{Z}\langle x, y \rangle$ the eigenrings can be very different, as Ex. 3.3.16 shows.

Exercises 4.6

1. In an atomic 2-fir characterize the elements whose eigenring has zero radical.

2. (Cohn [69]) Show that in the free algebra $\mathbb{R}\langle x, y \rangle$ the element $a = xy^2x + xy + yx + x^2 + 1$ is an atom, but does not remain one under extension to $\mathbb{C}$. Deduce that the eigenring of a is $\mathbb{C}$. Find an element in the idealizer mapping to i .

3. (Roberts [82]) Let $k = \mathbb{F}_3(z)$, and in $k\langle x, y \rangle$ examine $c = x^2yxyx^2 + x^2yxxyx + x^2y^2x^2 + xyxyx^2 + zx^3 + x^2y + xyx + yx^2 + 1$. Verify that c is an atom but splits on adjoining a cube root of z . In $\mathbb{Q}\langle x, y \rangle$ consider the element obtained from c by replacing z by 2; show that it is an atom, but splits on adjoining a cube root of 2.

4⁰. (Ikeda [69]) In a free algebra of finite rank, is every ideal which is maximal as left ideal finitely generated as left ideal? (By Cor. 6.9 and Cor. 2.5.4 this is equivalent

to the question: is every skew field which is finitely generated as k -algebra necessarily finite-dimensional over k ?)

5*. (Bass) Let R be a commutative principal ideal domain containing a field k . Show that R is a persistent PID if and only if every prime ideal in $R[t]$ that is not minimal among the non-zero primes (i.e. of height > 1) meets $k[t]^*$. Deduce that the condition that R/p be algebraic over k for any non-zero prime ideal p is sufficient as well as necessary for $R \otimes k(t)$ to be a principal ideal domain.

6. Apply Ex. 5 to test whether R is a persistent PID over k in the following cases: (i) $R = k(x)[y]$, (ii) R is a PID whose residue class fields are algebraic over k and (iii) $R = k[[t]]$.

7. (Bergman) In the ring of integral quaternions show that the eigenring of each atom is commutative, but that this need no longer hold for general (non-zero) elements.

8. Let p be an atom in a free k -algebra. Show that the eigenring of $p \oplus p$ is not commutative.

Notes and comments on Chapter 4

Most of the results in this chapter are due to Bergman and the author, and were first published in the first edition. In particular, Th. 2.3 was obtained by the author in 1964 (unpublished) and he conjectured that it applied to free algebras. This conjecture was proved by Bergman in 1966. Much of the chapter is contained in Bergman [67], especially the later version, and 4.4 and 5 follow this source (and other unpublished work of Bergman in the 1960's) rather closely. The results of 4.4, of course, go back further, cf. e.g. Birkhoff [67], though our presentation follows Bergman [67]. The latter also contains Th. 5.2, while Th. 5.3 is taken from another unpublished manuscript of Bergman (ca. 1968).

The material of 4.1 went through several versions and was improved as a result of discussions with Bergman and Stephenson; in this edition we have followed Camillo [75] and Stephenson [74] in presenting properties of distributive modules. The characterization in Prop. 1.6 is based on a result by Roos [67]. Th. 2.4 generalizes (and simplifies) a result of Noether and Schmeidler [20].

A 2-fir such that two similar right commensurable elements are right associated was called 'uniform' in the first edition, but this term is used in other senses now and has therefore been discarded. A 2-fir R was called 'conservative' if R and $R \otimes k(t)$ are 2-firs and R is 1-inert in $R \otimes k(t)$; the place of this term has now been taken by the terms 'absolute' and 'persistent' as in the text. Thus Th. 3.3, the main result of 4.3, originally stated that a conservative 2-fir is uniform. This result survives in Th. 3.3(i), while (ii) is taken from Cohn [82']. Th. 6.4 was first proved for elements in commutative principal ideal domains by H. Bass (in a letter to the author in 1964). The first edition contained a version for elements of a persistent 2-fir, the present version is taken from Cohn [85]. Prop. 6.10 is due to Bergman (unpublished).

5. Modules over firs and semifirs

Just as firs form a natural generalization of principal ideal domains, so there is a class of modules over firs which generalizes the finitely generated modules over principal ideal domains. They are the positive modules studied in 5.3; they admit a decomposition into indecomposables, with a Krull-Schmidt theorem (in fact this holds more generally for finitely presented modules over firs), but it is no longer true that the indecomposables are cyclic. On the other hand, there is a dual class, the negative modules, and it is shown how the general finitely presented module is built up from free modules, positive and negative modules. A basic notion is that of a bound module; this and the duality, essentially the transpose, also used in the representation theory of algebras, are developed in 5.1 and 2 in the more general context of hereditary rings. In the special case of free algebras, the endomorphism rings of finitely presented bound modules are shown to be finite-dimensional over the ground field. This result, first proved by J. Lewin, is obtained here by means of a normal form for matrices over a free algebra due to M. L. Roberts, and his work is described in 5.8.

A second topic is the rank of matrices. Several notions of rank are defined, of which the most important, the inner rank, is studied more closely in 5.4. Over a semifir the inner rank obeys Sylvester's law of nullity. This leads to a natural generalization of semifirs: the Sylvester domains, first defined by W. Dicks and E. Sontag. They and some variants form the subject of 5.5 and 6.

In 5.7 we compare the different factorizations of a rectangular matrix over a semifir. Here the results are less complete, although in some ways parallel to the square case. There is an analysis of factorizations, which throws some light on the limitations to be expected.

The role of full matrices in the construction of fields of fractions (Ch. 7) is illuminated by the specialization lemma, which is proved in 5.9 and is applied there to give a simple proof of the existence of 'free fields'.

The remainder deals with various chain conditions (5.10) and the intersection theorem for ideals in firs, in 5.11.

5.1 Bound and unbound modules

Let R be any ring and $\mathcal{T}, \mathcal{F}$ two classes of left R -modules such that

- (i) $X \in \mathcal{T}$ if and only if $\text{Hom}_R(X, Y) = 0$ for all $Y \in \mathcal{F}$ and
- (ii) $Y \in \mathcal{F}$ if and only if $\text{Hom}_R(X, Y) = 0$ for all $X \in \mathcal{T}$.

If we view Hom as a bifunctor (i.e. a functor in two arguments) on the category of modules, (i) and (ii) express the fact that $\mathcal{T}$ and $\mathcal{F}$ are annihilators of each other, and we shall sometimes write $\mathcal{T} = {}^\perp \mathcal{F}$, $\mathcal{F} = \mathcal{T}^\perp$. There is a certain parallel here with the concept of orthogonality in a metric linear space, but by contrast to that case, Hom is *not* symmetric in its two arguments. Any $\mathcal{T}$ and $\mathcal{F}$ satisfying (i) and (ii) are called a *torsion class* and its associated *torsion free class* respectively, and the pair $(\mathcal{T}, \mathcal{F})$ is called a *torsion theory*. Given any class $\mathcal{C}$ of R -modules, we obtain a torsion theory by setting $\mathcal{F} = \mathcal{C}^\perp$, $\mathcal{T} = {}^\perp \mathcal{F}$; this is called the *torsion theory generated* by $\mathcal{C}$. Analogously the torsion theory *cogenerated* by $\mathcal{C}$ is formed by setting $\mathcal{T} = {}^\perp \mathcal{C}$, $\mathcal{F} = \mathcal{T}^\perp$. We shall be particularly interested in the torsion theory cogenerated by R .

Thus we define an R -module M to be *bound* if

$$M^* = \text{Hom}_R(M, R) = 0.$$

This means that there are no linear functionals on M apart from 0. The modules in the corresponding torsion free class are said to be *unbound*. An unbound module can also be defined as a module with no non-zero bound submodule. For if N satisfies this condition and M is any bound module, then so is any homomorphic image of M , hence $\text{Hom}_R(M, N) = 0$, and so N is unbound. Conversely, if N has a bound submodule $N' \neq 0$, then $\text{Hom}_R(N', N) \neq 0$.

In every torsion theory the classes $\mathcal{T}, \mathcal{F}$ admit certain operations; for the bound and unbound modules this is easily verified directly:

PROPOSITION 1.1. *Over any ring R , the class of bound modules is closed under the formation of homomorphic images, module extensions, direct limits and, in particular, direct sums. ■*

PROPOSITION 1.2. *Over any ring R , the class of unbound modules contains all free modules and is closed under the formation of submodules and arbitrary direct products (hence under inverse limits and direct sums), and under module extensions. ■*

Let M be any R -module; by Prop. 1.1, M has a unique maximal bound submodule M_b , viz. the sum of all bound submodules of M , and M/M_b has no non-zero bound submodules, i.e. is unbound. Dually, M_b may also be characterized as the least submodule of M with unbound quotient.

To give an example, over $\mathbb{Z}$ (or more generally, any principal ideal

domain), M_b is just the torsion submodule of M . Moreover, for a finitely generated $\mathbb{Z}$ -module M we have a decomposition

$$M = M_b \oplus F, \quad (1)$$

where F is free. Such a decomposition exists in fact over any semifir:

THEOREM 1.3. *Let R be a semifir. Then any finitely generated R -module M has a decomposition (1), where M_b is the maximal bound submodule and F is free; here M_b is unique and F is unique up to isomorphism. Moreover, M^* is free and $\text{rk } F = \text{rk } M^*$.*

Proof. Let us write $M = M_0 \oplus F$, where F is a free summand of maximal rank. This is possible because the rank of F is bounded by the number of generators of M . If M_0 is not bound, there is a non-zero homomorphism $f: M_0 \rightarrow R$, and since R is a semifir, $\text{im } f$ is free. Thus we can split off a free module from M_0 but this contradicts the maximality of $\text{rk } F$. Hence M_0 is bound and $M/M_0 \cong F$ is unbound, therefore $M_0 = M_b$ is the maximal bound submodule, and (1) is established. The uniqueness is clear, and dualizing (1), we find that $M^* \cong F^*$, hence M^* is free and $\text{rk } F = \text{rk } F^* = \text{rk } M^*$. ■

The unique submodule M_b in (1) is called the *bound component* of M .

It is clear that a corresponding result holds for n -generator modules over n -firs. In particular, this immediately leads to a condition for a module to be bound:

COROLLARY 1.4. *A module over a fir (or a finitely generated module over a semifir, or an n -generator module over an n -fir) R is bound if and only if it does not contain R as a direct summand. ■*

For the moment let us write $\mathcal{B}, \mathcal{U}$ for the classes of bound and unbound modules, respectively (over any ring R). Sometimes we wish to consider a wider class than $\mathcal{U}$ (and a corresponding narrower class than $\mathcal{B}$). Given $n \geq 1$, if every bound n -generator submodule of a module N is zero, N is said to be *n-unbound*, and the class of all such N is written $\mathcal{U}_n$. It is clear that

$$\mathcal{U}_1 \supseteq \mathcal{U}_2 \supseteq \dots \supseteq \mathcal{U}.$$

The class $\mathcal{B}_n = {}^\perp \mathcal{U}_n$ consists of all modules all of whose non-zero quotients have a non-zero bound n -generator submodule, and it is easily verified that $\mathcal{B}_n^\perp = \mathcal{U}_n$. Thus $(\mathcal{B}_n, \mathcal{U}_n)$ is a torsion theory and

$$\mathcal{B}_1 \subseteq \mathcal{B}_2 \subseteq \dots \subseteq \mathcal{B}.$$

It follows that Prop. 1.1 also holds for $\mathcal{B}_n$ and Prop. 1.2 for $\mathcal{U}_n$.

Over an integral domain the 1-unbound modules are just the torsion-free modules. This follows from the next result, which describes the n -unbound modules over n -firs:

THEOREM 1.5. *Let R be an n -fir. Then an R -module M is n -unbound if and only if every n -generator submodule of M is free.*

Proof. Let M be n -unbound, $n > 0$, and assume the result for integers less than n . Any submodule N of M generated by n elements is a homomorphic image of R^n , say $N = (R^n)f$. Let $g: N \rightarrow R$ be a non-zero homomorphism, then $fg: R^n \rightarrow R$ is non-zero and by Th. 1.1.1 (f), $\text{im}(fg)$ is free of some rank $r > 0$; thus $\text{im}(fg) = R^r$ and applying an appropriate automorphism to R^n we find maps

$$R^r \oplus R^{n-r} \rightarrow N \rightarrow R^r,$$

such that the first map is still surjective, while the composition is the projection onto R^r . Thus we obtain the decomposition $N = R^r \oplus (R^{n-r})f$; by induction $(R^{n-r})f$ is free and hence so is N . The converse is clear from the definition of n -unbound. ■

Of particular interest for our purpose is the fact that by Prop. 1.2 for $\mathcal{U}_n$, all direct powers R^I are n -unbound, so that, by Th. 1.5, when R is an n -fir, any n -generator submodule of R^I is free. This sharpens Cor. 1.4.6.

Bound modules satisfy chain conditions under fairly mild restrictions. To state them let us recall that a left hereditary ring is certainly weakly semihereditary, so we can apply Th. 0.3.7, which tells us that any projective (left or right) module over a left hereditary ring is a direct sum of finitely generated modules.

THEOREM 1.6. *Let R be a left hereditary ring, and let M be a finitely related left R -module. Then each submodule of M is a direct sum of a finitely generated module and a projective module.*

Proof. By hypothesis, $M = F/N$, where F is free and N is finitely generated. Every submodule of M has the form P/N , where $N \subseteq P \subseteq F$ and here P is projective, because R is left hereditary. By Th. 0.3.7, P is a direct sum of finitely generated modules and so it contains a finitely generated direct summand P' containing N (because N was finitely generated). Write $P = P' \oplus P''$, then we have the exact sequence

$$0 \rightarrow P'/N \rightarrow P/N \rightarrow P/P' \cong P'' \rightarrow 0.$$

Since P'' is projective, this sequence splits and P'/N as image of P' is finitely generated, while P'' is projective. ■

Since a bound module contains no non-zero projective submodule as a direct summand, we see that in the situation of Th. 1.6 every bound submodule of M is finitely generated. Thus we find

COROLLARY 1.7. *Any finitely related left module over a left hereditary ring satisfies ACC on bound submodules.* ■

Taking M itself to be bound, we obtain

COROLLARY 1.8. *If M is a finitely related bound left module over a left hereditary ring, then every bound submodule of M (in particular, M itself) is finitely presented.* ■

In fact, every finitely related module is a direct sum of a finitely presented module and a free module (as is easily seen), hence a finitely related bound module over any ring is finitely presented.

Exercises 5.1

1. Show that for any ring R and any right ideal a of R , R/a is bound if and only if the left annihilator of a is zero.
2. Show that the class of bound left R -modules admits submodules if and only if the injective hull of R (as left module) is unbound.
3. Let R be a left hereditary ring and M a finitely related left R -module. Show that the projective submodules of M form an inductive system.
4. Let R be an integral domain and M an R -module. Show that a homomorphism $M \rightarrow R$ annihilates all torsion elements; deduce that any module consisting entirely of torsion elements is bound. Show that for finitely generated right modules over left Ore domains the converse holds. (Hint: Use Prop. 0.9.2.)
5. Show that over any ring R a finitely related bound module is finitely presented. Deduce that a projective module is finitely generated provided that it has a finitely generated submodule with bound quotient. Hence obtain another proof of Th. 1.6 (without using Th. 0.3.7).
6. Find an example of an n -unbound module for all n , yet not unbound. (Hint: Over a semifir this requires a module M whose finitely generated submodules are free, but $M^* = 0$.)

7. Let K be a field and R a subring containing two right R -linearly independent elements a, b . Verify that the submodule $Ra^{-1} + Rb^{-1}$ of K is bound, as left R -module. Deduce that K is not semifree as left R -module (cf. Ex. 0.9.3).

8. Let R be an n -fir and $\mathcal{C}_n$ the class of all n -unbound modules with ACC_n . Show that $\mathcal{C}_n$ contains all free modules and is closed under (a) the formation of submodules, extensions and unions of ascending sequences with quotients in $\mathcal{C}_n$ (and hence direct sums) and (b) inverse limits (and hence direct products). Show that R must have left ACC_n for $\mathcal{C}_n$ to be non-trivial.

9. (Bergman [67]) (a) Let k be a field with an endomorphism α which is not surjective and define $R' = k[y; \alpha]$. Show that yR' is a two-sided ideal which is free as left ideal; further, show that if (u_i) is a left basis for k over k^α , then (yu_i) is a left basis for yR' over R' .

(b) Let $R = k^\alpha + yR'$ be the subring of polynomials in y with constant term in k^α . Verify that R is a right Ore domain with the same field K of fractions as R' , but R is not left Ore (note that even R' is not left Ore).

(c) Take $x \in k \setminus k^\alpha$ and let $f: R'^2 \rightarrow R'$ be the linear functional $(a, b) \mapsto axy - by$. Then f also defines a linear functional f_R (by restriction) on R^2 . Show that $\ker f_R = yR'(1, x) = \sum R'yu_i(1, x)$, and deduce that a submodule of an R -module of rank 1 can have infinite rank (for a suitable choice of k and α). Take a particular suffix $i = 1$ say and put $M = R^2/R'yu_1(1, x)$. Verify that f_R induces a map $f_M: M \rightarrow R$ and by showing that $\ker f_M$ is embeddable in R^2 , verify that M is unbound. Show also that $K \otimes M = K^2/K(1, x)$ and hence that M , although unbound, is not embeddable in $K \otimes M$.

10. Let R be a left Ore domain, M and N right R -modules and $f: N \rightarrow M$ a homomorphism. If M is finitely generated and $\text{coker } f$ is bound, verify that $\text{rk } N \leq \text{rk } M$. Does this remain true without the condition that M be finitely generated? (Hint: Try $M = K$.)

11. Show that for any finitely generated modules M, N over a semifir R , if $M \oplus R \cong N \oplus R$, then $M \cong N$.

12⁰. Describe the rings R such that every unbound module can be embedded in a direct power R^I .

5.2 Duality

We shall now establish a duality for bound modules over hereditary rings, from which it will follow that any finitely presented bound module satisfies DCC as well as ACC on bound submodules. We shall present the actual results so as to apply to arbitrary rings and then specialize at the end.

Let us call a module M over any ring R *special* if it is finitely presented and of projective dimension at most 1. Thus a special module M is a module with a presentation

$$0 \rightarrow Q \rightarrow P \rightarrow M \rightarrow 0,$$

where P, Q are finitely generated projective R -modules. Over a semifir the

special modules are just the finitely presented modules. If we dualize the above exact sequence, we get the exact sequence of right R -modules

$$0 \rightarrow M^* \rightarrow P^* \rightarrow Q^* \rightarrow \text{Ext}(M, R) \rightarrow 0,$$

bearing in mind that $\text{Ext}(P, R) = 0$, because P is projective. Let us assume that M is bound, so that $M^* = 0$. Dualizing once more and defining the transpose as $\text{Tr}(M) = \text{Ext}(M, R)$, we obtain the following commutative diagram with exact rows:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & Q & \longrightarrow & P & \longrightarrow & M & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ & & \alpha & & \beta & & \gamma & & \\ 0 & \longrightarrow & \text{Tr}(M)^* & \longrightarrow & Q^{**} & \longrightarrow & P^{**} & \longrightarrow & \text{Tr}^2(M) \longrightarrow 0 \end{array}$$

Here α and β are isomorphisms because P and Q are finitely generated projective. This allows us to define γ and to show that it is an isomorphism too (by the 5-lemma); likewise we conclude that $\text{Tr}(M)^* = 0$. These remarks suggest the truth of

PROPOSITION 2.1. *For any ring R , there is a duality Tr between the categories of special bound left R -modules and special bound right R -modules, such that if $M = \text{coker}(Q \rightarrow P)$, where P, Q are finitely generated projective and $Q \subseteq P$, then $\text{Tr}(M) = \text{coker}(P^* \rightarrow Q^*)$.*

Proof. Consider a mapping between two special bound left R -modules, given by presentations of the above form

$$\begin{array}{ccccccccc} 0 & \longrightarrow & Q & \xrightarrow{i} & P & \xrightarrow{j} & M & \longrightarrow & 0 \\ & & \downarrow h & & \downarrow g & & \downarrow f & & \\ 0 & \longrightarrow & Q' & \xrightarrow{i'} & P' & \xrightarrow{j'} & M' & \longrightarrow & 0 \end{array}$$

By the projectivity of P , jf lifts to a map $g: P \rightarrow P'$ and $h = g|_Q$ maps into Q' . The map f is completely determined by the pair (g, h) , for any such pair of maps induces a map of the cokernels. Moreover, two such pairs (g, h) and (g_1, h_1) give the same map f if and only if there exists $e: P \rightarrow Q'$ such that $g_1 - g = ei$, $h_1 - h = ie$ (thus (g, h) and (g_1, h_1) are 'homotopic'). Hence the category of special bound left R -modules is equivalent to the category of maps $Q \rightarrow P$, which are injective and have injective duals under homotopy-equivalence, as defined above. This category is clearly dual to the corresponding category of projective right R -modules, and hence to the category of special bound right R -modules. ■

We observe that the duality of Prop. 2.1 is given explicitly by the functor $\text{Ext}(-, R)$. In the special case $R = \mathbb{Z}$ it is just the familiar duality of abelian groups given by $A \mapsto \text{Hom}(A, \mathbb{Q}/\mathbb{Z})$. We indicate briefly the conditions under which this simplification can be made.

A module M is said to be *strongly bound* if M and all its submodules are bound. For any ring R , let E be the injective hull of R , as left R -module, and put $T = \text{coker}(R \rightarrow E)$, so that we have the exact sequence

$$0 \rightarrow R \rightarrow E \rightarrow T \rightarrow 0. \quad (1)$$

Clearly M is strongly bound if and only if $\text{Hom}(M, E) = 0$. By (1) we have, since E is injective, the exact sequence

$$0 \rightarrow \text{Hom}(M, R) \rightarrow \text{Hom}(M, E) \rightarrow \text{Hom}(M, T) \rightarrow \text{Ext}(M, R) \rightarrow 0. \quad (2)$$

However, when M is strongly bound, then $\text{Hom}(M, E) = 0$ and we have

PROPOSITION 2.2 *Let R be any ring and define E, T as above. Then for any strongly bound module M ,*

$$\text{Ext}(M, R) \cong \text{Hom}(M, T). \quad \blacksquare \quad (3)$$

Over a principal ideal domain (commutative or not) every bound module is strongly bound and we can therefore use Prop. 2.2 to express the duality in terms of Hom . Moreover, in this case E is just the field of fractions of R . This is an R -bimodule, hence T is also an R -bimodule and it follows that (for a left R -module M) (3) is actually a right R -module isomorphism. When R is a fir but not principal, there will generally be modules that are bound but not strongly bound. Even then we can use the exact sequence (2) to describe $\text{Ext}(M, R)$ as the cokernel of the mapping $\text{Hom}(M, E) \rightarrow \text{Hom}(M, T)$. However, we cannot in general expect to write $\text{Ext}(M, R)$ in the form $\text{Hom}(M, I)$ for some I , i.e. the functor $\text{Ext}(-, R)$ is not representable, because $\text{Ext}(-, R)$ is not left exact, unless we restrict the class of bound modules further.

Returning to Prop. 2.1, let us apply the result to hereditary rings. In the first place, every finitely presented module is now special. Moreover, by Cor. 1.7, every finitely presented bound module satisfies ACC on bound submodules (necessarily finitely presented by Cor. 1.8) and applying the above duality, we find that the module satisfies DCC for bound submodules. Thus we obtain

THEOREM 2.3. *A finitely presented bound module over a (left and right) hereditary ring satisfies both chain conditions for bound submodules. ■*

Exercises 5.2

1. What becomes of the duality of Prop. 2.1 in the case where R is left self-injective, i.e. injective as left R -module?

2. For this exercise only, let us call a module M *extra-special* if it has a presentation $0 \rightarrow P \rightarrow P \rightarrow M \rightarrow 0$, where P is finitely generated projective. Show that in the duality of Prop. 2.1, extra-special bound modules correspond to extra-special bound modules.

3. Let R be a left fir. If every bound left R -module is strongly bound, show that R is left principal.

4. Let R be a left Ore domain and K its field of fractions. Show that K is the injective hull of R . Does this remain true for more general rings that are embeddable in fields?

5*. Let R be a two-sided fir and E its injective hull as left R -module. Describe the R -bimodule structure of E and compare it with the injective hull of R as right R -module.

6. Let R be a left hereditary ring and M a finitely related left R -module. Show that if P is a maximal projective submodule of M , then M/P is strongly bound. Deduce that for every finitely related module M there is an exact sequence

$$0 \rightarrow P \rightarrow M \rightarrow Q \rightarrow 0,$$

where P is projective and Q is finitely generated and strongly bound.

7. Give an example of a strongly bound module over $k\langle x, y \rangle$ which is not finite-dimensional over k .

8. Let R be an integral domain and $\mathfrak{a}$ a left ideal. Show that $R/\mathfrak{a}$ is strongly bound if and only if R is an essential extension of $\mathfrak{a}$ (i.e. every non-zero left ideal of R meets $\mathfrak{a}$ non-trivially).

9. Let R be a two-sided fir and E its injective hull as left R -module. Show that for any finitely related left R -module M , there exists a set X such that $\text{Hom}(M, E) \cong E^X$. Let $R = k\langle x_1, x_2, \dots \rangle$, $M = R/Rx_1$; show that $P = \sum Rx_i/Rx_1$ is a maximal projective submodule of M . Find X in this case.

10*. Does Th. 2.3 hold for fully atomic semifirs? (Hint: See Prop. 10.6.)

11. (A. H. Schofield) Show that for bound left R -modules M, N over a hereditary ring R , $\text{Hom}_R(M, N) \cong \text{Tor}_1^R(\text{Tr } M, N)$. (Hint: Resolve M and note that $\text{Hom}_R(P, N) \cong P^* \otimes N$ for finitely generated projective P .)

5.3 Positive and negative modules over semifirs

We have already met torsion modules in 3.3, where we saw that over a principal ideal domain they reduce to finitely generated modules of torsion

elements, while many of the properties of the latter carry over to torsion modules over semifirs. In this section we apply the results of 5.1 and 2 to study the two classes of finitely presented modules over semifirs: the positive modules, which over principal ideal domains correspond to finitely generated modules, and their duals, the negative modules, which have no analogue in the classical case.

Throughout this section all modules occurring will be finitely presented modules over semifirs; in that case the characteristic of a module, as defined in 0.6, is an integer, and by Prop. 0.6.1, characteristics are additive on short exact sequences. Moreover, for a semifir R , the category of all finitely presented left R -modules is an abelian subcategory of ${}_R\mathcal{M}$, by Th. A.9.

Definition. Let R be a semifir and M an R -module.

(i) If M is finitely presented and $\chi(M') \geq 0$ for all submodules M' of M , then M is said to be *positive*.

(ii) If M is finitely presented and $\chi(M'') \leq 0$ for all quotients M'' of M , then M is said to be *negative*.

(iii) If either M is positive and $\chi(M') > 0$ for all non-zero submodules M' , or M is negative and $\chi(M'') < 0$ for all non-zero quotients M'' , then M is said to be *prime*.

We remark that a torsion module is just a module that is both positive and negative. Writing ${}_R\text{Pos}$, ${}_R\text{Neg}$, ${}_R\text{Tor}$ for the categories of positive, negative and torsion modules, we thus have

$${}_R\text{Tor} = {}_R\text{Pos} \cap {}_R\text{Neg}.$$

Over a principal ideal domain any finitely generated module is positive and there are no negative modules apart from torsion modules, because the characteristic does not assume negative values in this case (cf. 1.1). By contrast, a semifir which is not an Ore domain will always have modules of arbitrary negative characteristic, by Prop. 0.8.9.

Any submodule of characteristic 0 of a positive module is a torsion module, and the same holds for any quotient of characteristic 0 of a negative module. Therefore a prime module may also be defined as either a positive module with no non-zero torsion submodule or a negative module with no non-zero torsion quotient.

We have seen that torsion modules (over semifirs) are presented by full matrices, and we now examine the presenting matrices of positive, negative and prime modules. Let M be a left module presented by the $m \times n$ matrix C , which can be taken to be left regular; it is clear that M is bound if and only if C is also right regular. Let M' be a submodule with quotient $M'' = M/M'$,

then M', M'' may be represented by A, B , respectively, such that $C = AB$ (Prop. 0.6.1). Recalling the definitions of left (right) full and prime matrices from 3.3, we obtain

PROPOSITION 3.1. *Let R be a semifir and M a finitely presented left R -module, with presenting matrix C . Then C is left regular, and it is also right regular if and only if M is bound. Further, M is positive if and only if C is left full, and negative if and only if C is right full. Finally, M is prime if and only if C is (left or right) prime. ■*

These results suggest that there should be a duality between positive and negative modules, and this is in fact the case, provided that we restrict ourselves to bound modules. We note that in the decomposition of a module,

$$M = M_b \oplus R^n,$$

if M is positive or prime, then so is M_b , the bound component. Of course, a negative module is always bound, by definition, because R has no non-zero submodules of characteristic ≤ 0 .

THEOREM 3.2. *Let R be a semifir. Then $\text{Tr} = \text{Ext}(-, R)$ provides a duality between the category of bound positive left R -modules and the category of negative right R -modules.*

The proof is a straightforward application of Prop. 2.1. ■

In order to establish chain conditions we shall assume that R is a fir. Then we can apply Th. 2.3 to obtain

THEOREM 3.3. *Let R be a (two-sided) fir. Then any positive bound or negative module satisfies both chain conditions on bound submodules. ■*

Later, in 5.10, we shall meet other chain conditions valid over certain semifirs.

PROPOSITION 3.4. *Let R be a semifir. Then Pos, Neg and the class of bound modules all admit extensions and hence finite direct sums. Further, Pos admits submodules and Neg admits quotients within the category of all finitely presented R -modules.*

Proof. In the short exact sequence

$$0 \rightarrow M' \rightarrow M \rightarrow M'' \rightarrow 0, \quad (1)$$

assume that M', M'' are positive and let $N \subseteq M$. Then $\chi(N \cap M') \geq 0$ and

$$N/(N \cap M') \cong (N + M')/M' \subseteq M/M' \cong M'',$$

hence $\chi(N/N \cap M') \geq 0$, therefore $\chi(N) = \chi(N/N \cap M') + \chi(N \cap M') \geq 0$. This shows M to be positive. Now assume M', M'' to be bound and let $f: M \rightarrow R$ be a homomorphism; then $f|_{M'} = 0$, hence f is induced by a homomorphism $\tilde{f}: M'' \rightarrow R$, which must be 0, so $f = 0$ and M is bound. Thus Pos and the class of bound modules both admit extensions; by duality the same holds for Neg. The remaining assertions are clear from the definitions. ■

Let us consider a finitely presented bound module M over a fir R ; by Th. 2.3, any chain of bound submodules in M is finite. Moreover, any submodule of M is finitely related; it follows that any submodule of minimal characteristic is finitely generated and so has finite characteristic. Let M_1 be a submodule of minimal characteristic, $\chi(M_1) = h$ say. Then M_1 is negative, for any submodule N of M_1 satisfies $\chi(N) \geq h$, hence $\chi(M_1/N) \leq 0$. Let M^- be maximal among all submodules of characteristic h ; since these submodules are bound, M^- exists. We claim that it contains every negative submodule of M ; if N is negative, consider $N + M^-$. As a homomorphic image of $N \oplus M^-$ it is again negative, by Prop. 3.4, so $\chi((N + M^-)/M^-) \leq 0$, and hence $\chi(N + M^-) \leq h$. Here equality must hold, by the definition of h , and now $N + M^- = M^-$, by the definition of M^- . Thus M^- is the unique maximal negative submodule of M . Any submodule N of M satisfies $\chi(N) \geq \chi(M^-)$, so if $N \supseteq M^-$, then $\chi(N/M^-) \geq 0$, with strict inequality unless $N = M^-$; this shows N/M^- to be prime positive.

Dually we can find a least submodule M^+ with bound positive quotient M/M^+ , and M^+ is prime negative, therefore $M^+ \subseteq M^-$. Of course, M^-/M^+ is both positive and negative and hence is a torsion module. Thus we have

THEOREM 3.5. *Let M be a finitely presented bound module over a fir R . Then there is a chain*

$$0 \subseteq M^+ \subseteq M^- \subseteq M, \quad (2)$$

where M^- is the greatest negative submodule of M , and M^+ the least submodule with positive quotient M/M^+ . Moreover, M/M^- is prime positive, M^+ is prime negative and M^-/M^+ is a torsion module. ■

This result has an interpretation in terms of matrices which we shall meet in Prop. 4.7.

If we now impose the Ore condition, we have a principal ideal domain, and here the positive modules admit quotients as well as submodules. However, negative modules are absent; more precisely, they reduce to modules of

torsion elements, as do bound positive modules. To avoid confusion we shall use the term 'torsion module' only in the sense defined here (in 3.3) and refer to the classical notion (as defined, for example, in 0.9) as a module of torsion elements. Then in the Ore case we have

PROPOSITION 3.6. *The left torsion modules over a left Bezout domain are precisely the finitely presented modules consisting of torsion elements. Moreover, in an exact sequence (1), if M is a torsion module and M' is finitely generated, then M', M'' are torsion modules.*

Proof. Over a Bezout domain the characteristic is non-negative, by Cor. 0.3.4. Now a torsion module M is certainly finitely presented and $\chi(M) = 0$; if $x \in M$ is torsion free, then $\chi(Rx) = 1$, and so $\chi(M/Rx) = -1$, which is a contradiction, so all elements of M are torsion. Conversely, if M is finitely presented and consists of torsion elements, its rank is 0, and so is its characteristic. This remark also shows that $\chi(M') = \chi(M'') = 0$ in (1), and when M' is finitely generated, then both M' and M'' are finitely presented. ■

COROLLARY 3.7. *Over a left or right Bezout domain, every n -generator torsion module has a chain of torsion submodules of length n , whose quotients are cyclic torsion modules.*

For left Bezout domains this is immediate, by taking a cyclic submodule and using induction; for right Bezout domains it follows by duality. ■

Over principal ideal domains a more precise decomposition can be obtained, as is well known, see 8.1.

So far we have confined ourselves to finitely generated modules, as that is the most interesting case (for us). However, it is also possible to extend the notions defined here. Let us briefly mention the result for torsion modules; the extension to positive and negative modules is entirely similar.

For any semifir R we define the category ${}_R\text{Tor}^\dagger$ of *general torsion modules* as consisting of those modules M in which every finite subset is contained in a finitely generated torsion submodule. Then ${}_R\text{Tor}^\dagger$ (as full subcategory of ${}_R\mathcal{M}$) is again an abelian category; moreover it has exact direct limits and a generator, i.e. it is a *Grothendieck category* (cf. Cohn [70] for proofs) and may be obtained as the completion of ${}_R\text{Tor}$. Dually one defines the category of *protorsion modules* ${}_R\text{Tor}^\dagger$ to consist of all inverse limits of finitely generated torsion modules and all continuous homomorphisms (relative to the natural topology on the inverse limit). Now the functor $\text{Tr} = \text{Ext}(-, R)$ establishes a duality between the categories ${}_R\text{Tor}^\dagger$ and $\text{Tor}_R^\dagger$ (cf. Cohn [70]).

In 3.6 we saw that the Krull-Schmidt theorem applies to torsion modules.

This amounts to considering factorizations of (square) full matrices. When we come to consider rectangular matrices, we find that a similar result holds; we shall state it as a Krull-Schmidt theorem for finitely presented modules (Th. 3.9). Our first task is to prove a form of Fitting's lemma.

LEMMA 3.8. *Let M be a finitely presented bound indecomposable module over a fir R . Then $\text{End}_R(M)$ is a completely primary ring.*

Proof. We have to show that every endomorphism of M is either nilpotent or invertible. So let α be an endomorphism of M which is not an automorphism; then $M\alpha$ is a bound submodule of M and we have the descending chain

$$M \supseteq M\alpha \supseteq M\alpha^2 \supseteq \dots;$$

by Th. 2.3 this chain becomes stationary, say $M\alpha^n = M\alpha^{n+1} = \dots$. On $M\alpha^n$, α^n is a surjective endomorphism for any $r \geq 0$, and we have the exact sequence

$$0 \longrightarrow \ker \alpha^r \cap M\alpha^n \longrightarrow M\alpha^n \xrightarrow{\alpha^r} M\alpha^n \longrightarrow 0. \quad (3)$$

Write $N_r = \ker \alpha^r \cap M\alpha^n$, then by (3), $\chi(N_r) = 0$ and clearly,

$$N_1 \subseteq N_2 \subseteq \dots \quad (4)$$

If we take the bound components only, we get an ascending chain, by Th. 1.3, which again becomes stationary, by Th. 2.3, say $(N_m)_b = (N_{m+1})_b = \dots$. Since $\chi(N_m) = 0$, we have

$$N_m \cong (N_m)_b \oplus R^k, \quad \text{where } k = -\chi((N_m)_b),$$

and in $M/(N_m)_b$ we have the ascending chain of modules $N_\mu/(N_m)_b$, each isomorphic to R^k . But $M/(N_m)_b$ satisfies ACC_k , by Th. 1.2.3, so this chain becomes stationary, say $N_p = N_{p+1} = \dots$. Now the rest of the proof follows along the usual lines: let $r = \max(n, p)$, then for any $x \in M$, $x\alpha^r \in M\alpha^{2r}$, say $x\alpha^r = y\alpha^{2r}$, hence $x = y\alpha^r + z$, where $z \in \ker \alpha^r$, so we have

$$M = M\alpha^r + \ker \alpha^r. \quad (5)$$

If $x \in M\alpha^r \cap \ker \alpha^r$, then $x = y\alpha^r$ and $y\alpha^{2r} = x\alpha^r = 0$, so $y \in \ker \alpha^{2r} = \ker \alpha^r$, hence $x = y\alpha^r = 0$. This shows that the sum (5) is direct, and by the indecomposability of M , either $M\alpha^r = 0$ and α is nilpotent, or $\ker \alpha^r = 0$, $M\alpha^r = M$ and α is an automorphism. ■

We recall that any module M is indecomposable if and only if $\text{End}_R(M)$ contains no idempotents $\neq 0, 1$. In particular, when R is an integral domain, then $\text{End}_R(R) \cong R$ contains no idempotents $\neq 0, 1$ and so R is indecomposable as left (or right) R -module.

THEOREM 3.9. *Let M be a finitely presented module over a fir R . Then there exists a decomposition*

$$M \cong M_1 \oplus \dots \oplus M_r \oplus R^k, \quad (6)$$

where each M_i is bound indecomposable, and R is indecomposable. Given a second such decomposition of M :

$$M \cong M'_1 \oplus \dots \oplus M'_s \oplus R^l,$$

we have $k = l$, $r = s$ and there is a permutation $i \mapsto i'$ of $1, \dots, r$ such that $M'_i \cong M_{i'}$.

Proof. We have $M = M_b \oplus R^k$, and here k is uniquely determined as the rank of M/M_b (or also of M^*) because R has invariant basis number. Thus we need only decompose M_b ; by Th. 2.3 there exists a complete decomposition, and since each component has a local endomorphism ring, by Lemma 3.8, the conclusion follows by Azumaya's form of the Krull-Schmidt theorem (cf. e.g. Cohn [77], p. 353). ■

We also note an analogue of Schur's lemma. Let us call an R -module M *minimal bound* if M is non-zero bound, but no proper non-zero submodule is bound; over a semifir this means by Th. 1.3 that every finitely generated proper submodule is free. Now we have

PROPOSITION 3.10. *Let R be a semifir and M a finitely presented minimal bound R -module. Then $\text{End}_R(M)$ is a field.*

Proof. Consider an endomorphism $\alpha: M \rightarrow M$; its image is again bound, hence it is 0 or M . Suppose that $\alpha \neq 0$, then $\text{im } \alpha = M$ and we have an exact sequence

$$0 \longrightarrow \ker \alpha \longrightarrow M \xrightarrow{\alpha} M \longrightarrow 0. \quad (7)$$

Now $\ker \alpha \neq M$, hence it is free, and by comparing characteristics in (7), of rank 0, therefore $\ker \alpha = 0$, and α is an automorphism. ■

We conclude this section with an application of the above results (due to Bergman [a]), namely the embedding of any fir in a field. This question will be taken up again later, in Ch. 7, in a more general context, but the proof given here is more direct.

We have seen in 3.3 that the endomorphism ring of a simple torsion module is a field. For prime modules this need not hold; for example, if M is

any prime module, then so is $M \oplus M$, but its endomorphism ring has nilpotent elements. However, we have the following result:

PROPOSITION 3.11. *Let R be a semifir and M, N any positive prime R -modules, such that $\chi(M) = 1$. Then any non-zero homomorphism $f: M \rightarrow N$ is injective.*

Proof. We have an exact sequence

$$0 \longrightarrow \ker f \longrightarrow M \xrightarrow{f} N \longrightarrow \operatorname{coker} f \longrightarrow 0.$$

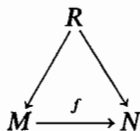
If $\ker f \neq 0$, then $\chi(\ker f) > 0$, hence $\chi(\operatorname{im} f) = 1 - \chi(\ker f) \leq 0$, therefore $\operatorname{im} f = 0$. ■

In particular, taking $N = M$, we obtain

COROLLARY 3.12. *Let R be a semifir and M a prime R -module of characteristic 1. Then $\operatorname{End}_R(M)$ is an integral domain.* ■

Here $\operatorname{End}_R(M)$ need not be a field, since, for example, R itself is prime of characteristic 1, but the endomorphism ring is a field for bound modules over a semifir with ACC_n for all n , under the conditions of Cor. 3.12, cf. Ex. 9.

Now let R be a fir and consider the category $\mathcal{L}$ of all left R -modules which are prime, of characteristic 1 and with a canonical embedding of R (as left R -module). The morphisms of $\mathcal{L}$ are the homomorphisms $f: M \rightarrow N$ such that the triangle shown commutes. We observe that between any two modules M, N of $\mathcal{L}$ there can be at most one morphism; for if f_1 is another, then $f_1 - f: M \rightarrow N$ is a homomorphism vanishing on R and hence zero, by Prop. 3.11. Thus $\mathcal{L}$ is a preordering.



LEMMA 3.13. *The category $\mathcal{L}$ is a directed preordering.*

Proof. Let $T \rightarrow M, T \rightarrow N$ be given and consider $M \oplus N$. This is again prime positive and it contains the submodule $T_1 = \{(x, -x) | x \in T\}$ which is isomorphic to T , hence of characteristic 1. Let S be a maximal submodule of characteristic 1 containing T_1 and write $P = (M \oplus N)/S$, then we see (as in the proof of Th. 3.5) that P has characteristic 1 and is prime; moreover it is the largest quotient of $M \oplus N$ with this property and therefore is the pushout

$$\begin{array}{ccc}
 T & \longrightarrow & M \\
 \downarrow & & \downarrow \\
 N & \longrightarrow & P
 \end{array}$$

(in the category $\mathcal{L}$). This shows $\mathcal{L}$ to be directed. ■

Let U be the direct limit of the category $\mathcal{L}$; then U is a left R -module containing R as submodule. Fix any non-zero element u of U ; there is a prime submodule of characteristic 1 containing u and the submodule generated by u is free: $Ru \cong R$. Now U can also be obtained as the direct limit of all the prime modules of characteristic 1 containing Ru ; this means that U has an automorphism mapping 1 to u . Since u was any non-zero element of U , we see that $E = \text{End}_R(U)$ is transitive on the non-zero elements of U and it contains R as subring. Moreover, if $f, g \in E$ and $1f = 1g$, then $f - g$ is not injective and hence vanishes, by Prop. 3.11. Thus each $f \in E$ is determined by its effect on $1 \in R$. Given $f \in E^\times$, there exists $g \in E$ with $1 \cdot fg = 1$, hence $fg = 1$, and this shows E to be a field. Thus we obtain

THEOREM 3.14. *Every fir can be embedded in a field.* ■

Here the finiteness conditions holding in a fir were needed for the proof. In Ch. 7 we shall see that every semifir R can be embedded in a field; moreover, there is a universal field of fractions for R , having any other field as specialization.

Exercises 5.3

1. Let R be a semifir and $n \geq 1$. How are the characteristics of modules related which correspond under the category equivalence of ${}_R\mathcal{M}$ and ${}_{R^*}\mathcal{M}$?
2. Show that over a semifir, every torsion submodule of an n -torsion module can be generated by n elements. Deduce that a torsion module over a fully atomic semifir satisfies ACC for torsion submodules.
3. Let M be an n -generator module over a semifir and M' any submodule. Show that M' can be generated by $n - \chi(M/M')$ elements. (Hint: Use the diagram of Prop. 0.6.1.)
4. Let R be a fir and $c \in R^\times$; a non-unit left factor b of c is called *inessential* if c has a non-unit left factor left coprime to b , otherwise it is *essential*. Show that the cyclic torsion modules defined by two elements c, c' have isomorphic injective hulls if and only if c and c' have an essential left factor (up to similarity) in common.

5. Let R be a fir and $c \in R^\times$. If the elements with c as essential left factor are pairwise right commensurable, show that the injective hull of R/Rc can be expressed as a direct limit of cyclic modules. Give an example in $R = k\langle x, y \rangle$ where this fails.

6. Let $R = k\langle x, y, z \rangle$ and consider the torsion modules $M = R/Rxz$, $N = R/Ryz$. Verify that both have R/Rz as quotient, and denote the pullback by T . Show that T is not a torsion module, but that it has a unique torsion submodule. Show that the lattice of torsion submodules of T is finite and isomorphic to that of $R/R(xy+1)xz$.

7. Let R be a semifir, $a, b \in R^\times$ and $A = R/Ra$, $B = R/Rb$. Show that $\text{Ext}^1(B, A) \cong \text{Tr}(B) \otimes A \cong R/(Ra + bR)$. Deduce that every extension of A by B splits if and only if $Ra + bR = R$.

If $R = k\langle x, y \rangle$, take $a = x$, $b = y$ and find the non-split extensions apart from the cyclic one.

8. Let R be a semifir, $a, c \in R^\times$, $A = R/Ra$, $C = R/cR$. Show that $\text{Tor}_1(C, A) \cong (cR \cap Ra)/cRa$.

9. Let R be a semifir with right ACC_r . Show that if M is a bound prime left R -module generated by n elements, such that $\chi(M) = 1$, then $\text{End}_R(M)$ is a field.

10⁰. (Bergman [67]) A module is said to be *polycyclic*, if it has a finite chain of submodules with cyclic factors. Does there exist a semifir, not left or right Bezout, over which every torsion module is polycyclic?

11⁰. (Bergman [67]) Determine the class of semifirs over which every polycyclic torsion module can be written as a direct sum of cyclic torsion modules (observe that this includes all principal ideal domains, cf. 8.1).

12. Let M, N be prime negative modules over a semifir R such that $\chi(N) = -1$. Show that any non-zero homomorphism from M to N is surjective.

13*. Let M, N be non-zero bound modules over a semifir R . If for all non-zero bound submodules $M' \subseteq M$, $N' \subseteq N$, we have $\chi(M') + \chi(N') > \chi(M)$, show that the kernel of any non-zero homomorphism from M to N is free. Deduce that if P is a finitely presented non-zero R -module such that for every finitely generated non-zero proper submodule P' of P , $2\chi(P') > \chi(P)$, then $\text{End}_R(P)$ is an integral domain. Hence obtain another proof of Cor. 3.12.

14. Let R be a semifir and ${}_R\mathcal{F}$ the full subcategory of ${}_R\mathcal{M}$ whose objects are the finitely presented modules. Use the characteristic to give a direct proof that ${}_R\mathcal{F}$ is an abelian subcategory of ${}_R\mathcal{M}$.

15. Let R be a semifir which is right Bezout. Show that every finitely generated bound left R -module is negative.

16. Let R be a semifir. Show that if every left torsion R -module has finite length, then R is a principal left ideal domain.

5.4 The ranks of matrices

The rank of a matrix is a numerical invariant which is certainly familiar to the reader, at least for matrices over a field or a commutative ring. We shall need an analogue for matrices over more general rings; there is then more than one invariant that can lay claim to generalize the usual rank. This is not surprising, since even in the commutative case one defines row and column rank separately and *then* proves them equal. We shall introduce three different notions of rank; the row rank, column rank and inner rank. Of these, the last will mainly be used, but the others are sometimes of use, and our first task is to describe their relationship to each other. For a smoother development we begin with some remarks on the decomposition of free modules over n -firs.

Let M be a module over any ring and N a submodule, then we define the *closure* $\bar{N}$ of N in M as the intersection of the kernels of all the linear functionals on M which vanish on N . If $\bar{N} = M$, we say that N is *dense* in M ; this is the case precisely when M/N is bound in the sense of 5.1. If $\bar{N} = N$, we say that N is *closed*; this is so if and only if N is the kernel of a mapping into a direct power of the ring. Hence if N is closed in M , then M/N is unbound.

If R is an n -fir, then the submodules of R^n which are annihilators of finite families of linear functionals are direct summands of R^n , by Th. 1.1.1(f). Moreover, any direct sum decomposition of R^n has the form $R^n \cong R^r \oplus R^s$, for a unique pair of integers r and s satisfying $r+s=n$. It follows (by an easy induction argument) that any descending chain of direct summands of R^n is finite, of length at most n . But any closed submodule of R^n is the intersection of a descending chain of direct summands, and so is itself a direct summand. This proves

PROPOSITION 4.1. *For any n -fir R , the closed submodules of R^n are precisely the direct summands. They form a lattice in which every maximal chain has length n , and the height of any member of this lattice is its rank as a free module. Moreover, if M is any closed submodule of R^n , then the linear functionals vanishing on M form a free module of rank $n - \text{rk } M$. ■*

We give two examples to show that the relation between the ranks of a module and its closure is not very close in general (for other examples see the exercises).

Let R be any integral domain; if $x_1, \dots, x_m$ are *left* linearly independent elements of R , then the elements $(x_1, 0, \dots, 0), \dots, (x_m, 0, \dots, 0)$ of R^n generate a free submodule of rank m , but its closure $(R, 0, \dots, 0)$ has rank 1. Secondly, let $y_1, \dots, y_n$ be *right* linearly independent; then it is clear that no non-zero

linear functional of R^n vanishes on $(y_1, \dots, y_n)$, so the submodule generated by this element is dense in R^n . Here we have a submodule of rank 1, whose closure has rank n .

Let us now turn to matrices. An $m \times n$ matrix $A = (a_{ij})$ over a ring R may be interpreted in different ways, either as

- (a) a right R -module homomorphism of columns ${}^nR \rightarrow {}^mR$,
- (b) a left R -module homomorphism of rows $R^m \rightarrow R^n$,
- (c) an element of the additive group ${}^mR \otimes R^n$ or
- (d) a bilinear mapping $R^m \times {}^nR \rightarrow R$.

Each interpretation leads to a notion of rank, of which some coincide when the ring is suitably specialized.

PROPOSITION 4.2. *Let R be an n -fir and $A \in {}^mR^n$. Then the following numbers are equal and do not exceed n :*

- (i) *the rank of the submodule of ${}^mR$ spanned by the n columns of A , i.e. $\text{rk im } A$ under interpretation (a),*
- (ii) *$n - \text{rk ker } A$ under interpretation (a) and*
- (iii) *the rank of the closed submodule generated by the rows of A under interpretation (b).*

Proof. The equivalence of (i) and (ii) is clear, as well as the fact that the common value cannot exceed n . To prove the equivalence of (i) and (iii), suppose that the closure of the image of the map $A : R^m \rightarrow R^n$ has rank r , the number described in (iii). Then we can factor A as follows:

$$\begin{array}{ccc} & R^r & \\ B \nearrow & & \nwarrow C \\ R^m & \xrightarrow{A} & R^n \end{array}$$

Here B has dense image (its range is the closure of its image), while C is the inclusion in a direct summand. Dualizing, we obtain maps represented by the same matrices and still factoring A :

$$\begin{array}{ccc} & {}^rR & \\ B \nwarrow & & \nearrow C \\ {}^mR & \xleftarrow{A} & {}^nR \end{array}$$

where C is the projection onto a summand and B will be injective. Hence the image of A in ${}^mR$ is isomorphic to ${}^rR$, where r is the number defined in (i). ■

The common value of these numbers is called the *column rank* of A . This definition makes it clear that the column rank of a matrix is independent of the choice of bases. In particular, the column rank is unchanged by elementary transformations. However, we have had to restrict the class of rings (to n -firs) to ensure that the image is a free module. Sometimes one defines the *column nullity* of A as $\text{rk ker } A$; unlike the column rank this is invariant under stable association.

Over an m -fir we can similarly define the *row rank* of A . Now case (i) of Prop. 4.2 (and its left-right dual) relates these numbers to the usual row and column ranks of matrices, over a commutative ring, say. In the general case, an $m \times n$ matrix over a $\max(m, n)$ -fir will have both kinds of ranks and in general they will be unrelated. This is clear from the examples given earlier, if we note that the row rank of A equals the rank of $R^m \cdot A$, while the column rank of A equals the closure thereof. Denoting the row and column ranks by ρ_r, ρ_c , respectively, we have for a product of matrices:

$$\rho_r(AB) \leq \rho_r(A), \quad \rho_c(AB) \leq \rho_c(B).$$

The reader may also formulate characterizations of these two types of rank in terms of interpretations (c) and (d) of the matrix A , but these are less simple. Various versions of the definitions of row and column rank may also be formulated for matrices over rings other than n -firs, but they are not in general equivalent. However, there is a third notion, which is left-right symmetric and is naturally defined over any ring:

PROPOSITION 4.3. *Let R be any ring and $A \in {}^mR^n$. Then the following four numbers are equal and do not exceed $\min(m, n)$:*

- (i) *the least r such that the map A (under interpretation (a) or (b)) can be factored through R^r ,*
- (ii) *the least r such that A can be written $\sum_1^r b_i \otimes c_i$, under interpretation (c),*
- (iii) *the least r such that the image of A in R^n is contained in a submodule generated by r elements (interpretation (b)) and*
- (iv) *the least r such that the image of A in ${}^mR$ is contained in a submodule generated by r elements (interpretation (a)).*

Further, this number does not exceed $\rho_r(A)$ in an m -fir, and does not exceed $\rho_c(A)$ in an n -fir.

The proof is straightforward. ■

The number described in this proposition is called the *inner rank* of A . It will be denoted by $\rho(A)$, and is most conveniently determined as the least r such that

$$A = BC, \quad \text{where } B \in {}^m R^r, C \in {}^r R^n. \quad (1)$$

Any factorization of A as in (1) with $r = \rho(A)$ will be called a *minimal factorization*.

In terms of minimal factorizations it is easy to verify the following inequalities:

$$\rho(A \oplus B) \leq \rho A + \rho B, \quad \rho(AB) \leq \min\{\rho A, \rho B\}, \quad \rho(A', A'') \geq \max\{\rho A', \rho A''\}. \quad (2)$$

We also note that for any $a \in R$, $\rho a = 0$ if and only if $a = 0$, for then (and only then) a may be written as a product of a 1×0 by a 0×1 matrix. Like other ranks, the inner rank does not depend on the choice of bases and is unaffected by elementary transformations. Over a skew field, it reduces to the usual rank. More generally we have

PROPOSITION 4.4. *Over a left Bezout domain the inner rank of any matrix equals its row rank.*

Proof. We know that in any case $\rho A \leq \rho_r A$; we have to prove equality here. Let $\rho A = r$; using interpretation (b), we can factor A by R^r . The image in R^r is a finitely generated submodule of R^r , hence by Prop. 0.9.3, of rank at most r , therefore the image in R^n also has rank at most r , i.e. $\rho_r A \leq r$. ■

In a two-sided Bezout domain, by symmetry, the row rank, column rank and inner rank of any matrix are all equal. Recalling the definition of a left full matrix from 3.3, we have

COROLLARY 4.5. *A matrix over a left Bezout domain is left regular if and only if it is left full.* ■

This result will be generalized in 5.5, where the precise class of rings satisfying the conclusion of this corollary will be determined, at least in the commutative case.

We note that over any ring, the full matrices are just the $n \times n$ matrices of inner rank n . Over a ring with UGN, every invertible matrix is full. For weakly finite rings we have the following more general result which is sometimes useful.

PROPOSITION 4.6. *Let R be a weakly finite ring and consider a partitioned matrix over R :*

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}, \quad \text{where } A_1 \in R_r.$$

If A_1 is invertible, then $\rho A \geq r$, with equality if and only if $A_3 A_1^{-1} A_2 = A_4$.

Proof. Since the inner rank is unchanged on passing to associates, we can make the transformations

$$A \rightarrow \begin{pmatrix} I & A_1^{-1} A_2 \\ A_3 & A_4 \end{pmatrix} \rightarrow \begin{pmatrix} I & A_1^{-1} A_2 \\ 0 & A_4 - A_3 A_1^{-1} A_2 \end{pmatrix} \rightarrow \begin{pmatrix} I & 0 \\ 0 & A_4 - A_3 A_1^{-1} A_2 \end{pmatrix}.$$

If $\rho A = s$, this matrix can be written as

$$PQ = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} (Q_1 \quad Q_2),$$

where P_1 is $r \times s$ and Q_1 is $s \times r$. Thus $I = P_1 Q_1$ and by weak finiteness, $r \leq s$. Therefore $\rho A \geq r$, and if equality holds, then $Q_1 P_1 = I$, but $P_1 Q_2 = 0$, so $Q_2 = 0$; similarly $P_2 = 0$ and so $A_4 - A_3 A_1^{-1} A_2 = P_2 Q_2 = 0$. The converse is clear. ■

We also note the following restatement of Th. 3.5 in terms of matrices:

PROPOSITION 4.7. *Let A be a regular matrix of inner rank r over a fir R . Then there is a factorization*

$$A = P A_0 Q,$$

where $A_0 \in R_r$ is full, P is right prime and Q is left prime. Moreover, if $A = P' Q'$ is any minimal factorization of A , then $P' = PU$, $Q' = VQ$, $A_0 = UV$, for some $U, V \in R_r$.

The proof is a straightforward translation, which may be left to the reader. ■

Sometimes it is possible to describe the inner rank in terms of full matrices alone. This will be useful for us later; we shall need the following lemma.

LEMMA 4.8. *Let A be a left full matrix over any ring R and suppose that A does not remain left full when the first column is omitted. Then there is a factorization*

$$A = B \begin{pmatrix} 1 & 0 \\ 0 & C \end{pmatrix},$$

where B is square and C has one less row and column than A . In this case B is full and C is left full.

Proof. Let A be $m \times n$ and write $A = (a, A')$, where a is $m \times 1$, A' is $m \times (n-1)$ and A' is not left full. Then we can write $A' = DC$, where D is $m \times (m-1)$ and C is $(m-1) \times (n-1)$, and now

$$A = (a \quad A') = (a \quad D) \begin{pmatrix} 1 & 0 \\ 0 & C \end{pmatrix}$$

is the desired factorization. Clearly if B is not full or C is not left full, then A cannot be left full. ■

THEOREM 4.9. *Let R be a ring such that the set of all full matrices over R is closed under products (where defined) and diagonal sums. Then the inner rank of a matrix over R is the maximum of the orders of its full submatrices.*

Proof. We may take R to be non-zero, since otherwise there is nothing to prove. Let A be an $m \times n$ matrix of inner rank r , with a minimal factorization $A = BC$. We shall show by induction on $m+n$ that A has a full $r \times r$ submatrix (clearly the largest possible). Since the product of two full matrices is again full, by assumption, we need only show that B and C each have a full $r \times r$ submatrix, and by symmetry it suffices to consider C . Now C is left full; if it remains left full when the first column is omitted, then by induction the truncated matrix, hence C itself, has a full $r \times r$ submatrix. So we may assume that C does not remain left full when the first column is omitted. By Lemma 4.8 there is a factorization $C = D(1 \oplus E)$, where D is full and E is left full. By the induction hypothesis E has a full $(r-1) \times (r-1)$ submatrix, therefore $1 \oplus E$ has a full $r \times r$ submatrix, by assumption, and its product with D is the desired full $r \times r$ submatrix of C . ■

A ring homomorphism is said to be *honest* if it keeps full matrices full. With this definition we have

COROLLARY 4.10. *Every honest homomorphism of a ring into a field preserves the inner rank.*

For the hypothesis ensures that the class of full matrices over R is closed under products and diagonal sums; moreover, a non-full matrix necessarily maps to a non-full matrix, so the result follows by applying Th. 4.9. ■

We conclude with a remark relating the minimal number of generators of a projective module to the inner rank, which is sometimes useful.

PROPOSITION 4.11. *Let P be a finitely generated projective module (over any*

ring R) and let E be an idempotent matrix presenting P . Then the minimal number of generators of P is the inner rank of E .

Proof. Consider any factorization

$$E = AB, \quad \text{where } A \in {}^n R', B \in {}^r R^n. \quad (3)$$

Then $F = BABA$ is an idempotent matrix isomorphic to E , and the R -module defined by F is generated by r elements. Conversely, if P can be generated by r elements, then it can be represented by an $r \times r$ matrix F , thus $E = AFB$ and $\rho E \leq r$. Thus the minimal number of generators equals the least value of r in (3), i.e. ρE . ■

Exercises 5.4

1. Let R be any ring with a field of fractions K (i.e. a field containing R and generated by R as field). Define the rank $\text{rk } M$ of any right R -module M as the dimension over K of $M \otimes K$. If R is a left Ore domain with field of fractions K and M is a finitely generated right R -module, show that $\text{rk } M = \text{rk } M^*$. Show that $\text{rk } M$ is the maximum length of chains of closed submodules in M .

2. (Klein [72]) Let $n \geq 1$ and let R be a ring such that any chain of closed submodules of R^n has length at most n ; show that any nilpotent $n \times n$ matrix A over R satisfies $A^n = 0$. If R is an integral domain, prove the converse (cf. Prop. 1.5.5).

3. Let $R = \sum_1^n R x_i$; show that $(x_1, \dots, x_n)^T$ generates a direct summand of ${}^n R$ and hence a closed submodule which is free of rank 1.

4. (Bergman [67]) Let R be an integral domain and x, y right linearly independent elements of R . Show that each of $(x, y, 1)$, $(0, 0, 1)$ generate a closed submodule of R^3 , but their sum is dense in R^3 . Show that if R is a 3-fir, then the lattice of closed submodules of R^3 is not modular.

5. (Bergman [67]) Given rings $R \subseteq K$ and an R -module M , a submodule M' is said to be K -closed if it is the zero-set of a family of maps $M \rightarrow K$; thus ' R -closed' is the same as 'closed'. Show that every closed submodule is K -closed. If R is a left Ore domain, K its field of fractions and M a finitely generated right R -module, show that every K -closed submodule is closed. If R is a right Ore domain, K its field of fractions and M a right R -module, then there is a natural bijection between the set of K -closed submodules of M and submodules of $M \otimes K$. Deduce that if R is a left and right Ore domain with field of fractions K and M a finitely generated right R -module, then the lattice of closed submodules of M is isomorphic to $\text{Lat}_K(M \otimes K)$, and hence is modular.

6. Let R be a semifir and M a left R -module with generators $e_1, \dots, e_n$ and defining relations $\sum a_{vi} e_i = 0$, where $A = (a_{vi})$ is a given matrix. Show that M is bound if $\rho_c(A) = n$. When $R = k\langle x, y \rangle$, give an example of a bound module with positive characteristic.

7. Show that a ring R is Hermite if and only if it has IBN and for any $n \times n$ idempotent matrix E of inner rank r , $\rho(I-E) \leq n-r$. Show also that a non-zero ring R is Hermite if and only if, for any $A \in {}^nR^n$, $B \in {}^nR^r$ such that AB is invertible, and any $C \in {}^nR^m$, $AC = 0$ implies $\rho C \leq n-r$.

8. Find all rings over which every full matrix is invertible.

9. Show that over a semifir any matrix stably associated to a full matrix is again full.

10. Let R be an n -fir. If $\rho \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \leq n$, show that $\rho \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \geq \rho A + \rho B$. Show also that if (A, B, C) is any matrix over R such that $\rho(A, B) + \rho(A, C) \leq n$, then $\rho(A, B, C) \leq \rho(A, B) + \rho(A, C) - \rho(A)$.

11. Let R be a left Ore domain. Show that an $r \times s$ matrix over R , where $r > s$, annihilates a non-zero row. Define the notions of row and column rank in this case, and show that over a two-sided Ore domain, the row and column ranks of an $r \times s$ matrix are equal. (Their common value is called the *outer rank*.)

12. Show that a commutative integral domain for which the inner and outer ranks coincide has the property: $a|b_1b_2$ implies $a = a_1a_2$ and $a_i|b_i$. (Hint: Consider 2×2 matrices. Such rings are called *primal*.)

13. (McAdam and Rush [78]) Let R be a commutative integral domain with field of fractions K . Show that the mapping $R \rightarrow K$ preserves the inner rank of 2×2 matrices if and only if the inverse of every finitely generated fractional ideal of R is locally cyclic, or equivalently, R is primal.

14. Let $f: R \rightarrow S$ be a ring homomorphism and $f_n: R_n \rightarrow S_n$ the induced homomorphism of matrix rings. Show that if f preserves the inner rank, then so does f_n .

5.5 Sylvester domains

In 5.4 we saw that the various ranks defined there coincide for Bezout domains, and this raises the problem of determining the precise class of rings for which this holds. Even in the commutative case this includes more than just Bezout domains, since the ranks agree, for example, for the polynomial ring $k[x, y]$, as will follow from Th. 5.4. However, this fails for $k[x, y, z]$, as is shown by the matrix

$$\begin{pmatrix} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{pmatrix}, \quad (1)$$

which is full, yet a (left and right) zero-divisor.

In this section we shall introduce a class of rings, following Dicks and Sontag [78], which in the commutative case gives a complete answer to these questions. We begin by recalling part of Prop. 3.2.4, which we shall formulate as follows:

PROPOSITION 5.1. *If R is an n -fir and $P \in {}^rR^n$, $Q \in {}^nR^s$ are such that*

$$PQ = 0, \quad (2)$$

then

$$\rho P + \rho Q \leq n. \quad (3)$$

Proof. Let $P = P_0P_1$ and $Q = Q_1Q_0$ be minimal factorizations for P, Q , then P_0 is right full, hence right regular, similarly Q_0 is left regular, and so (2) implies the full relation $P_1Q_1 = 0$. Applying Prop. 3.2.4, we obtain (3). ■

A non-zero ring satisfying the conclusion of this proposition for all n is called a *Sylvester domain*. As an immediate consequence we obtain Sylvester's law of nullity for the inner rank:

COROLLARY 5.2. *Let R be a Sylvester domain and $A \in {}^rR^n$, $B \in {}^nR^s$. Then*

$$\rho A + \rho B \leq n + \rho(AB). \quad (4)$$

For if $AB = PQ$ is a minimal factorization, so that Q has $\rho(AB)$ rows, then

$$(A \ P) \begin{pmatrix} B \\ -Q \end{pmatrix} = 0,$$

hence by Prop. 5.1,

$$n + \rho(AB) \geq \rho(A \ P) + \rho \begin{pmatrix} B \\ -Q \end{pmatrix} \geq \rho A + \rho B. \quad \blacksquare$$

Taking $n = 1$ we see that there are no zero-divisors, so a Sylvester domain is indeed an integral domain. More generally, every left full matrix over a Sylvester domain is left regular. By Prop. 5.1, every semifir is a Sylvester domain, but as we shall see, the converse does not hold. We begin by deriving some consequences of the law of nullity.

LEMMA 5.3. *Let R be a Sylvester domain. Then*

(i) *for any matrices A, B over R ,*

$$\rho(A \oplus B) = \rho A + \rho B; \quad (5)$$

(ii) if A, B, C are any matrices over R with the same number of rows and if $\rho(A, B) = \rho(A, C) = \rho A$, then

$$\rho(A \quad B \quad C) = \rho A. \quad (6)$$

(iii) R has UGN.

Proof. By suitably partitioning a minimal factorization of $A \oplus B$, we have

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} P \\ P' \end{pmatrix} (Q \quad Q'),$$

where the number of columns of P is $\rho(A \oplus B)$. Since $PQ' = 0$, we have, by (3),

$$\rho(A \oplus B) \geq \rho P + \rho Q' \geq \rho(PQ) + \rho(P'Q') = \rho A + \rho B \geq \rho(A \oplus B),$$

and this proves (5). To establish (6) we partition minimal factorizations of (A, B) , (A, C) as $(A, B) = D(E, E')$, $(A, C) = F(G, G')$; by hypothesis, $A = DE = FG$ are also minimal factorizations of A . Thus the number of columns of (D, F) is $2\rho A$ and

$$(D \quad F) \begin{pmatrix} -E \\ G \end{pmatrix} = 0,$$

hence by (3),

$$2\rho A \geq \rho(D \quad F) + \rho \begin{pmatrix} -E \\ G \end{pmatrix} \geq \rho D + \rho G \geq \rho(DE) + \rho(FG) = \rho A + \rho A.$$

Since also $\rho(D, F) \geq \rho A$, $\rho(-E, G)^T \geq \rho A$, it follows that $\rho(D, F) = \rho A$, say $(D, F) = H(J, K)$, where the number of columns of H is $\rho(D, F) = \rho A$. Now we have

$$(A, B, C) = (DE, DE', FG') = (HJE, HJE', HKG') = H(JE, JE', KG').$$

This shows that $\rho(A, B, C) \leq \rho H \leq \rho A$, and (6) follows. Finally, to show that R has UGN we must show that I_n is full, for all n . Since $R \neq 0$, 1 has rank 1; by induction from (i), $\rho I_n = n$ and the result follows. ■

In order to study Sylvester domains, we shall for a moment consider a somewhat wider class: the rings whose full matrices are regular. We note that this leads naturally to a one-sided class, whereas the condition defining Sylvester domains is left-right symmetric. For any ring R , an R -module will be called *spacial* if for every n , any set of n R -linearly dependent elements lies in a submodule generated by $n-1$ elements. An easy induction shows that in a spacial module every finite subset is contained in a free submodule, i.e. every spacial module is semifree; in particular, a spacial module is always flat. We

shall define the *local rank* of any R -module M as the least integer n such that every finite subset of M lies in an n -generator submodule; of course M may very well not have a local rank. To give an example, $\mathbf{Q}$ as $\mathbf{Z}$ -module has local rank 1, while $\mathbf{R}$ has no local rank, and for a finitely generated module the local rank is just the minimal generating number.

THEOREM 5.4. *For any ring R the following are equivalent:*

- (a) *every full matrix is left regular,*
- (b) *every left full matrix is left regular,*
- (c) *every flat left R -module is spacial,*
- (d) *every free left R -module is spacial,*
- (e) *the right annihilator of any non-zero matrix $A \in {}^m R^n$ has local rank less than n ,*
- (f) *the right annihilator of any non-zero vector in R^n has local rank less than n .*

When these conditions hold, the kernel of any homomorphism between spacial right R -modules is again spacial.

Proof. (a) $\Rightarrow$ (b). Let A be a left full matrix, say A is $m \times n$ of inner rank m ; we shall prove by induction on n that A is left regular. If the matrix obtained by omitting the first column of A is left full, then A is left regular by the induction hypothesis. In the opposite case, there is a factorization $A = B(1 \oplus C)$, by Lemma 4.8, where B is full and so left regular, by (a), and C is left full and so left regular, by the induction hypothesis. Hence $1 \oplus C$ is left regular and so is the product A .

(b) $\Rightarrow$ (c). Let M be a flat left R -module, and take any finite linearly dependent subset of M , arranged as a column $X \in {}^n M$, say. By hypothesis there is a non-zero $A \in R^n$ such that $AX = 0$. Since M is flat, this comes from an R -relation, say $X = BY$, where $B \in {}^n R^m$, $Y \in {}^m M$ and $AB = 0$. Thus B is not left regular and by (b) cannot be left full, say $B = CD$, where $C \in {}^n R^{n-1}$, $D \in {}^{n-1} R^m$. Then $X = BY = C(DY)$, so the components of X lie in the $(n-1)$ -generator submodule of M generated by the entries of DY , and this shows M to be spacial.

(c) $\Rightarrow$ (d) is obvious, and (d) implies (b) by the following chain of statements, each of which implies the next. Every free left R -module is spacial. For all m, n , any set of m left R -linearly dependent elements of R^n lies in an $(m-1)$ -generator submodule. For all m, n , any $m \times n$ matrix which is not left regular is the product of an $m \times (m-1)$ by an $(m-1) \times n$ matrix. Every matrix which is not left regular is not left full. Every left full matrix is left regular.

(b) $\Rightarrow$ (e). Let A be a non-zero $m \times n$ matrix. Any finite set of columns in the

right annihilator of A can be written as a matrix B such that $AB = 0$. Now for any minimal factorization $B = CD$, D is left full and so left regular by (b), hence $AC = 0$. Again by (b), C is not left full, but by definition it is right full, therefore it has at most $n-1$ columns. Now the columns of $B = CD$ lie in the submodule (of the right annihilator of A) generated by the columns of C , which shows that the right annihilator of A has local rank at most $n-1$.

(e) $\Rightarrow$ (f) is clear, and (f) $\Rightarrow$ (a): Let A be an $n \times n$ matrix and suppose that $XA = 0$ for some non-zero $X \in R^n$. Then the columns of A lie in the right annihilator of X , which by (f) has local rank less than n . Thus the columns of A lie in an $(n-1)$ -generator submodule of ${}^nR$, so A is not full.

Assume now that (a)–(f) hold and that $\alpha: M \rightarrow N$ is a homomorphism between spacial right R -modules. Any finite linearly dependent subset of $\ker \alpha$ can be written as a row, say $X \in (\ker \alpha)^m$. We write $X = YA$, where $Y \in M^r$, $A \in {}^rR^m$ and A, Y are chosen so that r is minimal. Then A has inner rank r , i.e. it is left full and so left regular. Since M is spacial, we have $r < m$ and the elements of Y are linearly independent. Now $\alpha(Y) \in N^r$ and there is again a factorization $\alpha(Y) = ZB$, where the elements of Z are linearly independent. We have

$$0 = \alpha(X) = \alpha(YA) = \alpha(Y)A = ZBA;$$

since Z is linearly independent and A left regular, it follows that $B = 0$, so $\alpha(Y) = 0$. Hence $Y \in (\ker \alpha)^r$ and this shows $\ker \alpha$ to be spacial. ■

Thus, for the rings of this theorem, spacial, flat and semifree left modules are the same. We remark that the conclusions of Th. 5.4 are not left–right symmetric, i.e. there exist rings R such that R satisfies the conditions but R^o does not (cf. Dicks and Sontag [78]). And, of course, (a) shows that the conditions hold in any Sylvester domain; more precisely, we have

COROLLARY 5.5. *Every Sylvester domain has weak global dimension at most 2 and is projective-free.*

Proof. Over a Sylvester domain every full matrix is regular, so the conditions of Th. 5.4 and their left–right duals hold. It follows that every flat module is semifree, and the kernel of any homomorphism between flat modules is flat; hence $\text{w.gl.dim. } R \leq 2$. Moreover, any finitely generated projective module is flat, hence free; it has unique rank by Lemma 5.3, so R is projective-free. ■

COROLLARY 5.6. *Every projective-free ring which is right coherent and of weak global dimension at most 1 is a semifir.*

For in such a ring every finitely generated right ideal is finitely related flat, therefore projective, hence free and of unique rank. ■

For Ore domains there is a converse to Th. 5.4 and Cor. 5.5. In that case the class of rings described in Th. 5.4 coincides with the class of Sylvester domains.

THEOREM 5.7. *For any two-sided Ore domain R the following are equivalent:*

- (a) R is a Sylvester domain,
- (b) $\text{w.gl.dim. } R \leq 2$ and every flat R -module is semifree,
- (c) the right annihilator of any matrix is semifree,
- (d) the right annihilator of any row vector is semifree,
- (e) every full matrix is left regular,
- (f) if K is the field of fractions of R , the embedding $R \rightarrow K$ preserves inner rank.

(a⁰)–(f⁰) The left–right duals of (a)–(f).

Proof. For any ring R we clearly have (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d).

(d) $\Rightarrow$ (e) for any right Ore domain R . Let K be the field of fractions of R and let $0 \neq X \in R^n$. If the right annihilator N of X contains n right R -linearly independent elements, then they are right K -linearly independent, but this is impossible because $X \neq 0$. Since N is semifree, by (d), it has local rank at most $n-1$; now (e) follows by Th. 5.4.

(e) $\Rightarrow$ (f) for any left Ore domain R . Any full matrix over R is left regular over R and remains so over the field of fractions K of R , hence it becomes left invertible. By Cor. 4.10 the embedding $R \rightarrow K$ preserves the inner rank, so (f) holds.

(f) $\Rightarrow$ (a). Sylvester's law of nullity holds in K , therefore it holds in R , so R is a Sylvester domain. ■

In particular, this result gives an explicit description of commutative Sylvester domains, in (b). To illustrate the result, let R be any ring and M a left R -module, with a finite free resolution

$$\cdots \longrightarrow F_2 \xrightarrow{A_1} F_1 \xrightarrow{A_0} F_0 \longrightarrow M \xrightarrow{\quad} 0,$$

where F_i is free of rank n_i . If A_1 can be written as

$$A_1 = BC, \tag{7}$$

where B is right regular and C left regular, then the rows of C generate the left annihilator of A_0 and are left R -linearly independent, hence $\ker A_0$ is then

free and so $pd(M) \leq 2$. Thus for any finitely generated module (over a Noetherian ring, say) of homological dimension at least 3 we obtain a matrix A_1 which cannot be expressed as in (7). For example, if $R = k[x, y, z]$, where k is a commutative field, then there is a minimal resolution for k (Koszul resolution)

$$0 \longrightarrow R \xrightarrow{A_2} R^3 \xrightarrow{A_1} R^3 \xrightarrow{A_0} R \longrightarrow k \longrightarrow 0, \quad (8)$$

with corresponding matrices

$$A_2 = \begin{pmatrix} x & y & z \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{pmatrix}, \quad A_0 = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

If A_1 were not full, we could replace (8) by a shorter sequence, in contradiction with the fact that $pd_R k = 3$. Thus A_1 is full, but clearly it is a zero-divisor.

Sometimes we shall want to consider the nullity condition used to define Sylvester domains for infinite matrices; here the inner rank of a matrix with infinitely many rows or columns (or both) is again defined in terms of minimal factorizations. Thus we consider the following condition:

$$\text{if } A \in {}^r R^n, B \in {}^n R^l \text{ and } AB = 0, \text{ then } \rho A + \rho B \leq n. \quad (9)$$

Taking I finite, we see that any non-zero ring satisfying (9) is a Sylvester domain. Moreover, if $\rho B = t$, say, we have a minimal factorization $B = CD$, where C is $n \times t$ and D is $t \times l$. Applying (9) to the relation $AC \cdot D = 0$ and remembering that $\rho D = t$, we find that $AC = 0$. This shows R to be right coherent. Conversely, if R is a right coherent Sylvester domain and $AB = 0$ as in (9), then $B = CD$, where C is $n \times t$ (for some t) and $AC = 0$. Hence we have $\rho A + \rho B \leq \rho A + \rho C \leq n$, which shows that (9) holds in R . We thus obtain

PROPOSITION 5.8. *The class of non-zero rings satisfying (9) consists of all right coherent Sylvester domains. ■*

In order to describe coherent Sylvester domains we shall need the following characterization of rings of weak global dimension at most 2.

LEMMA 5.9. *A ring has weak global dimension at most 2 if and only if the dual of every finitely presented left module is flat.*

Proof. Let M be a finitely presented left R -module, say

$$F_1 \xrightarrow{\alpha} F_0 \longrightarrow M \longrightarrow 0, \quad (10)$$

where F_0, F_1 are free of finite rank; then we have the exact sequence

$$0 \longrightarrow M^* \longrightarrow F_0^* \xrightarrow{\alpha^*} F_1^* \quad (11)$$

for the dual M^* . If $\text{w.gl.dim. } R \leq 2$, then M^* is flat, as kernel of α^* . Conversely, if the condition of the lemma holds, then the right annihilator of every matrix is flat. To reach the conclusion it suffices to show that any homomorphism $F_1 \rightarrow F_0$ of free right R -modules has a flat kernel K . Since flatness is a local condition, it is enough to verify that any finite subset X of K lies in a flat submodule of K . Now X lies in a finitely generated free submodule F'_1 of F_1 and the image of F'_1 lies in a finitely generated free submodule F'_0 of F_0 . Thus the kernel of the map $F'_1 \rightarrow F'_0$ is the right annihilator of a matrix, hence flat, and as a submodule of K containing X it has the requisite properties. ■

Since any Sylvester domain has weak global dimension at most 2 (by Cor. 5.5), the conditions of the next result hold in any right coherent Sylvester domain.

PROPOSITION 5.10. *For any ring R the following are equivalent:*

- (a) *R is right coherent and has weak global dimension at most 2;*
- (b) *the dual of every finitely presented left R -module is finitely generated projective;*
- (c) *the dual of every right R -module is flat.*

Proof. (a) $\Leftrightarrow$ (b). R is right coherent if and only if the dual of every finitely presented left R -module is finitely presented (cf. Th. A.6, Appendix). Using Lemma 5.9, we see that (a) is equivalent to the dual of every finitely presented left R -module being finitely presented and flat, hence projective, i.e. (b).

(a) $\Rightarrow$ (c). Here we shall use the characterization of right coherent rings as rings for which the dual of every free right module is flat. For any right R -module M there is a presentation (10), where F_0, F_1 are free right R -modules. Dualizing gives M^* as the kernel of a homomorphism between flat modules (as in (11)), hence M^* is flat, because the weak global dimension is at most 2.

(c) $\Rightarrow$ (a) follows because coherence is clear from the characterization just used, while the condition on the weak global dimension follows from the left-right dual of Lemma 5.9. ■

By adding right coherence we get the following consequence of Th. 5.7.

THEOREM 5.11. *For any two-sided Ore domain the following are equivalent:*

- (a) *R is a right coherent Sylvester domain,*
- (b) *the dual of every right R -module is semifree,*
- (c) *R is right coherent, projective-free and of weak global dimension at most 2,*
- (d) *the dual of every finitely presented left R -module is free,*
- (e) *the right annihilator of every matrix is free,*
- (f) *the right annihilator of every row vector is free.*

Proof. (a) $\Rightarrow$ (b). When (a) holds, R has weak global dimension at most 2, so by Prop. 5.10 the dual of every right R -module is flat, and by Th. 5.7, semifree.

(b) $\Rightarrow$ (c) $\Rightarrow$ (d) is clear from Prop. 5.10, and (d) $\Rightarrow$ (e) $\Rightarrow$ (f) follow directly.

(f) $\Rightarrow$ (a). If the right annihilator of every row vector is free, then R is a Sylvester domain by Th. 5.7, and it is right coherent, because over a right Ore domain a free submodule of a free right module of finite rank again has finite rank. ■

Finally, we prove a result that provides a source of Sylvester domains that are not semifirs.

THEOREM 5.12. *Let A be a commutative principal ideal domain and X any set. Then $A\langle X \rangle$ is a Sylvester domain.*

Proof. Writing K for the field of fractions of A , we have a homomorphism $A\langle X \rangle \rightarrow K\langle X \rangle$ and it will be enough to prove that this map is inner rank preserving, for then the law of nullity will hold in $A\langle X \rangle$.

To prove that the inner rank is preserved it suffices to show that every matrix over $A\langle X \rangle$ is inert in $K\langle X \rangle$, and since $K = \bigcup A[c^{-1}]$, where c ranges over $A^\times$, we need only verify the inertia for $A[c^{-1}]\langle X \rangle$. Writing c as a product of atoms and using induction, we need to show that every matrix over $A\langle X \rangle$ is inert in $A[p^{-1}]\langle X \rangle$, where p is an atom in A . We put $S = A\langle X \rangle$; p is a central atom in S and $S/pS = k\langle X \rangle$, where $k = A/pA$ is a field. Hence $R = S/pS$ is a fir, and it is a GE-ring, therefore we can apply Lemma 4.6.3 to conclude that every matrix of S is inert in $S[p^{-1}]$. ■

Exercises 5.5

1. Show that every Sylvester domain is an Hermite ring (and hence is weakly finite).
2. Let R be a ring; show that for any non-zero matrix $A \in {}^m R^n$ the right annihilator is free of rank $< n$ and the left annihilator free of rank $< m$ if and only if R is left and right coherent and every full matrix is regular.

3. (Dicks and Sontag [78]) Show that every Sylvester domain of weak global dimension at most one is a semifir.

4. Let R be a commutative local domain with maximal ideal $\mathfrak{m}$ such that the ideals $c\mathfrak{m}$ ($c \in R$) form a chain. If $\mathfrak{m}$ is locally cyclic, show that $\mathfrak{m}$ is idempotent, hence principal and deduce that R is a Sylvester domain.

5. Let R be a commutative domain and $A \in R_3$ of row rank 2. If $\rho A = 2$, show that A can be written as BC , where C is 2×3 left regular. Use this and the fact that the homological dimension of k over $k[x, y, z]$ is 3 to show that A_1 in (8) is full.

6. In any Sylvester domain, prove Frobenius' inequality: if AB, BC are defined, then

$$\rho(AB) + \rho(BC) \leq \rho(ABC) + \rho B.$$

7. Let R be a right coherent Sylvester domain and a a non-zero finitely generated right ideal. Show that $\chi(a) \geq 0$; deduce that any negative R -module is bound.

8°. Let R be a right coherent Sylvester domain. Which finitely generated right R -modules of positive characteristic and homological dimension 1 can be embedded as right ideals in R ? Which can be embedded as submodules in free R -modules?

9°. Let R be a Sylvester domain, C its centre and k the field of fractions of C . What can be said about $R \otimes_C k$? Is it always a semifir?

10°. Which commutative rings occur as centres of Sylvester domains?

5.6 Stable inner rank and pseudo-Sylvester domains

Sylvester domains share at least some of the good properties of semifirs, the main one being that they have a universal field of fractions, to be described in Ch. 7. It turns out that this property holds for an even wider class of rings, the pseudo-Sylvester domains. Moreover, these rings arise naturally in the study of localizations; we shall therefore briefly discuss them here, together with the notion of stable rank, used to define them.

Let R be any ring; given any matrix A over R , it is clear that $\rho(A \oplus I_s) \leq \rho A + s$, hence on writing $\rho_s A = \rho(A \oplus I_s) - s$, we have

$$\rho A = \rho_0 A \geq \rho_1 A \geq \dots \quad (1)$$

Let us assume that R has UGN; then every unit matrix is full, so that

$$\rho(A \oplus I_s) \geq \rho I_s = s,$$

hence $\rho_s A \geq 0$, so the sequence (1) is bounded below and therefore becomes stationary. The limit is reached after a finite number of steps; it is denoted by $\rho^* A$ and is called the *stable rank* of A . From the definition we have

$$0 \leq \rho^* A \leq \rho A, \quad (2)$$

and

$$\rho^*(A \oplus I_r) = \rho^*A + r, \quad \text{for all } r \geq 0, \quad (3)$$

while

$$\rho^*(A \oplus I_r) = \rho(A \oplus I_r), \quad \text{for sufficiently large } r. \quad (4)$$

Thus by taking a diagonal sum of A with a unit matrix of sufficiently high order we can stabilize the rank, i.e. make it equal to the inner rank. We shall call A stabilized if $\rho^*A = \rho A$.

If $A \in {}^mR^n$, we shall call A stably left (right) full if $\rho^*A = m$ (resp. n) and stably full if $\rho^*A = m = n$. We note that this is more than saying that A is full, it asserts that $A \oplus I_s$ is full for all $s \geq 0$.

In terms of the stable rank we can give criteria for a ring to be weakly finite or Hermite.

PROPOSITION 6.1. *For any ring R with UGN the following are equivalent:*

- (a) R is weakly finite,
- (b) every non-zero matrix over R has positive stable rank,
- (c) every non-zero idempotent matrix over R has positive stable rank.

Proof. Let A be any matrix over R , say $A \in {}^mR^n$ and suppose that $\rho^*A = \rho(A \oplus I_s) - s = r - s$, say, then we can write

$$\begin{pmatrix} A & 0 \\ 0 & I_s \end{pmatrix} = \begin{pmatrix} B \\ B' \end{pmatrix} (C \quad C'), \quad (5)$$

where $B \in {}^mR^r$, $B' \in {}^sR^r$, $C \in {}^rR^n$, $C' \in {}^rR^s$. Thus we have $B'C' = I$, $BC' = 0 = B'C$, $BC = A$.

To prove (a) $\Rightarrow$ (b), assume that $\rho^*A = 0$, then $r = s$, so by weak finiteness, $C'B' = I$, hence $B = 0 = C$ and therefore $A = BC = 0$.

(b) $\Rightarrow$ (c) is clear; to prove (c) $\Rightarrow$ (a), assume that R is not weakly finite, then there exist $B', C' \in R_s$ with $B'C' = I$, $C'B' \neq I$, so (5) holds with $A = B = C = I - C'B'$, and this is a non-zero idempotent matrix of zero stable rank. ■

PROPOSITION 6.2. *A ring R is Hermite if and only if R has UGN and the stable rank equals the inner rank.*

Proof. Assume that R is an Hermite ring and let A be a matrix for which (5) holds. Then $r \geq s$; if $r = s$, then $A = 0$ by the case just treated, so we may assume that $r > s$. Since $B'C' = I$, there exist by hypothesis $B'' \in {}^{r-s}R^r$, $C'' \in {}^rR^{r-s}$ such that $\begin{pmatrix} B' \\ B'' \end{pmatrix}$ and (C', C'') are mutually inverse. Hence

$$B = B(C' \quad C'') \begin{pmatrix} B' \\ B'' \end{pmatrix} = BC''B'',$$

because $BC' = 0$. It follows that $A = BC = BC'' \cdot B''C$, hence $\rho A \leq r - s \leq \rho^* A$, and so $\rho^* A = \rho A$, by (2). Conversely, if the stable rank and inner rank coincide, then by Prop. 6.1, R is weakly finite. Let $AB = I$, where $A \in {}^e R^n$, $B \in {}^n R^r$ and put $F = I_n - BA$. Then $F^2 = F$ and so

$$\begin{pmatrix} F & 0 \\ 0 & I_r \end{pmatrix} = \begin{pmatrix} F \\ A \end{pmatrix} (F \quad B),$$

hence $\rho^* F = n - r$, so $\rho F = n - r$, say $F = B'A'$, where $B' \in {}^n R^{n-r}$, $A' \in {}^{n-r} R^n$.

Now $(B, B') \begin{pmatrix} A \\ A' \end{pmatrix} = I$. By weak finiteness (Prop. 6.1) these matrices are mutually inverse, so R is Hermite. ■

Definition. (i) An *S-ring* is a ring with UGN in which the stable rank satisfies the law of nullity. (ii) A *pseudo-Sylvester domain* is a weakly finite S-ring.

A pseudo-Sylvester domain is, in fact, an integral domain, for if $ab = 0$, then $\rho^* a + \rho^* b \leq 1$, hence $\rho^* a = 0$ or $\rho^* b = 0$ and by Prop. 6.1, either $a = 0$ or $b = 0$. Further, every Sylvester domain is a pseudo-Sylvester domain, as the terminology suggests; more precisely, we have

PROPOSITION 6.3. *A ring is a Sylvester domain if and only if it is an S-ring and an Hermite ring.*

Proof. Any Sylvester domain is clearly an S-ring, and being projective-free (Cor. 5.5), it is Hermite. Conversely, assume that R is an S-ring and Hermite, then by Prop. 6.2 the stable rank agrees with the inner rank, so we have a Sylvester domain. ■

Since an Hermite ring is weakly finite, a Sylvester domain is certainly a pseudo-Sylvester domain, but examples will soon show that the latter form a wider class. We shall find that much of the theory of Sylvester domains has a parallel for pseudo-Sylvester domains. We first note

PROPOSITION 6.4. *Over a pseudo-Sylvester domain every finitely generated projective module is stably free.*

Proof. Let P be a finitely generated projective left R -module, defined by the idempotent matrix E , then $P \oplus R^m \oplus 0^n$ is given by the idempotent matrix

$F = E \oplus 0_m \oplus I_n$ and we have $I - F = (I - E) \oplus I_m \oplus 0_n$. By taking m, n large enough we can ensure that F and $I - F$ are stabilized. If F is $N \times N$, $\rho F = r$, $\rho(I - F) = s$, then $r + s \leq N$. Let $F = A_1 B_1$, $I - F = A_2 B_2$ be minimal factorizations; then

$$I = (A_1 \quad A_2) \begin{pmatrix} B_1 \\ B_2 \end{pmatrix},$$

and weak finiteness shows that $N \leq r + s$, and $A = (A_1, A_2)$, $B = (B_1, B_2)^T$ are mutually inverse; therefore $B_1 A_1 = I$ and $P \oplus R^m$ is free. ■

In order to describe pseudo-Sylvester domains more closely we need an analogue of Th. 4.9.

LEMMA 6.5. *Let R be a ring such that the set of stably full matrices over R is closed under products (where defined) and diagonal sums, and 1 is stably full. Then the stable rank of a matrix over R is the maximum of the orders of its stably full submatrices.*

Proof. Let $A \in {}^m R^n$, $\rho^* A = r$, and choose s such that $A \oplus I_s = BC$, where $B \in {}^{m+s} R^{r+s}$, $C \in {}^{r+s} R^{n+s}$. We shall prove by induction on $m+n$ that $A \oplus I_s$ has a stably full submatrix of order $r+s$ including I_s as a submatrix. By omitting the rows and columns of I_s we then obtain a stably full submatrix of A of order r .

Since the product of two stably full matrices is stably full, we need only find a stably full submatrix of order $r+s$ of B including the last s rows and a stably full submatrix of order $r+s$ of C including the last s columns; by symmetry it is enough to consider C . Further, by increasing s sufficiently we may assume that all submatrices of C including the last s columns are stabilized. Now C is stably left full; if it remains so when the first column is omitted, then by induction the truncated matrix, hence C itself, has a stably full submatrix of order $r+s$ including the last s columns. In the contrary case $C = (C_1, C')$, where C_1 is the first column and C' is not left full. Hence, by Lemma 4.8, we have $C = D(1 \oplus E)$, where D is full and E is left full, $(r+s-1) \times (n+s-1)$. Since C is stably left full, D must be stably full and E stably left full. Now by induction E has a stably full submatrix of order $r+s-1$ including the last s columns, hence $1 \oplus E$ has a stably full submatrix of order $r+s$ including the last s columns, while D is stably full. It follows that C has a stably full submatrix of order $r+s$ including the last s columns, which is what we had to find. ■

We also need an analogue of Lemma 5.3.

LEMMA 6.6. *Let R be an S -ring. Then*

(i) *for any matrices A, B over R ,*

$$\rho^*(A \oplus B) = \rho^*A + \rho^*B,$$

(ii) *if A, B, C have the same number of rows and $\rho^*(A, B) = \rho^*(A, C) = \rho^*A$, then*

$$\rho^*(A \quad B \quad C) = \rho^*A.$$

Proof. (i) Write $A' = A \oplus I_m, B' = B \oplus I_n$, where m, n are chosen so large that $A', B', A' \oplus B'$ are all stabilized. Now the proof of Lemma 5.3 shows that $\rho(A' \oplus B') = \rho A' + \rho B'$, hence $\rho^*(A \oplus B) = \rho^*A + \rho^*B$, as claimed. Similarly, in (ii) we replace A by $A' = A \oplus I_n$ with n chosen so large that $A', (A', B'), (A', C')$ are stabilized, where $B' = (B, 0)^T, C' = (C, 0)^T$. ■

By Prop. 6.4, over a pseudo-Sylvester domain every finitely generated projective module is stably free; further, it can be shown that any pseudo-Sylvester domain R satisfies $\text{w.gl.dim. } R \leq 2$, and for Ore domains this condition is sufficient as well as necessary for R to be a pseudo-Sylvester domain (cf. Cohn and Schofield [82]). Thus, for any field D the polynomial ring $D[x, y]$ in two central indeterminates is a pseudo-Sylvester domain, likewise the first Weyl algebra $A_1(k)$ over any commutative field k . Neither of these rings is a Sylvester domain, unless D is commutative.

Exercises 5.6

1. Verify that $A_1(k)$ is a pseudo-Sylvester domain but not a Sylvester domain.
2. Show that if for a ring R , ρ^* is not identically $-\infty$, then it is non-negative, and this is so precisely when R has UGN.
3. Use Lemma 6.6 to show that $\ker \rho^*$ is an ideal; by Prop. 6.1 deduce that for any ring R , $R/\ker \rho^*$ is the maximal weakly finite homomorphic image. Hence obtain another proof of Prop. 0.2.2.

5.7 The factorization of matrices over semifirs

In 3.3 we described the factorization of full matrices over semifirs and we saw there that with the appropriate finiteness assumptions matrix rings over semifirs behave very much like UFD's. By contrast the factorization of non-full or rectangular matrices is less well explored and no complete analogue of Th. 3.3.7 is known. We begin our study with a closer analysis of comaximal relations.

Given a comaximal relation over any ring R ,

$$C = AB' = BA', \quad \text{where } A \in {}^rR^m, B \in {}^rR^n, A' \in {}^nR^s, B' \in {}^mR^s, \quad (1)$$

we saw that this relation is full, hence by Prop. 3.2.4,

$$r + s \leq m + n. \quad (2)$$

For a proper comaximal relation equality holds in (2), so that $i(A) = i(A') = m - r$, $i(B) = i(B') = n - r$. For a general comaximal relation (1) we have

$$i(A') \leq i(A), i(B') \leq i(B), \quad i(A') + i(B') \leq i(C) \leq i(A) + i(B), \quad (3)$$

with equality in any one place (and hence in all) if and only if the relation is proper. The role of improper comaximal relations is made clear by

PROPOSITION 7.1. *Let R be an Hermite ring. Given a comaximal relation*

$$C = AB' = BA', \quad A \in {}^rR^m, B \in {}^rR^n, A' \in {}^nR^s, B' \in {}^mR^s, \quad (4)$$

if this is improper, then there exists $C'' \in {}^rR^t$, where $r + s + t = m + n$, such that

$$(C \ C'') = A(B' \ B'') = B(A' \ A''),$$

and the right-hand equation is a proper comaximal relation. Dually, we can add rows to C to obtain an $(r + t) \times s$ matrix with two factorizations forming a proper comaximal relation, or more generally, add t' rows and t'' columns ($t' + t'' = t$), to obtain an $(r + t') \times (s + t'')$ matrix.

Proof. By hypothesis (A, B) has a right inverse, say $AB_1 - BA_1 = I$, similarly there exist A_0, B_0 such that $A_0B' - B_0A' = I$, and if we put

$$P = \begin{pmatrix} A & B \\ A_0 & B_0 \end{pmatrix}, \quad Q = \begin{pmatrix} B_1 & B' \\ -A_1 & -A' \end{pmatrix},$$

then P is $(r + s) \times (m + n)$ and Q is $(m + n) \times (r + s)$ and

$$PQ = \begin{pmatrix} I & 0 \\ F & I \end{pmatrix}, \quad \text{where } F = A_0B_1 - B_0A_1.$$

Replacing A_0 by $A_0 - FA$ and B_0 by $B_0 - FB$, we find that $PQ = I$, hence by the Hermite property we can find A'_0, B'_0, A'', B'' such that

$$\begin{pmatrix} A & B \\ A_0 & B_0 \\ A'_0 & B'_0 \end{pmatrix}, \quad \begin{pmatrix} B_1 & B' & B'' \\ -A_1 & -A' & -A'' \end{pmatrix}$$

are mutually inverse, and now $A(B', B'') = B(A', A'')$ is a proper comaximal relation of the desired form. The rest follows by symmetry. ■

In a semifir, any relation may be analysed into a comaximal relation. In order to do this we need a normal form for matrices in terms of the column rank ρ_c . We observe that the result gives nothing new when the given matrix is right full, or more generally, right regular.

PROPOSITION 7.2. *Let R be a semifir and $A \in {}^tR^m$ a matrix of column rank t . Then $t \leq m$ and A may be written as*

$$A = PA', \quad \text{where } P \in {}^tR^t, A' \in {}^tR^m, \quad (5)$$

P is right regular and A' is right invertible, and this factorization is rigidly unique. Moreover, there is a proper comaximal relation

$$A'B'' = 0,$$

where B'' is left invertible, such that for any matrix B with m rows, $AB = 0$ if and only if

$$B = B'C, \quad \text{for some } C. \quad (6)$$

Proof. Given A as stated, we can find $T \in \text{GL}_m(R)$ such that AT has its first t columns right linearly independent and the rest 0. By suitably partitioning T , we have

$$A(T_1 \ T_2) = (P \ 0),$$

where $P \in {}^tR^t$ is right regular, and if $T^{-1} = (U, V)^T$, then $A = (P, 0)T^{-1} = PU$. Thus we obtain (5), and U is right invertible because $UT_1 = I$. If we also have $A = QW$, where W is right invertible and Q is right regular, denote by U', W' right inverses for U, W , then $Q = PUW' = PE$, say, similarly $P = QF$, hence $P = QF = PEF$, so $EF = I$ and similarly $FE = I$, hence (5) is rigidly unique.

Since A' is right invertible, there exist A'', B', B'' such that $(A', A'')^T$ and (B', B'') are square and mutually inverse. Hence $A'B'' = 0$ is a proper comaximal relation. Moreover, if $AB = 0$, then $PA'B = 0$, hence $A'B = 0$, therefore $B = (B', B'')(A', A'')^T B = B''A''B$, so that (6) holds with $C = A''B$. ■

We can now prove the basic lemma on the analysis of relations in a semifir.

LEMMA 7.3. *Let R be a semifir. Given a relation*

$$AB' = BA', \quad A \in {}^tR^m, B \in {}^tR^n, A' \in {}^nR^s, B' \in {}^mR^s, \quad (7)$$

let (A, B) have column rank t and $(B', A')^T$ row rank u . Then

$$t + u \leq m + n, \quad (8)$$

and there exists $P \in {}^tR^t$ such that P is right regular and $A_1 \in {}^tR^m$, $B_1 \in {}^tR^n$, such that

$$(A \ B) = P(A_1 \ B_1), \quad (9)$$

and there is a proper comaximal relation $A_1 B'_1 = B_1 A'_1$ such that $B' = B'_1 Q$, $A' = A'_1 Q$ for some Q .

The proof is an immediate consequence of Prop. 7.2, by writing (7) as $(A, B)(-B', A')^T = 0$, while (8) follows by the law of nullity. ■

We remark that A'_1, B'_1 depend only on A, B and not on A', B' .

In factorizing a rectangular matrix the role of an atom is taken over by an unfactorable matrix; in order to define this concept we have to look more closely at factorizations.

Let R be a semifir as before. A factorization of a matrix C :

$$C = AB \quad (10)$$

is said to be *proper* if A has no right inverse and B has no left inverse; for square matrices this agrees with the previous usage. By a trivial modification of a product AB we mean the replacement of A, B by $AU, U^{-1}B$, where U is an invertible matrix.

Given any factorization (10) of C , we can always ensure by a trivial modification that A becomes $(A_1, 0)$, where A_1 is right regular. If B becomes $(B_1, B_2)^T$, then $C = AB = A_1 B_1$, where now A_1 has fewer columns than A if A was a left zero-divisor. By repeating this operation on B we obtain a factorization (10) of C where A is right and B left regular. Moreover, the operations carried out on A , viz. multiplying on the right by an invertible matrix and omitting some columns, do not affect the property of having a right inverse; likewise for B , so a proper factorization remains proper. This shows the truth of

PROPOSITION 7.4. *Let R be a semifir and C any matrix over R . Then any factorization $C = AB$ can, by trivial modification and omitting some zero columns of A and corresponding rows of B , followed by trivial modification and omitting some zero rows of B and the corresponding columns of A , be brought to the form $C = A_1 B_1$, where A_1 is right and B_1 left regular. Moreover, A_1 and B_1 are non-units if and only if (10) was proper.*

The last remark follows because all changes preserve the property of being proper, and a right regular matrix is right invertible if and only if it is a unit. ■

Definition. A matrix C is said to be *unfactorable* if it is a regular non-unit and has no proper factorization.

It is clear that this property is preserved by stable association. We also note that since a one-sided unit is always a zero-divisor, an unfactorable matrix has neither a left nor a right inverse. Moreover, given any matrix C , we can always find a left invertible matrix P , a right invertible matrix Q and a regular matrix C_0 such that

$$C = PC_0Q,$$

by applying Prop. 7.2 and its left-right dual. It is easily seen that a finitely presented module is minimal bound if and only if its matrix is unfactorable.

We can now prove a refinement of the analysis of relations in Lemma 7.3, for use with unfactorable matrices.

LEMMA 7.5. *Let R be a semifir and C a regular matrix over R . Given two factorizations of C :*

$$C = AB' = BA', \quad (11)$$

where A is unfactorable and BA' is a proper factorization, either there is a matrix U such that $B = AU$, $B' = UA'$, or there is a proper comaximal relation

$$AB_1 = BA_1, \quad \text{and } A' = A_1Q, B' = B_1Q. \quad (12)$$

Proof. By Lemma 7.3 we have $(A, B) = P(A_0, B_0)$, $(B', A')^T = (B_1, A_1)^T Q$ and $A_0B_1 = B_0A_1$ is proper comaximal. Now either P is right invertible, then so is (A, B) (because (A_0, B_0) is right invertible) and we have a proper comaximal relation $AB_1 = BA_1$, or A_0 is left invertible. In the latter case, by a trivial modification we can take $A = P(I, 0)^T$, hence $P = (A, P')$. Writing $B_0 = (U, V)^T$, with a partition corresponding to that of P , we have $B = PB_0 = AU + P'V$, hence $AB' = BA' = (AU + P'V)A'$, i.e. $A(B' - UA') - P'VA' = 0$. But $P = (A, P')$ is right regular, hence $B' = UA'$, $VA' = 0$. Since A' is left regular, $V = 0$ and so $B = PB_0 = AU$ as claimed. ■

COROLLARY 7.6. *If in (11) both factorizations are proper and A, B are both unfactorable, then we may take U to be a unit, so that A, B are right associated.*

Thus either AB' and BA' can be obtained from each other by trivial modification, or we have a proper comaximal relation $AB_1 = BA_1$, where $A' = A_1Q$, $B' = B_1Q$.

For in this case, if there is no proper comaximal relation (12), we have $B = AU$, $A = BT$, hence $A = BT = AUT$ and A is right regular, hence $UT = I$, similarly $TU = I$ and so U is invertible. ■

We saw that a comaximal relation (1) is proper if and only if A and A' have the same index, and by Prop. 0.6.5, in any weakly finite ring two matrices A, A' are stably associated if and only if they occur in a proper comaximal relation (1). If the product AB' occurs in a factorization and we replace it by BA' , related to AB' by a proper comaximal relation (1), we shall call this change a *comaximal transposition*. It is clear that for full (square) matrices this reduces to the definition given in 3.5.

Let C be any regular matrix over a semifir R and consider a factorization

$$C = A_1 A_2 \dots A_r. \quad (13)$$

By trivial modification we may assume that $A_1, \dots, A_r$ are regular. If any A_i is a unit, we shall combine it with either of its neighbours; thus we may suppose that the A_i 's are regular non-units. Any such factorization of C corresponds to a chain of submodules with bound quotients and both chain conditions hold for such chains, by Th. 2.3; hence there is always a maximal refinement of such a chain. In terms of the factorization (13) this means that the A_i are unfactorable, since a matrix is unfactorable precisely if the module defined by it is minimal bound. Thus we have

THEOREM 7.7. *Over a fir, any regular matrix has a factorization into unfactorable matrices. More generally, any factorization into non-units can be refined to a factorization into unfactorable matrices. ■*

This proves the existence of complete factorizations; in order to compare two factorizations we apply Lemma 7.3; here another condition is needed, which unfortunately does not always hold:

LEMMA 7.8. *Let R be a semifir and C a regular matrix over R which has two factorizations*

$$C = AB' = BA', \quad (14)$$

such that

$$i(A) + i(B) \leq i(C). \quad (15)$$

Then either A, B have a common left factor which is not right invertible or A', B' have a common right factor which is not left invertible or (14) is a comaximal transposition.

We recall that if (14) is comaximal, then $i(A) + i(B) \geq i(C)$, so in that case (15) is just the condition for (14) to be proper.

Proof. Suppose that no common left or right factor exists in (14). Let C be $r \times s$ and let (A, B) have column rank t . By Lemma 7.3 there exists $P \in {}^e R^t$, such that P is right regular and a common left factor of A and B . By hypothesis, $PQ = I$ for some Q , hence $P(QP - I) = 0$, but P is right regular and so is a unit and may therefore be absorbed in A, B . By Lemma 7.3 there is now a proper comaximal relation $AB_1 = BA_1$ and $A' = A_1Q, B' = B_1Q$ for some Q . Further, $i(Q) = i(AB') - i(AB_1) = i(C) - i(A) - i(B) \geq 0$, by (15). By hypothesis Q has a left inverse, so by weak finiteness it is square and hence invertible. It follows that (14) is, indeed, a comaximal transposition. ■

This result suggests that we can compare two factorizations of C provided that the left-hand factors are not too 'skew', i.e. do not both have a large index. As a typical example where a comparison is impossible (owing to the failure of (15)), consider the free k -algebra on eight generators a, b, c, d, p, q, r, s and write $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, Y = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$. Then by taking the first row of the comaximal relation $X \cdot (I + YX) = (XY + I) \cdot X$, we obtain the improper relation

$$(a \ b) \begin{pmatrix} 1 + pa + qc & pb + qd \\ ra + sc & 1 + rb + sd \end{pmatrix} = (1 + ap + br, aq + bs) \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Exercises 5.7

1. Let R be a weakly finite ring. If $AB' = BA'$ is a proper comaximal relation, show that AB' is a least common right multiple of A and B .
2. Let R be the $n \times n$ matrix ring over a $2n$ -fir and let a, b be elements of R with a full common right multiple. Show that $aR \cap bR$ and $aR + bR$ are both principal.
3. Show that an unfactorable $n \times n$ matrix over any ring R is an atom in R_n and that the converse holds when R is commutative.
4. Show that a matrix over a semifir is unfactorable if and only if the module defined by it is minimal bound. Deduce that every minimal bound module over a semifir is prime.

5. Let R be a semifir and $A \in {}^m R^n$ an unfactorable matrix; show that the submodule of R^n generated by the rows of A is a direct summand in every proper finitely generated submodule of R^n containing it. Does the converse hold?

6. Let A be an invertible matrix over a semifir and consider a block decomposition $A = (A', A'')^T$ with the corresponding decomposition $A^{-1} = (B', B'')$. If by omitting some rows from A' we obtain (A_1, A_2) and by omitting some columns from B'' we get $(-B_1, B_2)^T$ (correspondingly partitioned), then $A_1 B_1 = A_2 B_2$ is a comaximal relation, but not proper, unless no rows or columns were omitted.

7⁰. Develop a theory of unique factorization rings that allows for the factorization of zero-divisors, taking e.g. the factorization of matrices over a principal ideal domain as a model, cf. 8.1.

5.8 A normal form for matrices over a free algebra

Let k be a commutative field; in the polynomial ring $k[x]$ it is easy to write down a normal form for polynomials under association: each non-zero polynomial is associated to precisely one monic polynomial. In the free k -algebra of rank greater than 1 the polynomials have a more complicated form, but now it is more natural to allow matrices as well and ask for a normal form under stable association. In particular, this allows us to take the matrix to be linear in the variables. This is the process of 'linearization by enlargement', also called Higman's trick (cf. Higman [40]). As a typical case let us take the (m, n) -entry of an $m \times n$ matrix. If this has the form $f + ab$, then by taking the diagonal sum with 1 and applying elementary transformations, we obtain successively

$$\begin{pmatrix} f+ab & 0 \\ 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} f+ab & a \\ 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} f & a \\ -b & 1 \end{pmatrix},$$

where only the last two entries of the last two rows are shown. By repeated application this allows us to reduce any matrix to one in which the elements of X occur to at most the first degree. Such a matrix is called *linear*; taking $X = \{x_1, \dots, x_d\}$ for simplicity, we see that every linear matrix has the form

$$A = A_0 + \sum_1^d A_i x_i, \quad \text{where } A_0, A_i \in {}^m k^n. \quad (1)$$

The matrix A in (1) is called *right monic* if $A_1, \dots, A_d$ are right comaximal (i.e. $(A_1, \dots, A_d)$ has a right inverse); a *left monic* linear matrix is defined similarly and a *monic* matrix is one that is left and right monic.

We remark that if A is right monic linear, then the terms of highest degree in A are left regular. For if $A = A_0 + \sum A_i x_i$ and $\sum C A_i x_i = 0$, then $\sum \bar{C} A_i x_i = 0$, where $\bar{C}$ are the terms of highest degree in C . Equating left cofactors of x_i we obtain $\bar{C} A_i = 0$, so $\bar{C}(A_1, \dots, A_d) = 0$ and therefore $\bar{C} = 0$, since the A_i are

right comaximal. This shows $\bar{A} = \sum A_i x_i$ to be left regular; it follows that A itself is left regular, for if $CA = 0$, then $\bar{CA} = 0$, hence $\bar{C} = 0$ and so $C = 0$.

We begin with a lemma which represents the weak algorithm for matrices.

LEMMA 8.1. *Let A be left monic linear and B right monic linear over $R = k\langle X \rangle$. If P, Q are any matrices over R such that*

$$PB = AQ, \quad (2)$$

then there exists a matrix C over R such that

$$P = AC + P_0, \quad Q = CB + Q_0, \quad (3)$$

where P_0, Q_0 are matrices over k .

Proof. By the remark preceding the lemma, (2) shows that P, Q have the same degree. Let us write $A = A_0 + \sum A_i x_i$, $B = B_0 + \sum B_i x_i$ and again denote the matrix of highest terms in any matrix T by $\bar{T}$. On equating the highest terms in (2), we obtain

$$\sum \bar{P} B_i x_i = \sum A_i x_i \bar{Q}. \quad (4)$$

If $\deg Q = 0$, then $\deg P = 0$ by (4), because B is right monic, so there is then nothing to prove. We may therefore assume that $\deg Q \geq 1$ and write $\bar{Q} = \sum Q_i x_i$, where Q_i is over R . Equating left cofactors of x_i in (4), we find $\bar{P} B_i = \bar{A} Q_i$, hence

$$\bar{P}(B_1, \dots, B_d) = \bar{A}(Q_1, \dots, Q_d).$$

Let D be a right inverse of $(B_1, \dots, B_d)$ and write $C = (Q_1, \dots, Q_d)D$, then D is over k and we have

$$\bar{P} = \bar{P}(B_1, \dots, B_d)D = \bar{A}C, \quad (5)$$

hence by (2),

$$(P - AC)B = A(Q - CB).$$

This is an equation of the same form as (2), but $P - AC$ has lower degree than P , by (5), and $Q - CB$ has lower degree than Q . Now the result follows by induction on the degree of Q . ■

We can now establish the existence of a normal form for matrices over $k\langle X \rangle$.

THEOREM 8.2. *Let $R = k\langle X \rangle$ be a free k -algebra. Any matrix over R is stably associated to a matrix $A \oplus 0$, where A is monic linear. Moreover, if $A \oplus 0$,*

$A' \oplus 0$ are two matrices that are stably associated and such that A, A' are monic linear, then the number of rows of A, A' agree, as do those of $A \oplus 0, A' \oplus 0$, likewise for columns and there exist invertible matrices P, Q over k such that $PA' = AQ$.

The uniqueness should be compared with the assertion of Prop. 0.7.5.

Proof. By the linearization process described earlier we reach the form $A \oplus 0$, where A has the linear form (1). Suppose that $m+n$ is minimal; we claim that then A is monic. For if $(A_1, \dots, A_d)$ has no right inverse say, then this $m \times nd$ matrix over k has rank less than m and so by elementary row operations the last row can be reduced to 0. If now the last row of A_0 is also 0, we can reduce m by 1, contradicting the minimality (this amounts to writing A as $B \oplus N$, where N is the 1×0 zero matrix). Hence the last row of A_0 is not 0, and by further row operations over R and column operations over k we find that A is associated to $A' \oplus 1$, where A' is again linear. So A is stably associated to A' , but this again contradicts the minimality of $m+n$. Hence A is right monic; by symmetry it is also left monic. Thus the existence of the form (1) is established. We remark that if A is invertible, it is stably associated to the 0×0 matrix.

Suppose now that A and A' are both monic linear and $A \oplus 0, A' \oplus 0$ are stably associated, then $F \oplus 0$ and $F' \oplus 0$ are associated, where $F = A \oplus I_r$, $F' = A' \oplus I_s$, for suitable r and s . Thus there exist invertible matrices U, V such that for appropriate partitioning

$$\begin{pmatrix} U_1 & U_2 \\ U_3 & U_4 \end{pmatrix} \begin{pmatrix} F & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} F' & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1 & V_2 \\ V_3 & V_4 \end{pmatrix}. \quad (6)$$

It follows that $U_3 F = 0 = F' V_2$. Now A, A' are regular, hence so are F, F' and therefore $U_3 = 0 = V_2$. Let F be $m \times n$, F' $m' \times n'$ and $F \oplus 0, F' \oplus 0$ both $t \times u$. Then U_3 is $t - m' \times m$, so by Prop. 3.2.3, $t - m' + m \leq t$, i.e. $m \leq m'$. Similarly V_2 is $n' \times u - n$, so $n' + u - n \leq u$, and hence $n' \leq n$. By symmetry we have $m' \leq m, n \leq n'$ and so F, F' are both $m \times n$. It follows that U_1, U_4, V_1, V_4 are all square and are invertible, since this is true of U, V . Moreover, $U_1 F = F' V_1$, so F, F' are associated and A, A' stably associated. We thus have a comaximal relation

$$AB' = BA'.$$

Hence there exist matrices C, D, C', D' over R such that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \text{ and } \begin{pmatrix} D' & -B' \\ -C' & A' \end{pmatrix} \quad (7)$$

are mutually inverse. Further, by Lemma 8.1 there exists P such that

$$B = AP + B_0, \quad B' = PA' + B'_0,$$

where B_0, B'_0 are over k . Hence on multiplying the matrices (7) by $\begin{pmatrix} I & -P \\ 0 & I \end{pmatrix}$ on the right and $\begin{pmatrix} I & P \\ 0 & I \end{pmatrix}$ on the left, respectively, we obtain the pair of mutually inverse matrices

$$\begin{pmatrix} A & B_0 \\ C & D_1 \end{pmatrix} \text{ and } \begin{pmatrix} D'_1 & -B'_0 \\ -C' & A' \end{pmatrix}.$$

We also have $A'C = C'A$, so by another application of Lemma 8.1 we obtain a pair of inverse matrices

$$\begin{pmatrix} A & B_0 \\ C_0 & D_2 \end{pmatrix} \text{ and } \begin{pmatrix} D'_2 & -B'_0 \\ -C'_0 & A' \end{pmatrix},$$

where C_0, C'_0 are over k . Thus we have the equation

$$AD'_2 - B_0C'_0 = I.$$

Equating highest terms we find $\sum A_i x_i \overline{D'_2} = 0$. Taking right cofactors of x_i we have $A_i \overline{D'_2} = 0$; and since $(A_1, \dots, A_d)^T$ is right regular, we conclude that $\overline{D'_2} = 0$, hence $D'_2 = 0$ and so $-B_0C'_0 = I$. Thus B_0 has a right inverse over k , and by symmetry B'_0 has a left inverse. Hence $i(B) \geq 0 \geq i(B')$; since A, A' are stably associated, they have the same index, hence so do B, B' and therefore $i(B_0) = i(B'_0) = 0$. It follows that B_0, B'_0 are invertible over k , and we have $AB'_0 = B_0A'$, as claimed. ■

The proof shows that once we have reached a linear form for our matrix, if the latter is left regular, we can achieve a right monic linear form by elementary row operations over R and column operations over k , and similarly on the other side. Thus we have

COROLLARY 8.3. *If A is a right regular linear matrix over $R = k\langle X \rangle$, then there exist invertible matrices P over k and U over R such that $PAU = B \oplus I$, where B is left monic linear. ■*

We can now give the promised extension of Prop. 4.6.8 by proving that the eigenring of a regular matrix over a free k -algebra is finite-dimensional over k . Equivalently, we may show that the endomorphism ring of a finitely presented bound module is finite-dimensional. We shall first prove a more general result:

THEOREM 8.4. *Let $R = k\langle X \rangle$ and let M, N be finitely presented R -modules of which M is bound. Then $\text{Hom}_R(M, N)$ is finite-dimensional over k .*

Proof. As in Prop. 4.6.8 we may take R to be of finite rank. Let M, N be left R -modules with defining matrices A, B , which may be taken to be monic linear by Th. 8.2, and A is regular (because M is bound), while B is left regular. A homomorphism $f: M \rightarrow N$ corresponds to a pair of matrices P, Q such that

$$AQ = PB, \quad (8)$$

and conversely, such a pair of matrices defines a homomorphism, while P, Q and P', Q' define the same homomorphism if and only if $P - P' = AC$, $Q - Q' = CB$ (cf. 0.7). By Lemma 8.1, if (8) holds, we can write $Q = CB + Q_0$, $P = AC + P_0$, where P_0, Q_0 are over k . Thus the given homomorphism may be represented by P_0, Q_0 . But the space of these matrices is a finite-dimensional k -space, hence $\text{Hom}_R(M, N)$ is finite-dimensional. ■

To obtain an estimate for the dimension, let $A \in {}^m R^n$, $B \in {}^r R^s$. As we saw in 0.7, we have

$$\text{Hom}_R(M, N) \cong \mathbf{I}(A, B)/\mathfrak{b}, \quad (9)$$

where $\mathbf{I}(A, B) = \{P \in {}^m R^s \mid AP = PB \text{ for some } P' \in {}^m R^r\}$ and $\mathfrak{b}$ is the left R -module spanned by the rows of B . Now Th. 8.4 shows that

$$\mathfrak{b} \subseteq \mathbf{I}(A, B) \subseteq \mathfrak{b} + {}^n k^s;$$

hence by the modular law,

$$\mathbf{I}(A, B) = \mathfrak{b} + (\mathbf{I}(A, B) \cap {}^n k^s).$$

Since B is left regular, it may be taken to be right monic; then its linear part is left regular and it follows that $\mathfrak{b} \cap {}^n k^s = 0$. We therefore have a k -space isomorphism

$$\mathbf{I}(A, B)/\mathfrak{b} \cong \mathbf{I}(A, B) \cap {}^n k^s. \quad (10)$$

This shows that the dimension of $\text{Hom}_R(M, N)$ is at most ns . We also remark that the condition that M be bound cannot be omitted, since, for example, $\text{Hom}_R(R, N) \cong N$, but a finitely presented module, even bound, need not be finite-dimensional over k ; e.g. if $R = k\langle x, y \rangle$, then R/Rx is bound but has the k -basis $1, y, xy, x^2y, \dots$

Taking $N = M$ in Th. 8.4 we obtain

COROLLARY 8.5. *Let M be a finitely presented bound module over $R = k\langle X \rangle$. Then $\text{End}_R(M)$ is finite-dimensional over k . ■*

As already noted, this generalizes Prop. 4.6.8.

The monic normal form can also be used to describe factorizations of full matrices over $k\langle X \rangle$. First we need a bound on the degree of the factors.

LEMMA 8.6. *Let $R = k\langle X \rangle$ be the free algebra on $X = \{x_1, \dots, x_d\}$ and let C be a full matrix over R . If*

$$C = AB$$

is a factorization of C into square matrices, then there is an invertible matrix U such that $d(AU) \leq d(C)$.

Here the degree of a matrix $C = (c_{ij})$ is defined as $d(C) = \max\{d(c_{ij})\}$.

Proof. Let us write $C = (c_{ij})$ and likewise for A, B ; it is clear that these three matrices are all of the same order, n say. We consider the free k -algebra S on $x_1, \dots, x_d, y_1, \dots, y_n$. There is an embedding of R in S which is honest, since R is a retract of S . We extend the degree on R to S by putting $d(y_i) = 1$.

Consider the right ideal $\mathfrak{a}$ of S generated by the $p_j = \sum y_i a_{ij}$; since A (like C) is full, it is regular and so $\mathfrak{a}$ is free on the p_j . Let $f_1, \dots, f_n$ be a right d -independent basis of $\mathfrak{a}$ such that $d(f_1) \leq \dots \leq d(f_n)$ (cf. 2.2). We have

$$p_j = \sum y_i a_{ij}, \quad f_j = \sum y_i d_{ij},$$

where A and $D = (d_{ij})$ are right associated, because the p_j and f_j are two bases of $\mathfrak{a}$. We claim that $d(f_n) \leq 1 + d(C)$. For if $d(f_n) > 1 + d(C)$, then since $q_j = \sum y_i c_{ij}$ can be expressed as $\sum f_i g_{ij}$, we have

$$d(\sum f_i g_{ij}) = d(q_j) \leq 1 + d(C) < d(f_n) \leq \max\{d(f_i) + d(g_{ij})\}.$$

This states that $f_1, \dots, f_n$ are right d -dependent, a contradiction, unless $g_{nj} = 0$ for all j . But then $q_j = \sum y_i c_{ij} = \sum_1^{n-1} f_i g_{ij} = \sum y_i d_{ih} g_{hj}$, hence $c_{ij} = \sum_1^{n-1} d_{ih} g_{hj}$ and this contradicts the fact that C is full. This then shows that $d(f_n) \leq 1 + d(C)$, and therefore $d(D) \leq d(C)$. Now $D = AU$ for an invertible matrix U over S . Setting the y_j equal to 0 provides a homomorphism from S to R which maps U to an invertible matrix U_0 and does not raise the degree, hence $d(AU_0) \leq d(C)$. ■

We can now describe the possible factorizations of a full linear matrix over $k\langle X \rangle$:

THEOREM 8.7. *Let $R = k\langle X \rangle$ and let $C \in R_n$ be a monic linear matrix which is full. Then C is not an atom if and only if either $n = 0$ or $n > 0$ and there exist*

$P, Q \in \text{GL}_n(k)$ and $r, s > 0$ such that $r + s = n$ and

$$PCQ = \begin{pmatrix} A & 0 \\ D & B \end{pmatrix}, \quad A \in R_r, B \in R_s, C \in {}^sR^r. \quad (11)$$

Proof. If (11) holds, then

$$PCQ = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ D & I \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & B \end{pmatrix},$$

and A, B are again monic linear, hence not invertible. We therefore have a non-trivial factorization of PCQ and hence of C .

Suppose, conversely, that $C = FG$ is a factorization of C . Any square factor of C is again full, and if it is a non-unit, its degree is positive; moreover by Lemma 8.6 we may take F to be of degree 1. Being full, F is regular, so by Cor. 8.3 there exist $P \in \text{GL}_n(k)$, $U \in \text{GL}_n(R)$ such that $PFU = A \oplus I$, where $A \in R_r$, and $0 < r < n$, because F and G are non-units. Hence

$$PC = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} G_1 & G_2 \\ G_3 & G_4 \end{pmatrix}. \quad (12)$$

Now PC is linear, hence (G_3, G_4) is linear and (G_1, G_2) has degree 0, since otherwise the leading term of (G_1, G_2) would be in the right annihilator of $\bar{A}$, whereas $\bar{A}$ is regular, because A is monic linear. Now (G_1, G_2) has r rows and since each factor in (12) is full, (G_1, G_2) has rank r , so there exists $Q \in \text{GL}_n(k)$ such that

$$(G_1, G_2)Q = (I_r, 0).$$

It follows that

$$PCQ = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ D & B \end{pmatrix} = \begin{pmatrix} A & 0 \\ D & B \end{pmatrix},$$

and this is of the required form. ■

We remark that if in a linear matrix A the cofactor of some x_i is the unit matrix, then A is necessarily full and monic. For such a matrix the criterion of Th. 8.7 takes the following form.

COROLLARY 8.8. *Let $R = k\langle X \rangle$ and let $A = A_0 + \sum A_i x_i$ be a linear $n \times n$ matrix over R , where $A_i = I$ for some $i > 0$. Then A is an atom if and only if $A_0, A_1, \dots, A_d$ act irreducibly on k^n .*

Proof. Clearly A is monic linear and full, so by Th. 8.7 it is not an atom if and

only if there exist $P, Q \in \mathbf{GL}_n(k)$ such that

$$PA_iQ = \begin{pmatrix} B_i & 0 \\ D_i & C_i \end{pmatrix}, \quad i = 0, 1, \dots, n. \quad (13)$$

By hypothesis $A_i = I$ for some $i > 0$, so in particular PQ has the block triangular form of (13), and hence so does $PA_iP^{-1} = PA_iQ(PQ)^{-1}$. ■

This result makes it easy to construct matrix atoms; we record one important case:

COROLLARY 8.9. *Let $R = k\langle X \rangle$ and let $A_1, \dots, A_N$ ($N = n^2$) be a k -basis of $\mathfrak{M}_n(k)$. Then $A = \sum A_i x_i$, where the x_i are distinct elements of X , is an absolute matrix atom, in fact it remains an atom under all commutative field extensions of k .*

Proof. Let $I_n = \sum \alpha_i A_i$ ($\alpha_i \in k$), where $\alpha_i \neq 0$, say. If we make a linear change of generators in R by writing $y_1 = \sum \alpha_i x_i$, $y_j = x_j$ ($j > 1$), then A satisfies the hypothesis of Cor. 8.8 relative to the y 's and hence is an atom; clearly it remains one under any extension of the ground field. ■

Let us return to eigenrings for a moment. We have seen that eigenrings of regular matrices (over free algebras) are finite-dimensional over k (Th. 8.4), a matrix atom has as eigenring a field (Prop. 3.3.4), and for a non-zero element the eigenring is commutative (Prop. 4.6.10). The latter no longer holds for matrices, for we shall see that any finite-dimensional k -algebra can occur as eigenring of a regular matrix, and any skew field (finite-dimensional over k) can occur as eigenring of a matrix atom.

THEOREM 8.10. *Let k be a commutative field and F a finite-dimensional k -algebra. Then*

(i) *there exists a torsion module M over a free k -algebra of finite rank, R , such that $\text{End}_R(M) \cong F$;*

(ii) *if F is also a field, then there is a $\mathcal{T}$ -simple module over the free k -algebra R of rank 2 such that $\text{End}_R(M) \cong F$.*

Proof. (i) Let $[F:k] = n$ and embed F in $E = \text{End}_k(F) \cong \mathfrak{M}_n(k)$ by letting F act on itself by left multiplications. We denote the image of F in E by F' and its centralizer in E by G . Since F acts bicentrally on itself, the centralizer of G is F' . Now G is finitely generated as k -algebra, by $A_0, A_1, \dots, A_m \in k_n$, say, where we may take $A_m = I$, without loss of generality. Let $R = k\langle x_1, \dots, x_m \rangle$ and put $A = A_0 + \sum A_i x_i \in R_n$; it is clear that A is full and monic linear. By (9)

and (10), we have

$$\begin{aligned}\text{End}_R(M) &\cong \mathbf{I}(A) \cap k_n \\ &= \{P \in k_n \mid PA_i = A_i Q, i = 0, 1, \dots, m \text{ and some } Q\}. \quad (14)\end{aligned}$$

Since $A_m = I$, we have $P = Q$ on the right of (14), therefore $\text{End}_R(M)$ is the centralizer of the A_i , hence of G , and so $\text{End}_R(M) \cong F$.

(ii) Now let F be a field, finite-dimensional over k . If k is a finite field, F is a commutative field extension of k (by Wedderburn's theorem), and we can write $F = k(\alpha)$ for some $\alpha \in F$. If p is the minimal polynomial of α over k , then F is the endomorphism ring of the simple torsion module $R/p(x)R$, where $R = k[x]$.

There remains the case when k is infinite. Let $[F:k] = n$ and let $f_1, \dots, f_n$ be a k -basis of F . Take n distinct elements $\lambda_1, \dots, \lambda_n$ of k and define matrices $A_0, A_1 \in F_n$ by

$$\begin{aligned}A_0 &= \text{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \\ A_1 &= \sum f_i e_{ii} + \sum e_{i1} + \sum e_{1j} = \begin{pmatrix} f_1 & 1 & \dots & 1 \\ 1 & f_2 & & 0 \\ \vdots & 0 & \ddots & \\ 1 & & & f_n \end{pmatrix}\end{aligned}$$

We claim that A_0, A_1 generate F_n as k -algebra. Since the λ_i are distinct, the subalgebra generated by A_0 contains all diagonal matrices over k , in particular, it contains each e_{ii} ; hence the subalgebra generated by A_0, A_1 also contains $e_{i1} = e_{ii}A_1e_{11}$ and $e_{1j} = e_{11}A_1e_{jj}$, and so also $e_{ij} = e_{i1}e_{1j}$. Thus it contains k_n ; it also contains $f_i e_{ii} = e_{ii}A_1e_{ii}$ and so contains all of F_n .

Now F_n has just one simple left F_n -module S say, up to isomorphism, and $[S:k] = n^2$. Consider the embedding $F_n \rightarrow \text{End}_k(S) \cong \mathfrak{M}_{n^2}(k)$; since F_n acts irreducibly on S , A_0 and A_1 act irreducibly on k^{n^2} and if A'_0, A'_1 are their images in $\mathfrak{M}_{n^2}(k)$, then $P = A'_0 + x_1 A'_1 + x_2 I$ is a matrix atom, by Cor. 8.8. Moreover, the centralizer of A'_0, A'_1 is the centralizer of F_n acting in $\text{End}_k(S)$ and so is isomorphic to F , so if M is the module over $R = k\langle x_1, x_2 \rangle$ defined by P , then M is $\mathcal{T}$ -simple and

$$\text{End}_R(M) = \mathbf{I}(P) \cap k_n.$$

As before this is the centralizer of A'_0, A'_1 , i.e. F , as we had to show. ■

Exercises 5.8

1. Extend Lemma 8.1 to matrices over $K_k\langle X \rangle$.

2. (M. L. Roberts) Let $R = K_k\langle X \rangle$ be a tensor K -ring. Show that every full matrix P over R is stably associated to a matrix

$$C + \sum A_i x_i B_i,$$

where $A_1, \dots, A_r$ are right comaximal, $B_1, \dots, B_r$ left comaximal and in each $A_i x_i B_i$ the columns of A_i are linearly independent over k , and likewise the rows of B_i .

3. Find a monic linear matrix over $k\langle X \rangle$ which is not full. (Hint: Try a hollow matrix.)

4⁰. Find a persistent fir R over k and finitely presented bound R -modules M, N such that $\text{Hom}_R(M, N)$ is not finite-dimensional over k .

5. Use Th. 8.10 to find a matrix over $\mathbf{R}\langle X \rangle$ with $\mathbf{C}$ as eigenring, and a matrix with the quaternions as eigenring.

6. (Cohn [76']) Let K be a field which is a k -algebra. Show that two square matrices A, B over K are conjugate over k if and only if $xI - A$ is stably associated to $xI - B$ over $K_k\langle x \rangle$.

7. (Cohn [76']) Show that if a given matrix P over $K_k\langle x \rangle$ is stably associated to $xI - A$ then A is determined by P up to conjugation by a matrix over k .

5.9 The specialization lemma

The inner rank and full matrices will come to play a basic role in the general construction of fields of fractions in Ch. 7. Briefly this may be explained as follows: a square matrix over a (skew) field is invertible if and only if it is full. Now let R be a ring and $\lambda: R \rightarrow K$ a homomorphism to a field; since the image under λ of a non-full matrix cannot be full, a matrix A over R has an image over K which can be non-singular only if A is full. So the most that we can expect to invert over a field are the full matrices over R ; for a free K -ring, $R = K_C\langle X \rangle$, there is actually a field U over which every full matrix (over R) has an invertible image. This field U is called the *universal field of fractions* of R ; it has indeed a universal property characterizing it which will be described in 7.2. In the present section we shall prove the existence of U for the free K -ring R ; this will provide an embedding of R in a field, because every non-zero element of R is full and so is inverted in U . Our main tool is a result of independent interest, the specialization lemma (Lemma 9.5). For the proof we shall need (i) a slight extension of Amitsur's theorem on generalized polynomial identities, (ii) the inertia theorem (Th. 2.9.15) and (iii)

some auxiliary results to be described below, as well as a lemma from Ch. 8 (Lemma 8.1.3). The existence proof of U in this section will also make use of the ultraproduct theorem, but we emphasize that the development in Ch. 7 is entirely independent of the results of this section.

We begin by recalling Amitsur's theorem. Let A be a k -algebra; by a *generalized polynomial identity* (GPI) one understands a non-zero element p of the free A -ring $A_k\langle X \rangle$ which vanishes under all maps $X \rightarrow A$. Here 'generalized' refers to the fact that the elements of A do not commute with the variables; for a 'polynomial identity' one takes an element of $k\langle X \rangle$. Amitsur [65] proved that a primitive k -algebra A satisfies a GPI if and only if it is a dense ring of linear transformations over a skew field of finite dimension over its centre, and A contains a non-zero transformation of finite rank. We shall only need the case where A is itself a skew field; in this case Amitsur's theorem takes the following form:

THEOREM 9.A. *A skew field satisfies a generalized polynomial identity if and only if it is of finite dimension over its centre.*

We shall need a generalization of this result in which the ground field, K say, need not be central or even commutative, and the element p of $A_K\langle X \rangle$ vanishes only when X ranges over the centralizer of K in A . The proof we shall give is similar to that of Th. 9.A; we shall follow the version in Herstein [76]. Until Th. 9.3, D will be a field with a subfield K ; the centralizer of K in D is again a field, denoted by K' , and its centralizer, written K'' and called the *bicentralizer* of K , necessarily includes K .

LEMMA 9.1. *Let D be a field and K a subfield. If $a_i, b_i \in D^\times$ satisfy*

$$\sum_1^n a_i x b_i = 0, \quad \text{for all } x \in K', \quad (1)$$

then the a_i are right K'' -linearly dependent.

Proof. Since $a_1, b_1 \neq 0$, we have $a_1 x b_1 \neq 0$, so we may take $n > 1$. By induction we may suppose that there is no such relation (1) for a proper subfamily of the a_i ; we define a mapping $\gamma_i: D \rightarrow D$ for $i = 1, \dots, n$ as follows. Any element of D can be written in the form $\sum u_j b_1 y_j$ for some $u_j \in K'$, $y_j \in D$ ($j = 1, \dots, r$), because $K'b_1D = D$. We define γ_i by

$$\gamma_i: \sum u_j b_1 y_j \mapsto \sum u_j b_i y_j. \quad (2)$$

To show that γ_i is well-defined we have to check that

$$\sum u_j b_1 y_j = 0 \Rightarrow \sum u_j b_i y_j = 0, \quad (i = 2, \dots, n).$$

Suppose then that $\sum u_j b_1 y_j = 0$; for any $v \in K'$ we have $\sum a_i v u_j b_i = 0$ by (1), hence

$$0 = -\sum a_1 v u_j b_1 y_j = \sum_{i=2}^n a_i v (\sum u_j b_i y_j).$$

This holds for all $v \in K'$ and is a shorter relation than (1); since $a_i \neq 0$, it follows that $\sum u_j b_i y_j = 0$ for $i = 2, \dots, n$, as we wished to show. Thus the mapping (2) is well-defined; clearly it is right D -linear and left K' -linear, hence it is a left multiplication by an element of K'' . By definition, $\gamma_i b_1 = b_i$, hence we have

$$0 = \sum a_i x b_i = \sum a_i x \gamma_i b_1 = (\sum a_i \gamma_i) x b_1,$$

therefore $\sum a_i \gamma_i = 0$, and $\gamma_1 = 1$, hence the a_i are right linearly dependent over K'' . ■

LEMMA 9.2. *Let D, K be as before and suppose that $a_1, \dots, a_n \in D$ are right K'' -linearly independent and $b_1, \dots, b_n \in D^\times$ are such that the set $\{\sum a_i x b_i \mid x \in K'\}$ is contained in a finite-dimensional left K -space. Then there exists $c \in D^\times$ such that KcK' is finite-dimensional as left K -space.*

Proof. If $n = 1$, we have $a_1 x b_1 = \sum \lambda_j(x) u_j$ ($\lambda_j \in K$), hence $\mu a_1 x = \sum \mu \lambda_j(x) u_j b_1^{-1}$ for any $\mu \in K$, $x \in K'$, and so the conclusion holds. We may therefore take $n > 1$ and use induction; thus assume that

$$\sum_1^n a_i x b_i = \sum \lambda_j(x) u_j,$$

where we may further suppose that $b_1 = 1$. Then we obtain for any $y \in K'$

$$\sum_2^n a_i x (b_i y - y b_i) = \sum \lambda_j(x) u_j y - \sum \lambda_j(xy) u_j.$$

Now the result follows by induction on n , unless $b_i y = y b_i$ for all $y \in K'$, but then $b_i \in K''$, $\sum a_i x b_i = (\sum a_i b_i) x$ and we are reduced to the case $n = 1$. ■

We can now prove the desired extension of Amitsur's theorem. For convenience in later applications we shall state it in contrapositive form.

THEOREM 9.3. *Let D be a skew field with a subfield K , such that (i) $K'' = K$, where K'' is the bicentralizer of K in D , (ii) for any $c \in D^\times$ the left K -space KcK' is infinite-dimensional over K . Then any non-zero element f of $D_K \langle X \rangle$ has a non-zero value for some choice of values of X in K' .*

Proof. Let $f \in D_K \langle X \rangle$ and suppose that $f \neq 0$ but that f vanishes for all values of X in K' . By linearization (i.e. replacing $f(\dots, x_i, \dots)$ by $f(\dots, x_i + y, \dots)$ —

$f(\dots, y, \dots)$, cf. e.g. Cohn [77], p. 453), we may take f to be multilinear of degree n , say. Let us single out the terms in which x_1 occurs last and write

$$f = \sum_1^h g_i x_1 b_i + \sum_1^k p_j x_1 q_j, \quad (3)$$

where $b_i \in D$ and q_j has positive degree. Without loss of generality we may take h, k minimal and $b_1 = 1$. We have

$$f(x_1, \dots, x_n)y - f(x_1y, \dots, x_n) = \sum_2^h g_i x_1 (b_i y - y b_i) + \sum p_j x_1 (q_j y - y q_j). \quad (4)$$

Since h was chosen minimal in (3), $1, b_2, \dots, b_h$ are left K -linearly independent, in particular $b_i \notin K$ for $i > 1$, so none of the terms $b_i y - y b_i$ in the first sum vanishes identically for y in K' . Choosing $y = y_0 \in K'$ such that $b_2 y_0 \neq y_0 b_2$, we obtain a GPI with a smaller h , unless $h = 1$, when the first sum on the right of (4) is absent. If $q_j y - y q_j$ vanishes identically for all $y \in K'$, then $q_j \in K'' = K$ for all values of the x 's in K' . Write $q_j = \sum c_i x_2 d_i$, where the values of $x_3, \dots, x_n$ in K' are so chosen that $q_j \neq 0$ (by an induction on n). Then $\sum c_i K' d_i$ is 1-dimensional over K , therefore KcK' is finite-dimensional over K , for some $c \in D^\times$, by Lemma 9.2. But this contradicts (ii).

We may thus assume that $q_j y - y q_j$ does not vanish identically, and for suitable $y = y_0 \in K'$, the left-hand side of (4) is then a GPI f_1 , again multilinear in $x_1, \dots, x_n$ with no terms in which x_1 is last. Moreover, each term in f_1 has the x 's in the same order as some term in f , so if x_v does not come last in any term of f , then the same is true for f_1 . We now apply the same reduction to $x_2, \dots, x_n$ in turn and finally get a GPI f^* in which no x_v comes last. This is impossible, so this case cannot occur. ■

We remark that if KcK' is n -dimensional over K , say, then we have a homomorphism (by right multiplication)

$$K' \rightarrow \text{End}_K(K^n) \cong \mathfrak{M}_n(K). \quad (5)$$

In particular, when K is commutative (and $K = K''$ as before), then it is the centre of K' and we have an embedding of K' in $\mathfrak{M}_n(K)$ as K -algebra, via (5), hence K' is then finite-dimensional over K . Thus if K is commutative and $[K':K] = \infty$, then KcK' cannot be finite-dimensional over K , for any $c \in D^\times$.

As already indicated, we shall need the inertia theorem, in the following form.

LEMMA 9.4. *Let D be a skew field with a subfield K . Then the embedding*

$$D_K \langle X \rangle \rightarrow D_K \llbracket X \rrbracket \quad (6)$$

is honest.

Proof. Let C be a full matrix over $F = D_K\langle X \rangle$ which is non-full over $\hat{F} = D_K\langle\langle X \rangle\rangle$, say $C = AB$, where A is $n \times r$, B is $r \times n$ and r is least; we shall assume that $r < n$ and obtain a contradiction.

By the inertia theorem (Th. 2.9.15) the inclusion (6) is totally inert, so there exists $P \in \text{GL}_r(\hat{F})$ such that on writing $A' = AP^{-1}$, $B' = PB$, the product of any row of A' by any column of B' lies trivially in F . Since C is full over F , A' and B' cannot have all their entries in F . If the $(1, 1)$ -entry of A' is not in F , then the first row of B' is 0, and on omitting the first column of A' and the first row of B' we can diminish r ; this contradicts the definition of r and the result follows. ■

We can now state the main result of this section.

LEMMA 9.5. (Specialization lemma). *Let D be a skew field whose centre C is infinite. Assume that K is a subfield of D which is its own bicentralizer and whose centralizer K' is such that KcK' is infinite-dimensional over K , for all $c \in D^\times$. Then any full matrix over $D_K\langle X \rangle$ is non-singular for some choice of X in K' .*

Proof. Let $A = A(X)$ be an $n \times n$ matrix over $D_K\langle X \rangle$ and let r be the supremum of its ranks as its arguments range over K' . By a translation $x \mapsto x + a$ ($a \in K'$) we may assume that the maximum rank is attained at the point $x = 0$, and by elementary transformations we may take the principal $r \times r$ minor of $A(0)$ to be invertible. Thus if

$$A(X) = \begin{pmatrix} B_1(X) & B_2(X) \\ B_3(X) & B_4(X) \end{pmatrix},$$

where B_1 is $r \times r$, then $B_1(0)$ is invertible. Given $a \in K'^X$ and any $\tau \in C$, we have $\rho A(\tau a) \leq r$, hence by Lemma 8.1.3, the rank of $A(\tau a)$ over $D(\tau)$ is at most r , and the same holds over $D((\tau))$. Now $B_1(\tau a)$ is a polynomial in t with matrix coefficients and constant term $B_1(0)$, a unit, hence $B_1(\tau a)$ is invertible over $D[[t]]$. By Prop. 4.6, the equation

$$B_4(\tau a) = B_3(\tau a)B_1(\tau a)^{-1}B_2(\tau a) \quad (7)$$

holds over $D[[t]]$, for all $a \in K'^X$. This then means that the matrix

$$B_4(tX) - B_3(tX)B_1(tX)^{-1}B_2(tX) \quad (8)$$

vanishes whenever X is replaced by any values in K' . Now (8) is a power series in t with coefficients that are matrices over $D_K\langle X \rangle$; thus the coefficients are generalized polynomial identities (or identically 0), so by Th. 9.3 the expression (8) vanishes identically as an element in the completion of

$R = D_K\langle X \rangle[t]$ in the t -adic topology. Hence the same equation holds in the (t, X) -adic topology, and also in the X -adic topology; for in each case the matrix $B_1(tX)$ has constant term $B_1(0)$ and so is invertible. Thus we may set $t = 1$ in (8) and find that

$$B_4(X) = B_3(X)B_1(X)^{-1}B_2(X) \quad \text{in } D_K\langle\langle X \rangle\rangle;$$

in other words, $A(X)$ is non-full over the power series ring. By Lemma 9.4, $A(X)$ is non-full over $D_K\langle X \rangle$, which is a contradiction, unless $r = n$. Thus every full matrix $A(X)$ is non-singular for some set of values in K' . ■

For the specialization lemma to hold, it is clearly necessary that K should coincide with its bicentralizer. For if $K'' \supset K$ and $\alpha \in K'' \setminus K$, then $\alpha x - x\alpha$ is non-zero and hence full, even though it vanishes for all values of x in K' .

We note the special case where K is the centre of D .

COROLLARY 9.6. *Let D be a skew field with infinite centre C and such that $[D : C] = \infty$. Then any full matrix over $D_C\langle X \rangle$ is non-singular for some set of values of X in D . ■*

If $[D : C]$ is finite, there are non-trivial identities over D , so this condition cannot be omitted. Whether the hypothesis that C be infinite can be relaxed is not known (cf. Ex. 3).

We can now prove the existence of a universal field of fractions for the tensor D -ring; the proof uses the ultraproduct theorem (cf. e.g. Cohn [81], Ch. VII), but no results in this book will be based on it, and another independent proof will be given in Ch. 7, without any restrictions on D and K .

THEOREM 9.7. *Let D be a skew field with infinite centre C and K a subfield such that $K'' = K$ and KcK' is infinite-dimensional over K for all $c \in D^\times$. Then $D_K\langle X \rangle$ can be embedded in a field U which is generated by $D_K\langle X \rangle$ and is such that every full matrix over $D_K\langle X \rangle$ becomes invertible over U .*

Proof. Consider the natural mapping

$$D_K\langle X \rangle \rightarrow D^{K'^X},$$

where $p \in D_K\langle X \rangle$ is mapped to (p_f) with $p_f = p(Xf)$, for any $f \in K'^X$. With each square matrix A over $D_K\langle X \rangle$ we associate the subset $\mathcal{D}(A)$ of K'^X consisting of all $f \in K'^X$ such that $A(Xf)$ is non-singular, the *singularity support* of A . Of course $\mathcal{D}(A) = \emptyset$ unless A is full, but by Lemma 9.5, $\mathcal{D}(A) \neq \emptyset$ for each full A . Now if A, B are non-singular, then so is $A \oplus B$, hence we have

$$\mathcal{D}(A) \cap \mathcal{D}(B) = \mathcal{D}(A \oplus B).$$

Moreover, the diagonal sum of full matrices is full, hence the family of $\mathcal{D}(A)$, where A is full, is closed under finite intersections and so is contained in an ultrafilter $\mathcal{F}$ on K^X . We thus obtain a homomorphism into an ultrapower

$$D_K\langle X \rangle \rightarrow D^{K^X}/\mathcal{F}, \quad (9)$$

and by definition every full matrix A over $D_K\langle X \rangle$ is non-singular on the set $\mathcal{D}(A)$ of $\mathcal{F}$, and is therefore non-singular in the ultrapower. Now the subfield of the ultrapower generated by the image of $D_K\langle X \rangle$ is the desired field of fractions. ■

The ultrapower on the right of (9) is again a field, V say, and from the construction it is clear that the embedding of $D_K\langle X \rangle$ in V induces an *elementary* embedding of D in V , i.e. a mapping preserving all first order sentences.

Exercises 5.9

1. Show that any field finite-dimensional over its centre satisfies a non-trivial GPI. Show also that for any finite field F there is a matrix A such that $A - xI$ is singular for all values of x in F .

2. Let k be an infinite field and let A be a full matrix over $k\langle X \rangle$. Show that there exists $n = n(A)$ such that for every central k -division algebra D of finite dimension at least n , A is non-singular for some set of values of X in D . (Hint: Use Kaplansky's theorem on polynomial identities.)

3°. Let K be an infinite field with finite centre, and A a square matrix over K . Does there always exist $\alpha \in K$ such that $A - \alpha I$ is non-singular?

4. Show that every matrix over $K_C\langle X, X^{-1} \rangle$, where K is a field with centre C , is stably associated to a linear matrix $A_0 + \sum A_i x_i$ ($x_i \in X$). If $[K:C] = \infty$, $|C| = \infty$, deduce that for any full matrix A over $K_C\langle X, X^{-1} \rangle$ there exist non-zero values of X in K for which A is non-singular. (Hint: Apply the specialization lemma to $A \oplus \text{diag}(x_1, \dots, x_r)$)

5.10 Ascending chain conditions

In 3.3 we saw that any $2n$ -fir with left and right ACC_n is n -atomic. To prove a corresponding result for rectangular matrices in 5.7 we needed to assume that R is a fir; but merely assuming left ACC_n will enable us to split off a 'maximal' left factor from any matrix of inner rank n :

PROPOSITION 10.1. *Let R be a $2m$ -fir and A an $m \times n$ matrix of inner rank t over R . Suppose that A can be written in the form*

$$A = PA', \quad \text{where } P \in {}^m R^t, A' \in {}^t R^n, \quad (1)$$

where A' is left prime. Then in any other factorization $A = P'A''$ of A into an $m \times t$ matrix P' by a $t \times n$ matrix A'' , there is a matrix V such that $A'' = VA'$, $P = P'V$.

Such an expression (1) for A (with A' left prime) exists whenever R has left ACC_m or is m -atomic.

Proof. Let $A = PA' = P'A''$, where P, P' have t columns and A' is left prime. Then $(P, -P')(A', A'')^T = 0$, hence by the law of nullity,

$$\rho(P, -P') + \rho(A', A'')^T \leq 2t;$$

but each summand is at least t , so $\rho(A', A'')^T = t$ and we have

$$\begin{pmatrix} A' \\ A'' \end{pmatrix} = \begin{pmatrix} U \\ V \end{pmatrix} C, \quad \text{where } U, V \in R_t, C \in {}^tR^n.$$

Since A' is left prime, U is a unit and so may be absorbed in C . Thus $C = A'$, $A'' = VA'$, $A = P'A'' = P'VA' = PA'$. Since A' is left full, it is left regular and so $P = P'V$, as claimed.

Now assume left ACC_m and write $A = P_0A_0$, where $P_0 \in {}^mR^t$, $A_0 \in {}^tR^n$ and A_0 is left full. Since $t \leq m$, we can choose a maximal t -generator submodule of R^n containing R^tA_0 ; this is of the form R^tA' , where the rows of A' are the generators. The equation (1) follows, and A' is left prime by the maximality of R^tA' .

When R is m -atomic, write $A = P_1A_1$, where P_1 is $m \times t$ and A_1 is $t \times n$; then A_1 is left full. Let A'_1 be a full $t \times t$ submatrix of A_1 , then in any factorization

$$A_1 = P_2A_2, \quad \text{where } P_2 \in R_t, A_2 \in {}^tR^n,$$

P_2 is a left factor of A'_1 and so the number of terms in a complete factorization of P_2 is bounded by the corresponding number for A'_1 . Therefore we can ensure that A_2 is left prime by taking P_2 with a maximal number of factors. ■

By restating the result in terms of modules we see that for the rings considered every finitely presented module has a largest positive quotient, and dually, a largest negative submodule; this was proved for the special case of firs in Th. 3.5.

THEOREM 10.2. *Let R be a $2m$ -fir which has left ACC_m or is m -atomic, and let M be an m -generator submodule of a free left R -module F . Denote by t the least integer for which there is a t -generator submodule between M and F , thus $t \leq m$. Then there is a greatest t -generator submodule N between M and F . ■*

Here 'greatest' is understood in the sense that N contains every t -generator submodule of F containing M . Thus N/M is the greatest negative submodule of F/M . A dual argument shows that every finitely presented right R -module has a largest positive quotient. Thus we have

COROLLARY 10.3. *Let R be a $2m$ -fir which has left and right ACC_m or is m -atomic, and let F be a free left R -module of rank at most m with an m -generator submodule M . Then F/M has a largest positive quotient and a largest negative submodule. ■*

We now consider another chain condition which entails ACC_n for all n and holds in all firs. A module M (over any ring R) is said to satisfy ACC_{ds} if every chain of finitely generated submodules of M with bound quotients (or equivalently: dense inclusions) must break off. If a ring R satisfies ACC_{ds} as left (or right) module over itself, we shall say that R satisfies *left* (or *right*) ACC_{ds} .

To clarify the connexion between ACC_{ds} and ACC_n we shall need a result on chains with bound quotients. Let us call two ascending sequences of elements in a partially ordered set *cofinal* if each element of either sequence precedes some element of the other sequence.

LEMMA 10.4. *In an n -fir R , let M be a module with a sequence*

$$M_1 \subseteq M_2 \subseteq \dots \quad (2)$$

of submodules, each free of rank at most n , and assume that (2) is not cofinal with any sequence of submodules all free of rank less than n . Then M_i/M_{i-1} is bound for all large i .

Proof. Any M_i/M_{i-1} that is not bound will have R as a direct summand, by Cor. 1.4, and its complement corresponds to a submodule M'_i of smaller rank than M_i , so of rank less than n , such that $M_{i-1} \subset M'_i \subset M_i$. If this happens for infinitely many i , then the sequence M'_i is cofinal, but of rank less than n . ■

Suppose now that R is an n -fir with left ACC_{ds} . Given any ascending chain of n -generator left ideals of R , let $m \leq n$ be the least integer for which our chain is cofinal with a chain of left ideals that are free of rank at most m . Applying Lemma 10.4 to this chain, we see that it has bound quotients and hence terminates. This proves

COROLLARY 10.5. *For any n -fir, left ACC_{ds} implies left ACC_n . In particular, a semifir with left ACC_{ds} satisfies left ACC_n for all n . ■*

The next result elucidates the role of ACC_{ds} in modules over semifirs.

PROPOSITION 10.6. *Let R be a ring and M an R -module. If every countably generated submodule of M is free, then M satisfies ACC_{ds} . Conversely, if R is a semifir, M satisfies ACC_{ds} and all finitely generated submodules of M are free, then every countably generated submodule of M is free.*

Proof. Let M be a left R -module whose countably generated submodules are free, and consider a chain

$$M_1 \subseteq M_2 \subseteq \dots$$

of dense inclusions of finitely generated submodules in M . Their union M' is countably generated and hence free. Since M_1 is a finitely generated submodule of M' , it involves only finitely many members of a basis of M' , and so is contained in a free direct summand of finite rank, N , of M' , say $M' = N \oplus N'$, where N' is also free. Let $p: M' \rightarrow N'$ be the projection onto N' . For any linear functional α on N' , $p\alpha$ is a linear functional on M' , zero on M_1 , hence by density zero on each M_i , and hence zero on M' ; it follows that $N' = 0$, so $M' = N$ is finitely generated and our chain must terminate.

Conversely, assume that R is a semifir, M satisfies ACC_{ds} and all its finitely generated submodules are free. Let N be a submodule generated by countably many elements $u_1, u_2, \dots$. Put $N_0 = 0$ and for each $i \geq 0$ let us recursively construct N_{i+1} as a maximal finitely generated submodule of N in which $N_i + Ru_{i+1}$ is dense; this is possible by ACC_{ds} . We claim that each N_i is a direct summand in N_{i+1} . Indeed, let N'_i be a direct summand of N_{i+1} containing N_i , which is free of least possible rank. Any linear functional on N'_i which is zero on N_i will have for kernel a direct summand of N'_i (and hence of N_{i+1}), which contains N_i and is free of smaller rank than N'_i (by Th. 1.1.1(f)), unless the functional is zero. But N'_i was of minimal rank, hence the functional must be zero and N_i is dense in N'_i . By construction of N_i we therefore have $N'_i = N_i$.

We thus have $N_{i+1} = N_i \oplus P_{i+1}$, say, where P_{i+1} , being finitely generated, is free. Hence $N = \bigcup N_i = P_1 \oplus P_2 \oplus \dots$ is also free. ■

Bearing in mind Th. 1.5, we can state the result as

COROLLARY 10.7. *Let R be a semifir. Then a countably generated R -module is free if and only if it is n -unbound for all n and has ACC_{ds} . ■*

From the first part we moreover obtain a sharpening of Th. 1.2.3.

COROLLARY 10.8. *If R is a left $\aleph_0$ -fir, then any free left R -module satisfies ACC_{ds} and hence ACC_n for all n . ■*

As a consequence we can show that $k\langle\langle X \rangle\rangle$ is not an $\aleph_0$ -fir; we saw in Ex. 3.4.7 that it is not a fir.

PROPOSITION 10.9. *If X is a set with more than one element, then $k\langle\langle X \rangle\rangle$ is not an $\aleph_0$ -fir; likewise the subring $k_{\text{alg}}\langle X \rangle$ of algebraic power series is not an $\aleph_0$ -fir.*

This is in contrast to $k_{\text{rat}}\langle X \rangle$, which will be shown to be a fir in Th. 7.10.7.

Proof. Let x, y be distinct elements of X and consider the element $v = \sum_0^\infty x^i y x y^i$. It is clear that v satisfies the equation

$$v = xvy + yx, \quad (3)$$

and moreover, (3) determines v uniquely, as an algebraic power series. Let a_n be the left ideal of $k\langle\langle X \rangle\rangle$ generated by xy^i ($i = 0, 1, \dots, n$) and vy^{n+1} . These elements are left linearly independent, for if $\sum a_i xy^i + bvy^{n+1} = 0$, then a_i must vanish as left cofactor of xy^i ($i \leq n$), so also $b = 0$. Hence a_n is free on these elements as basis. Now a_n/a_{n-1} is bound, for if λ is any linear functional which vanishes on xy^i ($i < n$) and vy^n , then

$$0 = (vy^n)\lambda = (xvy^{n+1} + yxy^n)\lambda = x(vy^{n+1}\lambda) + y(xy^n)\lambda,$$

therefore $vy^{n+1}\lambda = xy^n\lambda = 0$, and so $\lambda = 0$. On the other hand, $xy^n \notin a_{n-1}$, so $a_{n-1} \subset a_n$. This shows that $k\langle\langle X \rangle\rangle$ does not satisfy left ACC_{ds} and so by Cor. 10.8 is not a left $\aleph_0$ -fir.

Now consider $R_1 = k_{\text{alg}}\langle X \rangle$; here the argument is the same: we have a semifir, admitting transductions, and again obtain a sequence of left ideals violating left ACC_{ds} . ■

Thus, the power series ring is an example of a semifir satisfying ACC_n for all n , but not ACC_{ds} . By contrast, in Bezout domains, ACC_{ds} can be replaced by 'ACC $_n$ for all n ';

PROPOSITION 10.10. *Over a right Bezout domain R , any torsion-free (=1-unbound) left R -module with ACC_n for all n has ACC_{ds} .*

Proof. Let M be a torsion-free left R -module and take a chain

$$M_1 \subseteq M_2 \subseteq \dots \quad (4)$$

of finitely generated submodules with dense inclusions. By Prop. 1.1.9, any finitely generated torsion-free left R -module is free. If for some i , $\text{rk } M_{i-1} < \text{rk } M_i$, then the induced map $K \otimes M_{i-1} \rightarrow K \otimes M_i$ is not surjective, where K is the field of fractions of R , hence there is a non-zero K -linear functional on $K \otimes M_i$ which vanishes on $K \otimes M_{i-1}$. By right multiplication with an appropriate element of R we may assume that the induced map takes M_i into R . Thus we have a linear functional on M_i which is zero on M_{i-1} without vanishing, and this contradicts the fact that M_i/M_{i-1} is bound. Thus all the ranks in (4) must be equal, to n say, and the sequence terminates by ACC_n . ■

Prop. 10.10 cannot be simplified by taking ACC_n for a fixed n only, since even over $\mathbb{Z}$ the conditions ACC_n ($n = 1, 2, \dots$) are independent (cf. Ex. 9). If we combine Prop. 10.10 with Cor. 10.7, we obtain a somewhat surprising conclusion:

COROLLARY 10.11. *Let R be a right Bezout domain with left ACC_n for all n . Then any countably generated left R -module embedded in a direct power R^I is free. ■*

This shows, for example, that every countably generated subgroup of $\mathbb{Z}^I$ is free although, of course, $\mathbb{Z}^I$ itself is not free, unless I is finite (by Lemma 3.4.13).

We conclude this section with a result on what may be called ' α -complete' direct limits of α -firs. We shall call a partially ordered set α -directed if every subset of cardinal at most α has an upper bound. A directed system over an α -directed set is also called an α -directed system.

PROPOSITION 10.12. *Let α be any cardinal greater than 1. Then the direct limit of any α -directed system of α -firs is again an α -fir.*

Proof. Let $\{R_i\}$ be the given system of α -firs, with maps $f_{ij}: R_i \rightarrow R_j$, and write $R = \varinjlim R_i$. When α is finite, the assertion is simply that the direct limit of any directed system of n -firs is an n -fir; this follows easily from the characterization of n -firs given in Th. 1.1.1(a). Thus we may take α to be infinite. Further, by the finite case, R will be a semifir, so it remains to show that given any set $X \subseteq R$, of cardinal at most α , RX will be a free left ideal of R .

Since our system is α -directed we can find a ring R_{i_0} such that each $x \in X$ has an inverse image x' in R_{i_0} . Write X' for the set of all these x' . For each finite subset Y of X' and each $i \geq i_0$ consider the rank of the left ideal $R_i(Yf_{i_0i})$. For fixed Y this rank is non-increasing in i , so it ultimately equals

a minimum, which it attains for some i depending on Y . Since there are no more than α finite subsets Y of X' , we can find $i_1 \geq i_0$ such that all the ranks $\text{rk}(R_i Y f_{i_0 i})$ have their minimum value for $i \geq i_1$. Put $X'' = X' f_{i_0 i_1}$ and let B be a basis for the left ideal $R_{i_1} X''$ of the α -fir R_{i_1} . We assert that for all $i \geq i_1$ and each finite $C \subseteq B$, $C f_{i_1 i}$ is left linearly independent in R_i . It will follow that the image of C in R is left linearly independent, hence the image of B will be left linearly independent, and so will form a basis of RX .

Thus assume that $C f_{i_1 i}$ is linearly dependent. Pick a finite subset Y of X'' such that $C \subseteq R_{i_1} Y$; then $R_{i_1} C$, being a direct summand in $R_{i_1} B \supseteq R_{i_1} Y$, will be a direct summand in $R_{i_1} Y$. Hence C can be extended to a basis C' of $R_{i_1} Y$. But $C f_{i_1 i}$ is linearly dependent, so $\text{rk}(R_{i_1} Y f_{i_1 i}) < |C'| = \text{rk}(R_{i_1} Y)$, and this contradicts our choice of i_1 . ■

Exercises 5.10

1. (Continuation of Ex. 1.8) If R is a semifir and $\mathcal{C}$ the class of R -modules that are n -unbound for all n and have ACC_{ds} show that $\mathcal{C}$ is closed under the operations listed in (a) but not (b). Show that R must have left ACC_{ds} for $\mathcal{C}$ to be non-trivial. If R is right Bezout, show that $\mathcal{C} = \cap \mathcal{C}_n$ and hence in this case $\mathcal{C}$ admits (b) too.

2. Let R be a right Ore domain. Show that if a torsion-free left R -module satisfies ACC on submodules of rank at most n , for all n , then it also satisfies ACC_{ds} (use Prop. 10.10).

3*. (Bergman [67]) Let R be an integral domain which is not a right Ore domain. Show that $R^{\mathbb{N}}$ as left R -module does not have ACC_{ds} . (Hint: Let $a, b \in R$ be right linearly independent and define $e_i, f_i \in R^{\mathbb{N}}$ by $e_i \pi_j = \delta_{ij}$, $f_i \pi_j = a^{j-i} b$, where $a^r = 0$ for $r < 0$, and π_j is the projection on the j th factor. Verify that $R f_i$ is dense in $R e_i + R f_{i+1}$ and deduce that $M_i = R e_i + \dots + R e_i + R f_{i+1}$ form a strictly ascending chain of dense inclusions.)

4. Verify that the first part of Prop. 10.1 holds for Sylvester domains.

5. Let R be a semifir and $\mathfrak{a} = \cup \mathfrak{a}_i$, where $\mathfrak{a}_i$ is a finitely generated left ideal properly containing $\mathfrak{a}_{i-1}$ as dense submodule. Show that $\mathfrak{a}$ is countably generated but not free.

6. Let kF be the group algebra over k of the free group F on x, y, z and let R be the subalgebra generated by z and all $z^{-n} x, z^{-n} y$ ($n = 1, 2, \dots$). Show that R is a right but not left fir (cf. Th. 2.10.3). Show also that (x, y) cannot be written as $u(x', y')$, where (x', y') is left prime (cf. Cohn [82]).

7. Using Ex. 1.5, show that any left hereditary integral domain has left ACC_1 .

8. Prove the following converse of Prop. 10.10: if R is a semifir over which each torsion-free left R -module with ACC_n for all n satisfies ACC_{ds} , then R is right Bezout.

9*. (Bergman [67]) (a) For any prime number p denote by $\mathbb{Z}_p$ the ring of p -adic integers and by $\mathbb{Q}_p = \mathbb{Z}_p[p^{-1}]$ its field of fractions (the p -adic numbers), so that $\mathbb{Z}[p^{-1}] \cap \mathbb{Z}_p = \mathbb{Z}$. Let $1, x_1, \dots, x_n$ be any $\mathbb{Z}$ -linearly independent elements of $\mathbb{Z}_p$ and define a subgroup G of $\mathbb{Q}_p^{n+1}$ by the equation

$$G = (\mathbb{Z}_p^{n+1} + (1, x_1, \dots, x_n)\mathbb{Q}_p) \cap \mathbb{Z}[p^{-1}]^{n+1}.$$

Verify that any finitely generated subgroup of G can be generated by $n+1$ elements. Show further, that for any $k > 0$, there exists $a = (1, a_1, \dots, a_n) \in \mathbb{Z}^{n+1}$ such that $p^{-k}a \in G$ and deduce that G is not finitely generated. Hence show that G does not satisfy ACC_{n+1} .

(b) Let G be defined as in (a) and suppose that C is a union of n -generator subgroups in G . Show that C is annihilated by a non-zero $\mathbb{Q}$ -linear functional λ with coefficients in $\mathbb{Z}$. Show, further, that C is contained in $p^{-k}\mathbb{Z}^{n+1}$, where p^k is the highest power of p dividing $\lambda((1, x_1, \dots, x_n))$. Deduce that C is finitely generated and that G has ACC_n . This shows that the conditions ACC_n ($n = 1, 2, \dots$) are independent, even over $\mathbb{Z}$.

10°. Find an ACC (of the type ACC_{ds}) such that every semifir satisfying this condition on the left is a left fir.

11°. Find an example of a left fir which is elementarily equivalent to a fir, but is not two-sided (cf. Ex. 8.1.12).

12°. In a left Ore domain, does left ACC_n for some n , or left ACC_{ds} imply the corresponding condition for free modules?

13. In Prop. 10.4, show that the M_i/M_{i-1} are torsion modules for all large i .

14. An *involution* of a ring R is an antiautomorphism whose square is the identity. If $*$ is an involution of R , verify that $A \mapsto A^H$, where the entries of A^H are $a_{ij}^H = a_{ji}^*$ ($A = (a_{ij})$), is an involution of R_n . Let R be a semifir with involution $*$ and let $A \in R_n$ satisfy $A^H = A$. Show that if $\rho A = r$, then $A = P A_1 P^H$, where $P \in {}^n R^r$ and $A_1^H = A_1$. If moreover, R is a fir, show that P may be taken to be right prime.

5.11 The intersection theorem for firs

In this section we shall apply the ACC for bound submodules of finitely related modules over a fir to show that the intersection of the powers of an ideal in a fir is zero. In fact, we shall prove a somewhat more general statement which will be needed in Ch. 6.

We begin by considering an arbitrary ring R . If $a_1, a_2, \dots$ is any sequence of (left, right or two-sided) ideals of R , we define

$$\prod_{i \geq 1} a_i = \bigcap_{n \geq 1} a_1 a_2 \dots a_n.$$

Clearly the left-hand side is a (left, right, two-sided) ideal whenever all the a_i are.

LEMMA 11.1. Let α_1 be a right ideal and let $\alpha_2, \alpha_3, \dots$ be a sequence of two-sided ideals of a ring R , and assume that α_1 is free as right R -module, with basis (e_λ) , say. Then any $a \in \prod \alpha_i$ can be written as

$$a = \sum e_\lambda a_\lambda, \quad \text{where } a_\lambda \in \prod_{i \geq 2} \alpha_i. \quad (1)$$

Proof. By hypothesis, $a \in \alpha_1 \alpha_2 \dots \alpha_n$, so we have an expression $a = \sum u_{i1} u_{i2} \dots u_{in}$, where $u_{ir} \in \alpha_r$; in particular there is an expression (1) for a with

$$a_\lambda \in \alpha_2 \dots \alpha_n. \quad (2)$$

Since the e_λ are right linearly independent, the a_λ are unique and (2) holds for all n , which is the assertion. ■

LEMMA 11.2. Let R be a ring and α a proper non-zero free right ideal of R , with basis (e_λ) . Given $a_1, \dots, a_r \in \alpha$, left linearly independent over R , write

$$a_i = \sum e_\lambda b_{\lambda i}. \quad (3)$$

Then $\sum R b_{\lambda i} / \sum R a_j$ is a non-zero bound module.

Proof. Clearly $\sum R a_j \subseteq \sum R b_{\lambda i}$; if equality holds here, we have

$$b_{\lambda i} = \sum c_{\lambda i j} a_j;$$

on substituting into (3) we get $a_i = \sum e_\lambda b_{\lambda i} = \sum_{\lambda j} e_\lambda c_{\lambda i j} a_j$. Since the a_i are left linearly independent, we have $\sum e_\lambda c_{\lambda j j} = 1$, which contradicts the fact that α is proper. To show that $M = \sum R b_{\lambda i} / \sum R a_j$ is bound, take any linear functional on M ; this maps $b_{\lambda i}$ to $\beta_{\lambda i} \in R$ say, and by (3) we must have $\sum e_\lambda \beta_{\lambda i} = 0$. By the right linear independence of the e_λ , $\beta_{\lambda i} = 0$ for all λ, i , so our functional is 0, as we had to show. ■

We can now prove the intersection theorem:

THEOREM 11.3. Let R be a left fir and $\alpha_1, \alpha_2, \dots$ a sequence of proper two-sided ideals which are free as right ideals. Then $\prod_{i \geq 1} \alpha_i = 0$.

Proof. Suppose that $\prod \alpha_i \neq 0$ and let c_1 be a finitely generated non-zero left ideal contained in $\prod \alpha_i$. Let $a_1, \dots, a_r$ be a basis of c_1 and (e_λ) a basis of α_1 as right ideal, and write

$$a_i = \sum e_\lambda b_{\lambda i}.$$

Then $b_{\lambda i} \in \prod_{n \geq 2} \alpha_n$ by Lemma 11.1, hence $c_2 = \sum R b_{\lambda i} \subseteq \prod_{n \geq 2} \alpha_n$ and by Lemma 11.2, c_2/c_1 is bound and non-zero. By induction we obtain a sequence

of finitely generated left ideals c_n such that $c_n \subseteq \prod_{m \geq n} a_m$ and we have the strictly ascending chain

$$c_1 \subset c_2 \subset \dots$$

with bound quotients. Thus we have an infinite ascending sequence of bound submodules of R/c_1 , which contradicts Cor. 1.7. ■

In particular, taking R to be a two-sided fir, we find that the conclusion holds for any sequence of proper two-sided ideals. Taking all ideals equal, we obtain

COROLLARY 11.4. *In a two-sided fir the intersection of the powers of any proper two-sided ideal is zero.* ■

Exercises 5.11

1. Give a direct proof of Cor. 11.4 for principal ideal domains.
2. Let R be an integral domain with left ACC₁. Show that any proper two-sided ideal a that is principal as right ideal satisfies $\bigcap a^n = 0$.
- 3⁰. Let a be a finitely generated ideal in a commutative ring R . If $a^2 = a$, show that $a = eR$ for an idempotent e . Find a generalization to the non-commutative case.
4. Show that in a left fir R , if $ab \subseteq a'b$ for two-sided ideals a, a' and a left ideal b , then $a \subseteq a'$ (cf. 6.6).
5. Show that the conclusion of Cor. 11.4 does not hold for the one-sided fir constructed in 2.10.
6. Let F be a free group and kF the group algebra of F . If a is the augmentation ideal, show that $\bigcap a^n = 0$. Define the lower central series of F recursively by $\gamma_1 F = F$, $\gamma_{n+1} F = [\gamma_n F, F]$ (commutator subgroup), and show that for any $u \in \gamma_n F$, $u \equiv 1 \pmod{a^n}$; deduce that $\bigcap \gamma_n F = 1$ (Magnus' theorem).

Notes and comments on Chapter 5

The factorization of square matrices over firs which was one of the main objectives of this chapter in the first edition, has now been transferred to Ch. 3. In the present version there is a greater emphasis on modules, which is also reflected in the title (changed from 'Linear algebra over firs and semifirs').

Sections 5.1 and 2, based on Cohn [73''', 77'''], describe the background in the more general setting of hereditary rings. The notion of torsion class used in 5.1 is basic in the study of torsion theories (cf. e.g. Stenström [75]), but we need only the most elementary properties; in any case the usual treatment deals

mainly with hereditary torsion theories (in which the torsion class admits subobjects), and so does not apply here. The transpose $\text{Tr}(M) = \text{Ext}_k^1(M, R)$ has also been used by M. Auslander in the study of Artin algebras.

Torsion modules over firs were first described in Cohn [67] and in the first edition formed the basis of the factorization of full matrices. The positive and negative modules (corresponding to left and right full matrices) are studied in Cohn [77', 82]. The Krull-Schmidt theorem for finitely presented modules over firs (Th. 3.9) is new; it has also been obtained independently by A. H. Schofield (unpublished). The application in Th. 3.14, giving a conceptual proof of the embedding of a fir in a field, is due to Bergman [a].

The treatment of rank in 5.4 essentially follows Bergman [67], but Th. 4.9 (which is used in 5.5) is taken from Cohn [74]. Sylvester domains were introduced by Dicks and Sontag [78], and 5.5 is based on this source, but the presentation has been modified here so as to be independent of the results of Ch. 7. The law of nullity first occurs (for the case of fields) in Sylvester [1884]; in the case of semifirs it first appeared in the first edition. In studying localization (Ch. 7) we shall need to consider pseudo-Sylvester domains; their properties, described in 5.6 are taken from Cohn and Schofield [82]. The analysis of matrix relations and factorization in 5.7 is mostly new; see also Cohn [82''].

The normal form which is the subject of 5.8 is due to Roberts [84] and all the results of this section are taken from that source, except for Lemma 8.1, which is implicit in Roberts [82], where the more general normal form of Ex. 8.2 also occurs. Cor. 8.5 was first proved (by a different method) by Lewin [69]. The uniqueness of Th. 8.2 was proved in a special case in Cohn [76']. Th. 8.7 on the factorization of matrices over a free algebra, and Lemma 8.6 on which it depends, are due to A. H. Schofield, who is also responsible for the elegant proof of Th. 8.10 (replacing an earlier proof by Roberts which required larger rank in (ii)).

The specialization lemma (with K as the centre of D) was first proved in Cohn [72''] (in which the proof owed important simplifications to Bergman), and applied in Cohn [77']. The present version (with K non-central, giving incidentally in Th. 9.3 a strengthened form of Amitsur's theorem) is new. Th. 9.7 is taken from Cohn [84]; for other applications of the specialization lemma, to transfer first order properties, see Révész [81].

The study of ascending chain conditions in 5.10 is taken from Bergman [67], but there have been some additions in this edition, notably 10.1-3 from Cohn [82]. Prop. 10.6 was previously stated for ideals in $\aleph_0$ -firs, while Prop. 10.9 was an exercise in the first edition. Cor. 10.7 generalizes Pontryagin's theorem: 'A countably generated torsion free abelian group with ACC_n for all n is free' (Pontryagin [39], p. 168). In the case of $\mathbb{Z}$, related results have been obtained by Specker [50], who shows, for example, that in the subgroup B of $\mathbb{Z}^I$ consisting of all bounded sequences, every subgroup of cardinal at most $\aleph_1$ is free. Dubois [66] and independently Nöbeling [68] have shown that B itself is free, and more recently Bergman [72] has given a very brief proof of this fact. The intersection theorem in 5.11 was first proved in the case where all the α_i are equal, by Cohn [70'], and this appeared in the first edition. It was greatly generalized by Bergman [72', 73]; a special case of his results yields Th. 11.3, as well as a slight simplification of the proofs.

6. Centralizers, subalgebras and automorphisms

The first topic of this chapter is commutativity in firs. We shall find that any maximal commutative subring of a 2-fir with strong DFL is integrally closed (Cor. 1.2), and the same method allows us to characterize the centres of 2-firs as integrally closed rings and the centres of principal ideal domains as Krull domains (6.1 and 2). Further, the centre of a fir that is not principal is a field (6.3). Sections 6.1 and 6.4 are devoted to a study of invariant elements in 2-firs and their factors; the well-known result that a simple proper homomorphic image of a principal ideal domain is a matrix ring over a skew field is generalized here to atomic 2-firs (Th. 4.3).

Secondly, we look at subalgebras and ideals of free algebras, in 6.6; by way of preparation submonoids of free monoids are treated in 6.5. A brief excursion into coding theory shows how the Kraft–McMillan inequality can be used to find free subalgebras, and the fir property of free algebras is again derived (Th. 6.8). Section 6.7 is devoted to a fundamental theorem on free algebras: Bergman's centralizer theorem (Th. 7.7).

The tameness of automorphisms of free algebras of rank 2 was proved (by Makar–Limanov and Czerniakiewicz) just as the first edition went to press, too late to be included. In the present edition there is a fairly complete treatment of this work, as well as the corresponding result for polynomial rings (Jung and v.d.Kulk), on which it depends. These and a number of related questions, such as the Jacobian problem, are discussed in 6.8 and 9. Section 6.10 deals with invariants of automorphisms of free algebras, and 6.11 with the Galois correspondence between automorphism groups and free subalgebras, as described by Kharchenko; these topics are still being developed, and no more than a brief introduction could be given here.

6.1 Commutative subrings and central elements in 2-firs

Just as commutative 2-firs have a rather special form, so it is possible to say more about commutative subrings of general 2-firs. Assuming strong DFL, we shall show that maximal commutative subrings are integrally closed. We recall that if A, A' are commutative integral domains and $A \subseteq A'$, then an element $y \in A'$ is *integral* over A if there is a monic equation for y with coefficients in A : $y^n + a_1 y^{n-1} + \dots + a_n = 0$ ($a_i \in A$). Equivalently, the A -module

generated by the powers of y is finitely generated over A . The set of all elements of A' integral over A forms a subring $\bar{A}$ of A' , the *integral closure* of A in A' , and A is *integrally closed* in A' if $\bar{A} = A$. By a *finite integral extension* of A in A' we understand a ring between A and A' which is finitely generated as A -module. Suppose that B is a finite integral extension of A in its field of fractions k , then we can write $B = \sum_1^n Au_i$, where $u_i \in k$. The u_i may be brought to a common denominator, say $u_i = a_i b^{-1}$ ($a_i, b \in A$); then $B = \sum Aa_i b^{-1}$, and so $Bb \subseteq A$; conversely, when this holds, then B is finite integral over A . This means that the *conductor* of A in B , defined as the set

$$c = \{c \in A \mid Bc \subseteq A\},$$

is different from 0. We note that c may also be described as the largest ideal in A which is also an ideal in B .

We begin with a result extending embeddings of commutative rings to integral extensions.

PROPOSITION 1.1. *Let R be a 2-fir, A a commutative subring of R and B a finite integral extension of A in its field of fractions. Then there exists an injection $f: B \rightarrow R$ and $e \in R^\times$, such that*

$$ex = xf \cdot e, \quad \text{for all } x \in A. \quad (1)$$

If, moreover, R satisfies right strong DFL ($baR \cap a'R \neq 0$, where a, a' are similar implies $ba \in a'R$), then e can be taken to be 1.

Proof. Since B is finite integral over A , there exists $c \in A^\times$ such that $Bc \subseteq A$. Put $c = Bc$; this is a non-zero finitely generated ideal of A (and of B). Any two non-zero elements of c have a non-zero common left multiple in A , hence in R ; since R is a 2-fir, it follows that Rc is principal, say $Rc = Re$ for some $e \in R^\times$. We have $eA \subseteq RcA \subseteq Rc = Re$, hence there is a map $f_0: A \rightarrow R$ such that

$$ea = af_0 \cdot e, \quad \text{for all } a \in A.$$

Clearly f_0 is injective; moreover, if $a \in Bb$ ($a, b \in A$), then

$$af_0 \cdot e = ea \in Rea = Rca \subseteq RcBb \subseteq Rcb = Reb = Rbf_0 \cdot e,$$

hence $af_0 \in R \cdot bf_0$. By Prop. 0.8.8 there exists an injective homomorphism $f: B \rightarrow R$ extending f_0 , and this proves (1).

Suppose now that R satisfies strong DFL. Let $q \in B$, say $q = ab^{-1}$. Then $a = qb$, so $af = qf \cdot bf$, and hence

$$qf \cdot eb = qf \cdot bf \cdot e = af \cdot e = ea \in eR.$$

By strong DFL, $qf \cdot e \in eR$, say $qf \cdot e = e \cdot r$. So we have

$$erb = qf \cdot eb = ea,$$

therefore $rb = a$. Thus $a = qb$ implies $a = rb$ for some $r \in R$, and so by Prop. 0.8.8 (with $f = \text{identity}$) we find that $q \mapsto r$ is a well-defined homomorphism $f: B \rightarrow R$ such that $af = a$ for all $a \in A$, as claimed. ■

COROLLARY 1.2. *In a 2-fir with strong DFL every maximal commutative subring is integrally closed in its field of fractions.*

Proof. Let R be a 2-fir with strong DFL, A a maximal commutative subring and B a finite integral extension of A in its field of fractions. By Prop. 1.1, there is an injection $f: B \rightarrow R$ which reduces to the identity on A . But B is commutative, so by the maximality of A we have $B = A$. ■

COROLLARY 1.3. *The centre of a 2-fir is integrally closed in its field of fractions.*

Proof. Let R be a 2-fir, denote its centre by C and suppose that B is a finite integral extension of C in its field of fractions. By Prop. 1.1 there exists an embedding $f: B \rightarrow R$ such that $af = a$ for all $a \in C$ (because C is the centre), so we may assume that $C \subseteq B \subseteq R$. Let $b \in B$, say $a = cb$ ($a, c \in C$), then for any $r \in R$,

$$cbr = ar = ra = rcb = crb,$$

hence $br = rb$, so $b \in C$. This shows that $C = B$, so C is integrally closed in its field of fractions. ■

To obtain further information on the centres of atomic 2-firs we consider the set $I = I(R)$ of invariant elements of R , i.e. regular elements c such that $cR = Rc$. We begin by noting conditions under which invariant elements occur in general rings:

PROPOSITION 1.4. (i) *In any ring R , an element c is invariant if and only if c is regular and the left and right ideals of R generated by c are both two-sided.*

(ii) *If R is an integral domain, then any non-zero ideal that is principal both as left and as right ideal has an invariant generator.*

(iii) *If R is a 2-fir, then the ideals with invariant generators form a sublattice of the lattice of all ideals.*

Proof. (i) is clear; for (ii) take $aR = Ra'$ and let $a = ua'$, $a' = av$, then $a' = ua'v$. Now $a'v \in Ra'$, say $a'v = wa'$, so $a' = uwa'$, i.e. $uw = 1$, hence u is a unit and $aR = Ra$.

(iii) Let $aR = Ra$, $bR = Rb$ be two non-zero ideals in a 2-fir R ; then $0 \neq ab \in aR \cap bR$, hence $aR + bR = dR = Ra + Rb = Rd'$ and $aR \cap bR = mR = Ra \cap Rb = Rm'$. By (ii), d and m are invariant, as claimed. ■

Prop. 1.4 shows that in any 2-fir R , the monoid I of invariant elements is lattice-ordered by divisibility. Thus any two invariant elements have an HCF and this is the same whether calculated in I or in R ; similarly for the LCM.

We now add atomicity to our assumptions; we recall that an I -atom is an invariant element which is an atom within I . By applying Th. 3.1.3 to an atomic 2-fir we obtain a factorization theorem for invariant elements:

THEOREM 1.5. *Every atomic 2-fir has unique factorization of invariant elements, i.e. every non-unit invariant element c can be written as a product of I -atoms*

$$c = a_1 \dots a_r \quad (2)$$

and if

$$c = b_1 \dots b_s \quad (3)$$

is any other factorization of c into I -atoms, then $r = s$ and there is a permutation $i \mapsto j$ of $1, \dots, r$ such that b_{j_i} is associated to a_i . Moreover, any order of the factors in (2) can be realized. ■

In terms of ideals Th. 1.5 states that in an atomic 2-fir the ideals with invariant generator form a free commutative monoid under ideal multiplication. Of course, this is merely a reflexion of the fact that a lattice-ordered group with descending chain condition (on the positive elements) is necessarily free commutative (Birkhoff [67], p. 298).

In certain cases the last result can be extended to matrix rings. Let us call an $n \times n$ matrix C over a ring R *right invariant* if it is right invariant as element of R_n , thus C is regular and for each $X \in R_n$ there exists X' such that $XC = CX'$. We remark that over any Sylvester domain a right invariant matrix must be full. For if $C = PQ$, where P is $n \times r$, Q is $r \times n$ and r is as small as possible, suppose that $r < n$, then by Th. 5.4.9, P contains a full $r \times r$ submatrix, say in the first r rows, so there exists $P' \in {}^nR^{n-r}$ such that $P_1 = (P, P')$ is full and so right regular. Put $Q_1 = \begin{pmatrix} Q \\ 0 \end{pmatrix}$, then $C = P_1 Q_1$ and since C is right invariant, there exists $P_2 \in R_n$ such that $P_1 C = C P_2 = P_1 Q_1 P_2$, hence $C = Q_1 P_2$. But Q_1 is not left regular, so neither is C , a contradiction. This shows that C must be full. Over a Sylvester domain the right invariant matrices take the following form:

PROPOSITION 1.6. *Let R be a Sylvester domain. Then any right invariant matrix*

C over R has the form

$$C = \gamma U, \quad (4)$$

where γ is right invariant in R and U is an invertible matrix.

Proof. By the remark just made, C is full. Let e_{11} be the usual matrix unit; by right invariance there exists $E_{11} \in R_n$ such that

$$e_{11}C = CE_{11}. \quad (5)$$

Hence $e_{11}C = e_{11}^2C = CE_{11}^2$, so $E_{11}^2 = E_{11}$, and by comparing inner ranks in (5) we see that $\rho(E_{11}) = 1$. By Prop. 0.4.5 there is an invertible matrix V such that $V^{-1}E_{11}V = e_{11}$. By (5) we obtain $e_{11}CV = CVe_{11}$ and clearly CV is again right invariant. Hence $CV = c_1 \oplus C'$, where C' is $(n-1) \times (n-1)$ and c_1, C' are again right invariant. By induction on n we find that C is right associated to $D = \text{diag}(c_1, \dots, c_n)$. Now for each permutation matrix P we have $PD = DP$, hence $c_i \in c_jR$ for all i, j , thus all the c_i are right associated and so $D = \gamma W$, where W is invertible, hence $C = \gamma U$, where U is invertible and $\gamma \in R$ is right invariant. ■

In any ring R , if c is left or right invariant, then so is any right associate of c , hence we obtain

COROLLARY 1.7. *If, in Prop. 1.6, C is invariant, then it has the form (4), where γ is invariant. ■*

Combining this result with Th. 1.5, we obtain

THEOREM 1.8. *Over a fir R every invariant matrix can be written as a product of I -atoms in R and an invertible matrix over R , and the factorization is unique up to associates and the order of the factors. ■*

Let us return to the case of elements (Th. 1.5). In order to apply this result to study the centre of a 2-fir we need to recall some facts on valuations. Let A be a commutative integral domain and k its field of fractions; any homomorphism $v: A^\times \rightarrow \Gamma_+$ into the positive cone of a totally ordered additive group Γ (together with the convention $v(0) = \infty$), such that

$$v(a-b) \geq \min\{v(a), v(b)\},$$

is called a *general valuation*. Such a valuation can always be extended in just one way to a valuation of k , again written v , and the set $k_v = \{x \in k \mid v(x) \geq 0\}$ is a local Bezout domain containing A ; k_v is called the *valuation ring* of v .

Now the ring $\bar{A} = \bigcap k_v$, where v ranges over all general valuations on A , consists of those elements of k that are integral over A ; thus $\bar{A}$ is just the *integral closure* of A in k , in fact A is integrally closed precisely if $\bar{A} = A$ (cf. e.g. Cohn [77], p. 322).

There is a related construction which we shall need here. This arises if instead of general valuations we limit ourselves to $\mathbf{Z}$ -valued valuations. Let A be a commutative integral domain and k its field of fractions, as before. If there is a family V of $\mathbf{Z}$ -valued valuations on k such that (i) for any $x \in A^\times$, $v(x) \geq 0$ for all $v \in V$, with equality for almost all v and (ii) $A = \bigcap k_v$, then A is said to be a *Krull domain*. More generally, suppose that there is a family V of $\mathbf{Z}$ -valued valuations on A satisfying (i), then $\bar{A} = \bigcap k_v$ is clearly a Krull domain containing A .

From our earlier remarks it is clear that every Krull domain is integrally closed. Every commutative UFD is a Krull domain: we take V to be the family of valuations associated with the atoms of A . Likewise, every Noetherian integrally closed domain is a Krull domain; here V is the class of valuations associated with the minimal prime ideals of A (cf. Bourbaki [72]).

We now have the following sharpening of Cor. 1.3.

THEOREM 1.9. *The centre of an atomic 2-fir is a Krull domain.*

Proof. Let R be an atomic 2-fir, C its centre and k the field of fractions of C . By Th. 1.5, each $a \in C^\times$ has a decomposition into I -atoms

$$a = u \prod p^{\alpha_p} \quad (u \in \mathbf{U}(R), \alpha_p \geq 0).$$

Fix the I -atom p and consider the function v_p defined on C by

$$v_p(a) = \alpha_p.$$

Clearly this is $\mathbf{Z}$ -valued, in fact it is non-negative on C , and $v_p(a) = 0$ for almost all p . Hence $\bar{C} = \bigcap k_v$ is a Krull domain containing C ; we claim that $\bar{C} = C$. Let $d \in \bar{C}$, say $d = ab^{-1}$, where $a = u \prod p^{\alpha_p}$, $b = v \prod p^{\beta_p}$, $u, v \in \mathbf{U}(R)$ and $\alpha_p \geq \beta_p$. Since $a = db$, we have $u \prod p^{\alpha_p} = dv \prod p^{\beta_p}$; if $\beta_p > 0$, we can cancel p , replacing α_p by $\alpha_p - 1$ and β_p by $\beta_p - 1$. After finitely many steps we find $u_1 \prod p^{\alpha_p - \beta_p} = dv_1$, where u_1, v_1 are units, possibly different from u, v , because the p 's are merely invariant and not central. It follows that $d \in R$; now we find as in the proof of Cor. 1.3 that $d \in C$, i.e. $\bar{C} = C$. ■

As we shall see in 6.2, this theorem is best possible, in the sense that any Krull domain can occur as the centre of an atomic 2-fir. It follows in particular that any fir has a Krull domain as centre, but in 6.3 we shall see that the centres of non-Ore firs are much more restricted.

Exercises 6.1

1°. Given integral domains $A \subseteq B$, an element $y \in B$ is called *left integral* over A if the left A -module generated by the powers of y is finitely generated over A . Show that this is so if and only if y satisfies a monic equation with left coefficients in A . If every element of B is left integral over A , B is called *left integral* over A . Is the notion of left integral extension transitive?

2. A ring B is called a *finite left integral extension* of a subring A if B is finitely generated as left A -module, and similarly on the right. Show that if B is a finite left and right integral extension of A , then there exist $u_1, \dots, u_n \in B$ such that $B = \sum A u_i = \sum u_i A$.

3. Let A be a right Ore domain with field of fractions K , and let B be an A -subring of K . Define the *conductor* of B in A as $c = \{c \in A \mid Bc \subseteq A\}$ and show that $c \neq 0$ if and only if B is finite right integral over A . Show that c is the largest ideal in A which is also a left ideal in B .

4. Let $B = k\langle x, y \rangle$ and denote by A the subalgebra generated by x^2 and y^2 . Find the set $\bar{A}$ of elements of B left integral over A . Is this set closed under addition or multiplication? Is $\sum A a^i$ a subring for every $a \in \bar{A}$?

5. Let R be an atomic 2-fir with DFL. Show that if A is a (left and right) Ore subring of R and B is a finite left extension of A in its field of fractions, then there is an embedding of B in R whose restriction to A is the identity map. Deduce that a maximal Ore subring of R is integrally closed.

6. Let I be the set of all right invariant elements in an integral domain R and S the set of all left factors of elements of I . Show that S is a right Ore set.

7. Let α be an automorphism of a field K such that no power of α is inner. Show that the monic invariant elements of $K[x; \alpha]$ are the powers of x . If α^r ($r > 0$) is inner, but no lower power, find all monic invariant elements (cf. 8.3).

8. Show that a principal ideal domain is simple if and only if it has no non-unit invariant element.

9. Let K be a field of characteristic 0 and D an outer derivation of K . Show that $K[x; 1, D]$ is simple and hence has no non-unit invariant elements.

10. Let R be a 2-fir and c an invariant element. Find the condition on c for R/cR to be (i) simple and (ii) semisimple.

11. Let R be an atomic 2-fir and c a non-unit invariant element. Show that the Artinian ring $A = R/cR$ has the property that every left (or right) ideal is the annihilator of its annihilator in A (i.e. A is a *quasi-Frobenius ring*).

12°. Is every Artinian principal ideal ring a quasi-Frobenius ring?

13. Let R be a 2-fir; show that the elements $\alpha\beta^{-1}$ ($\alpha, \beta \in R, \beta \neq 0$) satisfying $\alpha\beta = \beta\alpha$ for all $x \in R$ form an integrally closed ring.

14*. Find a ring in which the multiplication of invariant principal ideals is non-commutative. (Hint: Use a non-commutative analogue of Ex. 1.1.8.)

6.2 2-Firs with prescribed centre

This section is devoted to proving the converse of Th. 1.9, in the following strong form: every integrally closed domain (resp. Krull domain) occurs as the centre of some Bezout domain (resp. principal ideal domain). Since every Bezout domain is a 2-fir, and every principal ideal domain an atomic 2-fir, this, with Th. 1.9, completely characterizes the centres of (atomic) 2-firs.

The proof proceeds in two stages.

(i) Given an integrally closed commutative domain C , we construct a commutative Bezout domain A with an automorphism of infinite order whose fixed ring is C ; moreover, when C is a Krull domain, A can actually be chosen to be a principal ideal domain.

(ii) Given a commutative Bezout domain A with an automorphism α of infinite order, we construct a Bezout domain containing A , whose centre is precisely the fixed ring of α . Moreover, when A is a principal ideal domain, the ring containing it can be chosen to be principal.

It is convenient to begin with (ii). The two cases considered require rather different treatment, and we therefore take them separately.

PROPOSITION 2.1. *Let A be a commutative Bezout domain with an automorphism α of infinite order, and denote by C the fixed ring of α acting on A . Then there exists a Bezout domain containing A , with centre C .*

Proof. Let K be the field of fractions of A ; then α extends in a unique way to an automorphism of K , again denoted by α (cf. Th. 0.8.11). Now form the skew polynomial ring $K[x; \alpha]$ and let R be the subring of all polynomials with constant term in A , thus $R = A + xK[x; \alpha]$. That R is a subring is clear; we claim that it has the desired properties. In the first place, if $f = \sum x^i a_i$ lies in the centre, then for any $b \in A$, $bf = fb$, and on equating coefficients we find that $a_i b = a_i b^{\alpha^i}$. Now for each $i > 0$ there exists $b \in A$ such that $b^{\alpha^i} \neq b$, hence $a_i = 0$ for $i > 0$ and $f = a_0$. Further, the equation $fx = xf$ shows that $a_0^\alpha = a_0$, so $f = a_0 \in C$, as claimed.

It remains to show that R is Bezout. Let $f, g \in R$; we must show that $fR + gR$ is principal, and here we may assume that f, g have no common left factor of positive degree in x , for if $f = df_1$, $g = dg_1$ and $f_1R + g_1R = hR$ has been established, then $fR + gR = dhR$. By looking at the highest common left factor of f, g in $K[x; \alpha]$ we find polynomials u, v in the latter ring such that $fu - gv = 1$. On multiplying up by a suitable element of A we obtain an

equation

$$fu - gv = \gamma, \quad \text{where } u, v \in R, \gamma \in A^\times. \quad (1)$$

If the constant terms of f, g are λ, μ , respectively, say $f = \lambda + f_1$, $g = \mu + g_1$, where f_1, g_1 have zero constant term, then $\lambda = f - \gamma(\gamma^{-1}f_1) \in fR + gR$ by (1), and similarly $\mu \in fR + gR$, hence the highest common factor δ , say, of λ and μ (in A) divides γ . It follows that δ is a left factor of f and g , and since $\delta \in \lambda R + \mu R$, we have $fR + gR = \delta R$. Thus any two elements generate a principal right ideal; by symmetry they also generate a principal left ideal and so R is a Bezout domain, as claimed. ■

By combining this result with Cor. 1.3 we get

COROLLARY 2.2. *The fixed ring of an automorphism of infinite order acting on a commutative Bezout domain is integrally closed.* ■

Of course, this result can also be obtained directly.

PROPOSITION 2.3. *Let A be a commutative principal ideal domain with an automorphism α of infinite order, and let C be the fixed ring of α acting on A . Then there is a principal ideal domain containing A , with centre C .*

Proof. We form the ring $R = A((x; \alpha))$ of skew formal Laurent series $f = \sum x^i a_i$, where $a_i \in A$ and $a_i = 0$ for i less than some k depending on f , with the commutation rule: $ax = \alpha(a)x$. The verification that the centre of R is C is as before: the equations $bf = fb$ show that $f = a_0$ and now $fx = xf$ shows that $a_0 \in C$. It remains to show that R is a principal ideal domain.

Given $f = \sum x^i a_i \neq 0$, let a_k be the first non-zero coefficient; thus k is the order and a_k the leading coefficient of f . Given $f, g \in R$ we must show that $fR + gR$ is principal. If f or g is 0, this is clear; otherwise, on multiplying f, g by suitable powers of x we may take them both to be of order 0, since x is a unit in R . Denote their leading coefficients by a_0, b_0 , respectively. If b_0 is a unit, then so is g , because $g = (1 + \sum_{i=1}^{\infty} x^i b_i b_0^{-1}) b_0$ can be inverted. We may therefore use induction on the length of b_0 , as an element of A . Let us put $a_0 A + b_0 A = d_0 A$; there exist $\alpha'_0, b'_0, u, v \in A$ such that

$$\alpha'_0 u - b'_0 v = 1, \quad (2)$$

and $a_0 = d_0 \alpha'_0$, $b_0 = d_0 b'_0$. On replacing f, g by $f^* = f b'_0 - g \alpha'_0$, $g^* = fu - gv$, we have $fR + gR = f^*R + g^*R$, because the matrix of transformation from f, g to f^*, g^* lies in A and has determinant 1, by (2). Moreover, the leading coefficient of g^* is d_0 and this has shorter length than b_0 unless $b_0 | a_0$. By our

induction hypothesis we may therefore assume the latter, say $a_0 = cb_0$ ($c \in A$). Then $f^* = f - gc$ has higher order than f . If by suitable choice of h , the order of $f - gh$ can be made arbitrarily large, then there is also an h such that $f - gh = 0$, and hence $fR + gR = gR$. Otherwise let us choose h so that $f_1 = f - gh$ has the largest possible order; then its leading coefficient will not be divisible by that of g and by repeating the process (of forming f^*, g^*) we can replace g by an element with a leading coefficient of shorter length. This shows that R is right Bezout, and by symmetry R is left Bezout. Now in any factorization of an element $f \neq 0$ the number of non-unit factors is bounded by the length of the leading coefficient of f , hence R is atomic, and so is a principal ideal domain. ■

Note that in the proof we have had to use the fact that A is principal even to show that R is Bezout. In fact, if we perform the construction with a Bezout domain A , we do not generally get a Bezout domain. Using Th. 1.9, we again have a corollary (which, as before, can be proved directly).

COROLLARY 2.4. *The fixed ring of an automorphism of infinite order acting on a commutative principal ideal domain is a Krull domain.* ■

We now come to step (i) of our programme. This is in effect the converse of Cors. 2.2 and 2.4.

PROPOSITION 2.5. *Every integrally closed commutative integral domain occurs as the fixed ring of an automorphism of infinite order acting on a commutative Bezout domain.*

Proof. We first give the basic construction which will in most cases produce the required ring, and then show how to modify it to get the result in all cases.

Let C be the given domain and k its field of fractions. By hypothesis, $C = \bigcap k_v$, where v ranges over the family V of all general valuations defined on C . We form the polynomial ring $k[t]$ in an indeterminate t . Each $v \in V$ can be extended to $k[t]$ by putting

$$v(\sum t^i a_i) = \min_i \{v(a_i)\}.$$

We assert that this is again a valuation. The rule $v(a-b) \geq \min\{v(a), v(b)\}$ is clear; to show that v is multiplicative, let $a = \sum t^i a_i$, $b = \sum t^j b_j$ and let a_r, b_s be the first coefficient attaining the minimum $v(a), v(b)$, respectively. The product $c = ab$ has coefficients $c_k = \sum a_i b_{k-i}$ and $v(c_k) \geq v(a_r) + v(b_s)$, with equality holding for $k = r + s$, as is easily verified. Thus v is a valuation on $k[t]$ and in fact can be extended to a valuation, again denoted by v , of the field of

fractions $k(t)$. We now define

$$A = \bigcap k(t)_v, \quad \text{where } v \text{ ranges over } V.$$

Thus A consists of all fractions f/g , $f, g \in C[t]$, such that $v(f) \geq v(g)$. A is sometimes called the *Kronecker functional ring*.

We claim that A is a Bezout domain: given two elements of A , on multiplying by a common denominator, we may take them to be $f, g \in C[t]$. Now take any n greater than the degree of f in t and form $h = f + t^n g$. Then clearly $v(h) \leq v(f), v(g)$, hence $f/h, g/h \in A$ and so $fA + gA = hA$.

Consider the map α of $k[t]$ defined by $f(t) \mapsto f(t+1)$. This is an automorphism, clearly $v(f(t+1)) \geq v(f(t))$ for any $v \in V$, and by writing down the same inequality with -1 in place of 1 , we see that $v(f^\alpha) = v(f)$, hence α extends to an automorphism of A . We observe that α is of infinite order precisely when k has characteristic 0. In that case the fixed field of α acting on $k(t)$ is k , hence the fixed ring of α acting on A is $k \cap A = C$, as required.

In the case of finite characteristic we modify our construction by starting, not with $k[t]$, but with $k[\dots, t_{-1}, t_0, t_1, \dots]$. As before, we get a Bezout domain, and for our automorphism α we can use instead of the translation $t \mapsto t+1$ the substitution $t_n \mapsto t_{n+1}$. Clearly this is of infinite order and the fixed ring in the action on $k(\dots, t_{-1}, t_0, t_1, \dots)$ is k , hence the fixed ring of A is C . Thus A is a Bezout domain with an automorphism of infinite order whose fixed ring is C . ■

Suppose now that C is a Krull domain. Let k be its field of fractions and V the family of valuations defining C . We form A as in the proof of Prop. 2.5, using the family V instead of the family of all valuations. Then it follows as before that A is a Bezout domain with fixed ring C . We claim that A is, in fact, a principal ideal domain. Given $a \in A$, it is clear that $v(a) \neq 0$ for only finitely many $v \in V$, say $v_1, \dots, v_r$. Now for any factor b of a , $0 \leq v(b) \leq v(a)$ for all $v \in V$, and if $v(b) = v(b')$ for all $v \in V$, then b and b' are associated. Hence there are only finitely many classes of factors of a , therefore A is atomic and so is a principal ideal domain. This proves

PROPOSITION 2.6. *Every Krull domain occurs as the fixed ring of an automorphism of infinite order acting on a commutative principal ideal domain.* ■

Putting all the results of this section together, we obtain

THEOREM 2.7. *Every integrally closed commutative integral domain occurs as the centre of a Bezout domain; every Krull domain occurs as the centre of a principal ideal domain.* ■

Exercises 6.2

1. Give a direct proof (e.g. by valuation theory) that the fixed ring of an automorphism acting on a commutative Bezout domain R is integrally closed, and is a Krull domain when R is a principal ideal domain.

2*. Does Prop. 2.1 still hold when the automorphism has finite order?

3. Examine why the proof of Prop. 2.1 fails for principal ideal domains and that of Prop. 2.3 fails for Bezout domains.

4. In the proof of Prop. 2.5, if k has finite characteristic, complete the argument by taking in place of $k[t]$ the ring $k[t, t^{1/2}, t^{1/4}, \dots]$ with automorphism $t \mapsto t^2$.

5. Verify that for any commutative field k , the polynomial ring $k[x_1, x_2, \dots]$ is a Krull domain, but not Noetherian. Deduce the existence of a Noetherian domain with a non-Noetherian centre.

6. Show that the centre of a 2-fir with right ACC_1 is a Krull domain.

6.3 The centre of a fir

The results of 6.2 give a complete description of the possible centres of principal ideal domains or Bezout domains, as well as some information on the centres of 2-firs, but they leave open the question whether, for example, any Krull domain can occur as the centre of a genuine, i.e. non-Ore, fir. As we shall see, once we assume that our rings are non-Ore, the centre is very much more restricted. Thus the centre of a non-Ore fir is necessarily a field. More generally, the conclusion will hold for any non-Ore 3-fir with right ACC_2 .

We begin by looking at the quotient of a 2-fir by an ideal with an invariant generator; this has a rather special form, which is described in

THEOREM 3.1. *Let R be a 2-fir and c a non-unit right invariant element. Then cR is an ideal in R and R/cR is a right Bezout ring (not necessarily a domain) which is a field if and only if c is an atom.*

If, moreover, c has finite length, then R/cR is Artinian and if, further, cR is maximal among ideals with a right invariant generator, then $R/cR \cong \mathfrak{M}_n(K)$, for some $n \geq 1$, where K is a field.

Proof. Since c is right large and R is a 2-fir, every finitely generated right ideal containing cR is principal, hence of the form dR , where d is a left factor of c .

Hence R/cR is right Bezout and is a field precisely if c is an atom. This also shows that R/cR is Artinian if c has finite length.

Now let cR be maximal among ideals with right invariant generator, write $S = R/cR$ and let α be a minimal right ideal of S . Then $S\alpha$ is two-sided, hence of the form dR/cR , where dR is two-sided, so d is right invariant and by the maximality of cR we have $dR = R$, i.e. $S\alpha = S$. Thus S is a sum of copies of the simple right S -module α , hence a direct sum, say $S \cong \alpha^n$ and so

$$S^o \cong \text{End}_S(\alpha^n) \cong \mathfrak{M}_n(L),$$

where $L = \text{End}_S(\alpha)$ is a field, by Schur's lemma. Writing $K = L^o$, we have $S \cong \mathfrak{M}_n(K)$, as claimed. ■

Later (in Th. 4.3) we shall see that when c is an I -atom, then R/cR is simple Artinian.

In the proof of the main result we shall need the following technical lemma.

LEMMA 3.2. *Let R be a 3-fir and let c be an invariant element of R . If α is a right ideal which is maximal among the free right ideals of rank 2 of R , then $cR \cap \alpha = \alpha c$.*

Proof. Since c is invariant, cR is right large and any finitely generated right ideal containing cR must be principal (as in Th. 3.1), therefore $\alpha + cR$ is principal, say $\alpha + cR = dR$. We thus have the exact sequence

$$0 \rightarrow \alpha \cap cR \rightarrow \alpha \oplus cR \rightarrow \alpha + cR \rightarrow 0. \quad (1)$$

From the invariance of c it follows that $\alpha \cap cR$ can be written as $\alpha'c$, where $\alpha' = \{x \in R \mid xc \in \alpha\}$ is a right ideal containing α . Since dR is free, (1) splits and because we are in a 3-fir, we can conclude that $\alpha'c = \alpha \cap cR$ is free of rank 2, hence so is α' . But α was maximal for this property, therefore $\alpha' = \alpha$. ■

In the applications we obtain maximal free right ideals of rank 2 by imposing right ACC_2 . There is a corresponding result for maximal non-finitely generated right ideals in any ring; here the existence of such ideals follows by Zorn's lemma (as in the proof of I. S. Cohen's theorem, that any maximal non-finitely generated ideal in a commutative ring is prime, cf. Kaplansky [70], p. 5).

We now come to the main result of this section.

THEOREM 3.3. *Let R be a 3-fir with right ACC_2 , but not a right Ore domain. Then every invariant element of R is a unit.*

Proof. Since R is a domain, but not right Ore, there exist free right ideals of

rank 2. By ACC_2 we can find a right ideal a which is maximal free of rank 2. Let c be any invariant element of R . By Lemma 3.2, $a \cap cR = ac$ and $a + cR = dR$ for some $d \in R$. Hence if $a = pR + qR$, then

$$dR/cR \cong (pR + qR)/(pcR + qcR) \cong (R/cR)^2. \quad (2)$$

Now c is determined by R/cR up to associates (Lemma 4.2.6); by (2), $(R/cR)^2$ is cyclic, which can only happen when c is left comaximal with itself (Prop. 3.6.8), hence c must be a unit. ■

Recalling that any right fir satisfies right ACC_n for all n and if it satisfies the right Ore condition, it must be a principal right ideal domain, we find

COROLLARY 3.4. *Let R be a right fir which is not a principal right ideal domain. Then every invariant element of R is a unit, in particular, the centre of R is then a field. ■*

Finally, we record, for later use, the form taken by the centre of the field of fractions of a principal ideal domain.

PROPOSITION 3.5. *Let R be a principal ideal domain, K its field of fractions and C the centre of K . Then C consists of all elements of the form ab^{-1} , $a, b \in R$, such that $b \neq 0$ and*

$$axb = bxa, \quad \text{for all } x \in R. \quad (3)$$

Proof. If $ab^{-1} \in C$, then $ab^{-1}y = yab^{-1}$ for all $y \in K$, hence on writing $y = bx$ we obtain (3). Conversely, when (3) holds, then $ab = ba$, hence in K , $ab^{-1} = b^{-1}a$ centralizes R and so also K . ■

Exercises 6.3

1. Let R be any ring, a a right ideal maximal among non-finitely generated right ideals of R and let c be an invariant element not in a . Show that $a \cap cR = ac$.

2. (Goursaud and Valette [75]) Show that if R is right hereditary and c an invariant element of R , then R/cR is right Noetherian. (Hint: Use a dual basis.)

3. Let R be a right hereditary integral domain, but not right Ore. Show that R has no non-unit invariant elements.

4. Show that the subalgebra of $\mathbb{Q}\langle x, y \rangle$ consisting of elements in which the constant term has odd denominator is a non-Ore semifir with non-unit central elements.

5°. Determine the possible centres of 2-firs with right ACC_1 .

6. Show that the subalgebra of $k\langle x, y, y^{-1} \rangle$ generated by $x, y, y^{-n}x$ is a non-Ore right fir with right invariant non-unit y (cf. 2.10).

7. Let K be a field with a non-surjective endomorphism α . Show that $K[x; \alpha]$ is a fir which is not left Ore and has the right invariant element x .

8. Show that the subalgebra of $k\langle x, y, y^{-1} \rangle$ generated by y and $y^{-m}xy^{-n}$ ($m, n \geq 0$) is a semifir with a non-unit invariant element y .

9°. Does there exist a non-principal fir with a non-unit right invariant element?

6.4 Bounded elements in 2-firs

Let us return to consider invariant elements in a 2-fir R . We have seen that in an atomic 2-fir the decompositions into I -atoms play an important role. In general there is no reason to suppose that I -atoms will be atoms, but at least we can factorize them into atoms, and this suggests that we look more closely at the factors of invariant elements. Such factors and the cyclic modules they define are called *bounded* (to be distinguished from bound modules, cf. Ex. 1). However, it will be convenient to begin with a general definition and specialize later.

Definition. A module M over a ring R is said to be *bounded* if there is a regular element c in R such that $Mc = 0$.

An equivalent definition is to require the annihilator of M in R to contain a regular element. This annihilator is a two-sided ideal in R , called the *bound* of M . An element a of R is said to be *right bounded* if R/aR is bounded. We remark that any right bounded element is right large, for if a is right bounded, say $Rd \subseteq aR$, and $b \in R^\times$ is given, then $bd \in aR$, so $bR \cap aR \neq 0$.

To describe the right bound of $a \in R$ more closely, let $a = bc$, then for any $d \in \text{Ann}(R/aR)$ we have $Rd \subseteq aR = bcR$; in particular, $bd \in bcR$, so $d \in cR$. The same is true if instead of a we take any element similar to a . Thus

$$\text{Ann}(R/aR) \subseteq \bigcap \{cR \mid c \text{ right factor of an element similar to } a\}. \quad (1)$$

In a 2-fir we can, by Prop. 3.5.3, rewrite this as

$$\text{Ann}(R/aR) \subseteq \bigcap \{cR \mid c \text{ similar to a right factor of } a\}. \quad (2)$$

Moreover, in this case we can prove that equality holds in (2).

THEOREM 4.1. *Let R be a 2-fir. Then for any right large element a ,*

$$\text{Ann}(R/aR) = \cap \{cR \mid c \text{ similar to a right factor of } a\}. \quad (3)$$

In particular, this holds for any right bounded element a .

Proof. We have seen that (2) holds for any right bounded element a ; if a is not right bounded, then the left-hand side of (2) is zero, so (2) holds for all $a \in R$. Now assume that a is right large and choose u in the right-hand side of (2). Given $x \in R^\times$, we have $xR \cap aR \neq 0$, hence there is a right coprime relation

$$xa' = ax',$$

and here a' is similar to a right factor of a , therefore $u = a'u_1$ for some $u_1 \in R$. Hence $xu = xa'u_1 = ax'u_1 \in aR$, so u annihilates $x \pmod{aR}$, and since x was arbitrary, we see that $u \in \text{Ann}(R/aR)$, i.e. (3) holds. The result holds for every right bounded a , since such an a is right large. ■

Suppose now that R is an atomic 2-fir and that a is right bounded. Then all the cR on the right of (3) contain a fixed non-zero element, d say, where $d \in \text{Ann}(R/aR)$, and since the lattice $L(dR, R)$ has finite length, it is complete and the intersection on the right of (3) is principal, say $\text{Ann}(R/aR) = dR$. The element d , unique up to right associates, is called the *right bound* of a .

COROLLARY 4.2. *In an atomic 2-fir R , an element a is right bounded if and only if a is a left factor of a right invariant element d . In that case the right bound of a is a left factor of d .*

Proof. We have seen that for a right bounded element a of R , $\text{Ann}(R/aR) = dR$. Here $Rd \subseteq dR$, so d is right invariant, and $d \in aR$. Conversely, if $d = ab$ for a right invariant element d , then $Rd \subseteq dR \subseteq aR$, hence $d \in \text{Ann}(R/aR)$, i.e. a is right bounded, with bound $d'R \supseteq dR$. ■

An element a in an integral domain R is said to be *bounded* if it is a left factor of an invariant element: $c = ab$; it is then also a right factor, for we have $cb = b'c = b'ab$, hence $c = b'a$. Thus the notion defined here is symmetric. A bounded element is clearly left and right bounded, and by Prop. 1.4 we have

$$\text{Ann}(R/aR) = \text{Ann}(R/Ra). \quad (4)$$

Conversely, if a is left and right bounded and (4) holds, then $\text{Ann}(R/aR) = dR = Rc$, hence by Prop. 1.4, $\text{Ann}(R/aR)$ has an invariant generator and a is then bounded.

Let R be an atomic 2-fir and let $a \in R$ be bounded; the invariant generator of $\text{Ann}(R/aR)$, a^* say, is then unique up to associates and is called the *bound* of a . Clearly a^* depends only on the similarity class of a . By Th. 4.1, a^* can also be defined by

$$a^*R = \bigcap \{cR \mid c \text{ similar to a right factor of } a\}. \quad (5)$$

The right-hand side can be taken to be a finite intersection, by the DCC in R/a^*R , and if we take it to be irredundant, we obtain a subdirect sum representation of R/a^*R , *qua* right R -module, by modules R/dR . This shows that every atomic factor of a^* is similar to a factor of a . By Th. 4.1 this characterizes the atomic factors of a^* as the atoms similar to factors of a .

Let $a \in R$ have a bound a^* ; if $a = bcd$, then c is again bounded, with bound dividing a^* . This follows by observing that R/cR is a quotient of a submodule of R/aR . Likewise, the product of any bounded elements is bounded. These facts may be expressed by saying that the modules R/aR , where a is bounded, form the objects of a *dense* subcategory of the category $\mathcal{T}_R^1$ of cyclic torsion modules.

To establish the connexion between I -atoms and the atoms they bound we return to the topic of Th. 3.1. In view of its importance in the study of quotients of a principal ideal domain we give a separate proof.

THEOREM 4.3. *Let R be an atomic 2-fir. (i) If c is an invariant element of R , then the quotient ring R/cR is simple if and only if c is an I -atom. (ii) Every I -atom is a product of similar bounded atoms. (iii) If p is a bounded atom, then its bound p^* is an I -atom whose atomic factors are precisely all the atoms similar to p ; moreover, the eigenring K of pR is a field and*

$$R/p^*R \cong K_n, \quad \text{where } n = l(p^*). \quad (6)$$

Proof. We shall prove the parts in reverse order, beginning with (iii). Let p have bound $p^* = a^*b^*$, where a^*, b^* are non-unit invariant elements. Then p divides either a^* or b^* and so has a smaller bound, a contradiction; thus p^* is an I -atom. By Th. 4.1, $p^*R = \bigcap p'R$, where p' runs over all elements similar to p , and here we can take a finite intersection. Thus R/p^*R is a submodule of $(R/pR)^N$ for some N (in fact, p^* is fully reducible, cf. 3.6), and since R/pR is $\mathcal{T}$ -simple, every submodule has the form $(R/pR)^n$ for some $n \leq N$. Hence

$$R/p^*R \cong (R/pR)^n, \quad (7)$$

as right R -modules. By comparing the lengths of composition series within $\mathcal{T}_R$ we see that $n = l(p^*)$, and comparing endomorphism rings in (7), we obtain the isomorphism (6), where K , the eigenring of pR , is a field by Schur's lemma (Prop. 3.3.4).

To prove (ii), let p^* be an I -atom and let p be an atom dividing p^* ; then p^* is the bound of p and, as (5) shows, an atom divides p^* if and only if it is similar to p .

Finally, to prove (i), we see from (6) that for an I -atom c , R/cR is simple. Conversely, if c is invariant but not an I -atom, then R/cR has non-trivial quotients and so cannot be simple. ■

We next look at the direct decompositions of R/aR , where a is bounded. Our first task is to separate out the bounded components in such a decomposition; for this we need a definition and a lemma. Let us call two elements a, b of a 2-fir R *totally coprime* if no non-unit factor of a is similar to a factor of b , i.e. if R/aR and R/bR have no isomorphic factors apart from 0.

LEMMA 4.4. *A bounded element in a 2-fir can be comaximally transposed with any element totally coprime to it.*

Proof. Let a, b be totally coprime in R and let a be bounded, with bound $a^* = a_1a$, say; then a^* and b are left coprime and hence right comaximal: $a^*u - bv = 1$. It follows that $u'a^* - bv = 1$ for some $u' \in R$, thus $u'a_1a - bv = 1$, and so by Lemma 3.4.3, a, b are comaximally transposable. By the symmetry of the situation, b, a are also comaximally transposable. ■

Let a be any element in an atomic 2-fir R and take an atomic factorization

$$a = p_1 p_2 \dots p_r. \quad (8)$$

Suppose that a bounded atom occurs in this factorization, say $p_{i_1}, \dots, p_{i_k}$ are all bounded similar atoms, while the remaining atoms in (8) are not similar to p_{i_1} . By repeated application of Lemma 4.4 we can write a as

$$a = p'_{i_1} \dots p'_{i_k} p'_{j_1} \dots p'_{j_h},$$

where the p_{i_1} but not the p_{j_1} are similar to p_{i_1} . Bracketing the first k factors together, and the last h , we have $a = bc$, where b, c are totally coprime and b , like p_{i_1} , is bounded. Applying Lemma 4.4 again, we have $a = c'b'$, where b' is similar to b and c' similar to c . Therefore $a = bc = c'b'$ leads to a direct decomposition of a :

$$aR = bR \cap c'R.$$

If we repeat the process with a replaced by c' , we eventually reach a direct decomposition of a into products of pairwise similar bounded atoms and a term containing no bounded non-unit factors. Let us call an element *totally*

unbounded if it has no bounded factor apart from units. Then our result may be stated as

THEOREM 4.5. *Any element a of an atomic 2-fir R has a direct decomposition*

$$aR = q_1R \cap \dots \cap q_kR \cap uR, \quad (9)$$

where each q_i is a product of similar bounded atoms, while atoms in different q_i 's are dissimilar and u is totally unbounded. Moreover, the q_i and u are unique up to right associates.

Only the uniqueness still remains to be proved. But (9) gives rise to a direct decomposition

$$R/aR \cong R/q_1R \oplus \dots \oplus R/q_kR \oplus R/uR.$$

Here the R/q_iR are uniquely determined as the homogeneous components corresponding to a given $\mathcal{T}$ -simple bounded isomorphism type, while R/uR contains all $\mathcal{T}$ -simple submodules of unbounded isomorphism type. Moreover, q_iR is unique as the intersection of all $pR \supseteq aR$, where p runs over all atoms bounded by q_i . ■

We note that neither u nor the q_i are in general indecomposable in the sense defined in 3.6.

For an invariant element c we are naturally interested in decompositions of R/cR with $I = \mathbf{I}(R)$. We therefore define an element c of I to be *I-decomposable* if it has a factorization

$$c = ab, \quad (10)$$

into non-unit invariant elements a, b which are left (hence also right) coprime; otherwise c is *I-indecomposable*. There is a simple relationship between *I*-indecomposable elements and bounded elements that are indecomposable as defined in 3.6.

PROPOSITION 4.6. *Let R be an atomic 2-fir. Then*

(i) *an invariant element is I-indecomposable if and only if it is a power of an I-atom and*

(ii) *if $q \in R$ is bounded and indecomposable, then its bound is I-indecomposable, hence q is then a product of similar atoms.*

Proof. (i) follows by Th. 1.5. To prove (ii) let q be bounded, with bound q^* . By Th. 4.5, q is a product of similar atoms, say $q = p_1 \dots p_n$, where the p_i are all similar; hence its bound q^* divides p_1^{*n} , say $q^* = p_1^{*r}$ and this shows q^* to be *I*-indecomposable. ■

Of course, an unbounded indecomposable element need not be a product of similar atoms, as the example xy in the free algebra $k\langle x, y \rangle$ shows. Further, the converse of Prop. 4.6(ii) is false, i.e. a decomposable bounded element need not have an I -decomposable bound; thus in Prop. 4.6(ii), q^* itself may well be decomposable (cf. Ex. 3).

Let q^* be an I -indecomposable element in an atomic 2-fir R , say $q^* = p^{*e}$, $l(p^*) = h$. By Th. 4.3, $R/p^*R \cong K_h$, where the field K is the eigenring of an atomic factor of p^* . Now $R/p^*R \cong Q/J(Q)$, where $Q = R/q^*R$ and $J(Q)$ is the Jacobson radical of Q . Since Q is Artinian (Th. 3.1), we can lift the matrix basis from R/p^*R to R/q^*R (cf. e.g. Cohn [77], p. 350), whence $Q \cong L_h$, where $L/J(L) \cong K$, i.e. Q is an Artinian matrix local ring over the scalar local ring L . Note that L , like Q , is Artinian, hence it is completely primary.

Now take a complete direct decomposition of R/q^*R as right R -module. The summands are necessarily cyclic; thus

$$R/q^*R \cong R/q_1R \oplus \dots \oplus R/q_kR. \quad (11)$$

Since $R/q^*R \cong L_h$ has a complete direct decomposition into h isomorphic right ideals, we see that $k = h$ and all the R/q_iR are isomorphic. Thus $R/q^*R \cong (R/qR)^h$, as right R -modules, for some indecomposable q . Since $l(q^*) = eh = hl(q)$, we see that $l(q) = e$. The result may be summed up as

THEOREM 4.7. *In an atomic 2-fir R , let $q \in R$ be bounded indecomposable, with bound q^* . Express q^* as a power of an I -atom, say $q^* = p^{*e}$. Then $l(q) = e$, and if $l(p^*) = h$, then*

$$R/q^*R \cong (R/qR)^h, \quad (12)$$

as right R -modules, while as a ring, R/q^*R is a total matrix ring over a completely primary ring:

$$R/q^*R \cong L_h, \quad \text{where } L = \text{End}_R(R/qR). \quad \blacksquare \quad (13)$$

If q, q^* are as in Th. 4.7, then q^*R is determined by the similarity class of q as the annihilator of R/qR , while R/qR is determined by q^*R as an indecomposable part of R/q^*R . Hence we have

COROLLARY 4.8. *In an atomic 2-fir R , two bounded indecomposable elements have the same bound if and only if they are similar. $\blacksquare$*

Next we turn to the question of deciding when a given product of similar bounded atoms is indecomposable. Let p^* be an I -atom; for any integer $e \geq 0$ we have by Th. 4.7, on decomposing p^{*e} ,

$$R/p^{*e}R \cong (R/q_eR)^h,$$

for some indecomposable element q_e of length e . Thus q_e is a product of e atomic factors which are all similar.

Conversely, if p is a bounded atom with bound p^* and p_i is similar to p for $i = 1, \dots, e$, then $q = p_1 \dots p_e$ is bounded by p^{*e} , and if its exact bound is p^{*e} , then q is indecomposable. For if q could be decomposed, it would have a smaller bound. This proves

PROPOSITION 4.9. *In an atomic 2-fir R , each bounded indecomposable element is a product of similar atoms. Let p be a bounded atom and p^* its bound. Then a product $q = p_1 \dots p_e$ of atoms similar to p is indecomposable if and only if p^{*e} is the exact bound of q . Moreover, for any e , there is a bounded indecomposable element q_e of length e such that*

$$R/p^{*e}R \cong (R/q_e R)^h, \quad \text{where } h = l(p^*). \blacksquare$$

COROLLARY 4.10. *In an atomic 2-fir, if a has the bound p^{*e} , where p^* is an I -atom, then a direct decomposition of a has at most $l(p^*)$ terms. $\blacksquare$*

If an element a has bound p^{*e} and b has bound p^{*f} , then ab is bounded by p^{*e+f} . Applying this remark to a product of similar bounded atoms, we obtain

COROLLARY 4.11. *In an atomic 2-fir, any factor of a bounded indecomposable element is again bounded indecomposable. $\blacksquare$*

Now a product of atoms $p_1 p_2$ is indecomposable if and only if they cannot be comaximally transposed. Hence any product $q = p_1 \dots p_n$ of bounded atoms has no indecomposable factors apart from atoms if and only if no pair of adjacent factors can be comaximally transposed, i.e. if q is rigid. By Prop. 4.9, two similar atoms p_1, p_2 with a common bound p^* are comaximally transposable if and only if $p_1 p_2 \nmid p^*$. Thus we obtain

COROLLARY 4.12. *Let R be an atomic 2-fir and $q = p_1 \dots p_e$ a bounded product of atoms. Then q is rigid if and only if the p_i have a common bound p^* and $p_{i-1} p_i \nmid p^*$ ($i = 2, \dots, e$). In particular, p^e is rigid if and only if $p^2 \nmid p^*$. $\blacksquare$*

There remains the problem of finding which elements are bounded. We shall confine ourselves to the case of an atomic 2-fir R . If p is an atom in R , then $K = \text{End}_R(R/pR)$ is a field and each element a of R defines a K -endomorphism of R/pR by right multiplication. If the natural mapping $R \rightarrow R/pR$ is written $x \mapsto \bar{x}$, then the K -endomorphism defined by a is

$$\rho_a: \bar{x} \mapsto \bar{x}a.$$

We want to find an upper bound for the K -dimension of $\ker \rho_a$. Clearly, if $a = a_1 \dots a_r$, then $\rho_a = \rho_{a_1} \dots \rho_{a_r}$ and by Sylvester's law of nullity for fields,

$$\dim \ker \rho_a \leq \dim \ker \rho_{a_1} + \dots + \dim \ker \rho_{a_r}. \quad (14)$$

We assert that ρ_a is injective when a has no factor similar to p . Thus assume that $\ker \rho_a \neq 0$, then there exists $x \notin pR$ such that $xa \in pR$, say

$$xa = py \quad (x \notin pR). \quad (15)$$

Since p is an atom, (15) is left coprime, hence $a = p'a'$, where p' is similar to p . This shows that ρ_a is injective when a has no factor similar to p .

Suppose now that a is similar to p ; then any element $x_1 \in R$ satisfying $x_1 a = py_1$ for some $y_1 \in R$, while $x_1 \notin pR$, defines an isomorphism $R/aR \rightarrow R/pR$, and any two such isomorphisms differ by an endomorphism of R/pR , i.e. an element of K . Hence $\ker \rho_a$ is 1-dimensional in this case. Going back to (14) we see that $\ker \rho_a$ has a dimension at most equal to the number of factors of a that are similar to p . Thus we obtain

THEOREM 4.13. *Let R be an atomic 2-fir, p an atom and $K = \text{End}_R(R/pR)$. Then for any $a \in R$, the mapping $\rho_a: \bar{x} \mapsto \bar{x}a$ is a K -endomorphism of R/pR and*

$$\dim_K \ker \rho_a \leq m, \quad (16)$$

where m is the number of factors in a complete factorization of a that are similar to p . ■

Suppose now that p is a right large atom of R . Then for any $x \notin pR$ there is a comaximal relation

$$xp' = px';$$

hence there exists p' similar to p and annihilating $\bar{x}$. Conversely, as we have seen, any p' similar to p annihilates some $\bar{x}$, hence when p is right large and q is any atom, then ρ_q is injective if and only if q is not similar to p . We derive the following consequence.

Let p be a right bounded atom, with right bound p^* of length m . Then p is certainly right large, and by (16),

$$\dim_K(R/pR) \leq m. \quad (17)$$

If this inequality were strict, we could find a product of fewer than m factors similar to p which annihilates R/pR , and hence, by Th. 4.1,

$$aR \subseteq \bigcap \{p'R \mid p' \text{ similar to } p\} = p^*R.$$

Thus $a \in p^*R$ and this contradicts the fact that $l(a) < m = l(p^*)$. Hence equality holds in (17); this shows that for a right bounded atom p , R/pR is finite-dimensional (and, of course, p is right large). Conversely, if p is right large and R/pR finite-dimensional, then we can as before find a annihilating R/pR , hence p is then right bounded. Thus we have

COROLLARY 4.14. *Let R be an atomic 2-fir. Then an atom p in R with eigenring K is right bounded if and only if p is right large and $\dim_K(R/pR)$ is finite. Moreover, in this case,*

$$\dim_K(R/pR) = l(p^*),$$

where p^* is the bound of p . ■

Exercises 6.4

1°. Let R be an integral domain. Show that every bounded R -module is bound, but in general the converse is false. For which class of domains does the converse hold (for finitely generated modules)?

2. Let R be a principal ideal domain and $a \in R^*$. If $\bigcap dR \neq 0$, where d runs over all elements similar to right factors of a , show that a is bounded. Show that this no longer holds for firs, by taking $R = k\langle x, y \rangle$, $a = x$.

3. (Jacobson [43]) Let K be a field of finite dimension over its centre and α an automorphism of K such that α^r is inner for some $r \geq 1$. Show that every non-zero element of the skew polynomial ring $K[x; \alpha]$ is bounded; illustrate this by the complex-skew polynomial ring. Give an example of a product of similar bounded atoms which is decomposable in the sense of 3.6.

4. Show that for a bounded element in an atomic 2-fir R , 'left indecomposable' = 'indecomposable'. If q is such an element, show that R/qR has only one composition series.

5°. Does (4) hold for a bounded element in an integral domain?

6. Let R be any ring and c an invariant element of R . Show that for any $a \in R$ the following are equivalent: (i) a is right bounded and a left factor of c , (ii) a is left bounded and a right factor of c and (iii) a is a left and right factor of c .

7. Let R be an atomic semifir and $n \geq 1$. If p is a bounded atom in $S = R_n$, show that its bound p^* is an I -atom of R whose atomic factors in any R_m are precisely the atoms of R_m stably associated to p . If K is the eigenring of p and $r = l(p^*)$, show that K is a field and $S/p^*S \cong K_{rn}$, where $S = R_n$. Deduce that every bounded atom in R_n is stably associated to an element of R .

8. Let R be a fir with infinite centre and put $S = R_n$. Show that any atom P in S

with eigenring K is right bounded if and only if S/PS is simple as right S -module and $\dim_K(S/PS)$ is finite.

9. Let R be an atomic 2-fir and c an invariant element. Show that R/cR is semisimple if and only if c is not divisible by the square of an I -atom.

10. Let R be a ring and for any R -module M define the *tertiary radical* (of 0 in M) as the set of all $x \in R$ which annihilates a large submodule of M . If R is an atomic 2-fir and c is bounded in R , find conditions for R/cR to have a tertiary radical of the form $bR \supseteq cR$. If c is indecomposable, show that $b = p^*$, where p^* is the I -atom corresponding to c .

11. Show that in a principal ideal domain the quotient by any maximal ideal is simple Artinian.

6.5 Free monoids

Before discussing subalgebras of free algebras it is helpful to look at free monoids, where we shall meet the same problems, but in a simplified form, since there is only one operation. It is particularly instructive to see what becomes of the weak algorithm in monoids; as we saw in 2.4, the weak algorithm may be used to characterize free algebras, and below we obtain a corresponding result for monoids.

The free monoid on a set X is denoted by X^* . Each element of X^* may be written as a finite sequence of elements of X —with the empty sequence representing 1—and multiplication consists of juxtaposition. Clearly this representation of the elements is a normal form; moreover, 1 is the only unit of X^* . In fact, we can say a little more: X^* is conical (cf. 3.1). Further, X is the precise set of atoms in X^* . Thus the generating set X is uniquely determined in terms of the monoid structure on X^* . It is this fact that accounts for the simplicity of the theory.

THEOREM 5.1. *A monoid S is free on the set X of its atoms if and only if it is generated by X and is rigid and conical.*

Proof. Let S be free on X ; clearly S is generated by X , it is conical and satisfies cancellation. If $ab' = ba'$, express a, b, a', b' in terms of X , say $a = x_1 \dots x_r$, $b = y_1 \dots y_s$, $a' = y_{s+1} \dots y_n$, $b' = x_{r+1} \dots x_m$, then $x_1 \dots x_m = y_1 \dots y_n$ and, since S is free on X , we have $m = n$ and $y_i = x_i$, so if $r \leq s$ say, then $b = ac$, where $c = x_{r+1} \dots x_s$. Thus S is also rigid.

Conversely, assume that the conditions hold. It will be enough to show that every element of S can be written in just one way as a product of elements of

X . Let $a \in S$, then there is at least one way of so expressing a ; suppose that

$$a = x_1 \dots x_m = y_1 \dots y_n, \quad x_i, y_j \in X. \quad (1)$$

By rigidity, $x_1 = y_1 b$ or $y_1 = x_1 b$ for some $b \in S$, say the former holds. Since x_1, y_1 are atoms, b must be a unit and so $b = 1$, i.e. $x_1 = y_1$. Cancelling the factor x_1 and applying induction on $\max(m, n)$ we find that the two expressions for a in (1) agree, so S is indeed free on X . ■

If we introduce the *length* $|a|$ of an element a , as the number of factors in a complete decomposition of a , we see that when $|a| \geq |b|$, then $ab' = ba'$ implies that $a = bc$ for some $c \in S$.

Th. 5.1 may be used to give criteria for a submonoid of a free monoid to be free. Consider, for example, the free monoid on one free generator x (the free cyclic monoid). The submonoid generated by x^2 and x^3 is commutative and, if it were free, would be cyclic, which is clearly not the case. The conditions of the next theorem make it easy to find other examples of non-free submonoids of free monoids:

THEOREM 5.2. *A submonoid T of a free monoid S is free if and only if one of the following three equivalent conditions holds:*

(a) *given $s \in S$, if $as, sa' \in T$ for some $a, a' \in T$, then $s \in T$; thus*

$$TT^{-1} \cap T^{-1}T \subseteq T,$$

(b) *T is rigid,*

(c) *given $a, b, b' \in T$, if $ab' = ba$, then $a = bc$ or $b = ac$ for some $c \in T$.*

A subset T satisfying (a) is also called an *anti-ideal* in S .

Proof. If (b) holds we can use Th. 5.1 to show that T is free: clearly T satisfies cancellation and it can have no units apart from 1. To express a non-unit a of T as a product of atoms of T , we simply write a as a product of members of T with as many factors as possible; the number of factors other than 1 is bounded by the length of a in S . Conversely, it is clear that a free submonoid T satisfies (b).

(a) $\Rightarrow$ (c). Let $ab' = ba$ ($a, b, b' \in T$), where we may take $b \neq 1$, without loss of generality. Let us put $a = b^r u$ ($u \in S$) with r chosen as large as possible; this is permissible since r is bounded by the length of a in S . Then $b^r u b' = b^{r+1} u$, hence

$$bu = ub'. \quad (2)$$

By the rigidity of S we have $b = uv$ or $u = bv$ for some $v \in S$; here the second

alternative conflicts with the definition of r , hence

$$b = uv, \quad b' = vu. \quad (3)$$

We conclude that $av = b^{r+1} \in T$, $va = b'^{r+1} \in T$; applying (a) we find that $v \in T$ and hence applying (a) again to (3) we conclude that $u \in T$. Thus if $r > 0$, we have $a = b \cdot b^{r-1}u$, while for $r = 0$, $a = u$ and $b = av$, so (c) holds.

(c) $\Rightarrow$ (b). Let $ab' = ba'$, then $b'a \cdot b'b = b'b \cdot a'b$; by (c), either $b'a = b'bc$ or $b'b = b'ac$ for some $c \in T$, and by cancelling b' we obtain (b).

(b) $\Rightarrow$ (a). Let $au = b$, $ua' = b'$, where $a, a', b, b' \in T$; then $aud' = ba' = ab'$, hence by (b), $a = bc$ or $b = ac$ for some $c \in T$. If $a = bc$, then $a = auc$ and so $u = c = 1$. If $b = ac$, then $ac = au$ and so $u = c$. In either case $u \in T$, which proves (a). ■

We note that whereas (a) refers to S , conditions (b) and (c) are intrinsic in T . In a free monoid, let us associate with every element $u \neq 1$ the shortest element of which it is a power (the 'least repeating segment') and call this the *root* of u . For example, $xyxyxy$ has the root xy . Every element $u \neq 1$ has a unique root, which may also be characterized as an element of the shortest positive length commuting with u . From the criteria of Th. 5.2 we easily obtain

COROLLARY 5.3. *Let a, b be two elements other than 1 in a free monoid S . Then the following conditions are equivalent:*

- (a) a, b form a free subset of S ,
- (b) $ab \neq ba$,
- (c) a, b have distinct roots.

Proof. If (c) fails, we have $a = u^r, b = u^s$, hence $ab = ba$ and (b) fails. If (b) fails, so does (a) and finally, if a, b are not free, then there is a non-trivial relation which after cancelling common left factors reduces to $a \dots = b \dots$. Suppose that $|b| \leq |a|$, then by rigidity, $a = bc$. Now the submonoid generated by a, b is also generated by b, c and so a, b have the same root, by induction on $|a| + |b|$. ■

This result shows that commutativity is an equivalence relation on the set of elements different from 1, and the equivalence classes, with 1 adjoined, are free cyclic monoids.

In 0.11 we saw that the free algebra $k\langle X \rangle$ may also be defined as the monoid algebra kX^* ; the weak algorithm holding in the free algebra may be regarded as the analogue of rigidity and it is clear that any free submonoid T generates a free subalgebra kT , though, of course, not every free subalgebra

of $k\langle X \rangle$ is of this special form. We shall briefly consider free generating sets of free submonoids; this is part of the theory of codes, an extended treatment of which would go far beyond the framework of this book (cf. Lothaire [83], Berstel, Perrin and Schützenberger [a]).

Let X be a finite set; by a *code* Y on X we understand a subset of $X^+ = X^* \setminus \{1\}$ which is a free generating set of the submonoid it generates. For example, if $X = \{x, y\}$, then $\{x, xy, xy^2\}$ is a code, but $\{x, xy, yx\}$ is not, because $x \cdot yx = xy \cdot x$. Suppose that Y is a subset of X^+ which is *not* a code, then we have an equation between two distinct words in the elements of Y and by cancellation we may take this to be of the form $yu = y'v$, where $y, y' \in Y, y \neq y'$. By the rigidity of X^* we have $y = y'z$ or $y' = yz$ for some $z \in X^*$, say the former holds. Then we say that y' is a *prefix* of y . A subset Y of X^+ is called a *prefix set* if no element of Y is prefix of another, and what has been said shows the truth of

PROPOSITION 5.4. *Every prefix set in a free monoid is a code. ■*

For this reason a prefix set in a free monoid is also called a *prefix code*. If N is a submonoid of a free monoid X^* with minimal generating set Y , then Y is a prefix code if and only if $Y^{-1}Y = Y$.

Suffix sets are defined by left-right symmetry; they are again codes, e.g. $\{x, xy, xy^2\}$ is a suffix set but not a prefix set. 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 (this is also known as a 'zero-delay code'). By contrast, if a code is not prefix, say if $y = y'u$, then at any occurrence of y' it is necessary to read past it to see whether y or y' is intended.

Let X be a set of r elements; then X^* contains r^n words of length n . We shall again write $|u|$ for the length of u , and for any subset Y of X^+ define its *weight* (possibly ∞) as

$$\mu(Y) = \sum_{y \in Y} r^{-|y|}. \quad (4)$$

Writing X^n for the set of all words of length n in X^* , we see that $\mu(Y \cap X^n)$ may be regarded as the 'weight' of $Y \cap X^n$ as a fraction of $\mu(X^n) = 1$, in particular, $\mu(Y \cap X^n) \leq 1$.

If Y is finite and $1 \notin Y$, then

$$\mu(Y) \leq \max\{|y| \mid y \in Y\}, \quad (5)$$

for if the right-hand side is N , then $Y \subseteq X \cup X^2 \cup \dots \cup X^N$.

For any two subsets Y, Z of X^+ let $YZ = \{yz \mid y \in Y, z \in Z\}$, then

$$\mu(YZ) \leq \mu(Y)\mu(Z), \quad (6)$$

and when the right-hand side is finite, equality holds if and only if each member of YZ factorizes uniquely as $yz, y \in Y, z \in Z$. In particular, for a code Y we have

$$\mu(Y^n) = \mu(Y)^n, \quad n = 1, 2, \dots \quad (7)$$

We can now state a necessary condition for a subset of X^+ to be a code:

THEOREM 5.5. *Let X be a set of r elements. If Y is a code in the free monoid X^* , then*

$$\mu(Y) \leq 1. \quad (8)$$

Proof. Suppose first that Y is finite, say $\max\{|y| \mid y \in Y\} = N$. Then by (7) and (5),

$$\mu(Y)^n = \mu(Y^n) \leq nN.$$

Hence $\mu(Y) \leq (nN)^{1/n}$ and letting $n \rightarrow \infty$, we obtain (8). Since any subset of a code is again a code, the result follows for all finite subsets of a code, and by passing to the limit we see that it holds generally. ■

Remarks. 1. If $Y = \{y_1, y_2, \dots\}$, where $|y_i| = n_i$, then (8) takes the form (Kraft–McMillan inequality)

$$\sum r^{-n_i} \leq 1. \quad (9)$$

This condition is not sufficient for Y to be a code, but when it holds, we can always find a code with elements of lengths $n_1, n_2, \dots$ (cf. Ex. 13).

2. A *maximal* code is a code which is not a proper subset of any code. By Th. 5.5 we see that any code Y such that $\mu(Y) = 1$ is necessarily maximal, e.g. if $X = \{x, y\}$, then $\{x, xy, xy^2, y^3\}$ or $\{x, xy, xy^2, \dots\}$ are maximal codes.

Th. 5.5 has an analogue for free algebras, which we shall now derive. For any subset Y of $k\langle X \rangle \setminus k$ let us define the *weight* as

$$\mu(Y) = \sum r^{-\deg y}.$$

To get a measure for the elements of degree n , we need to take a basis (rather than the set of all elements), so to obtain an analogue of (5) we shall assume that $Y \cup \{1\}$ is linearly independent. Then

$$\mu(Y) \leq \max\{\deg y \mid y \in Y\}. \quad (10)$$

For suppose that the right-hand side of (10) is N . If $y, y' \in Y$ have the same monomial as leading term, we can replace y' by $y' - \alpha y$, with $\alpha \in k$ chosen so that $y' - \alpha y$ has smaller leading term than y' . By repeating this process if

necessary, we can bring Y to a form where no two elements have the same leading term, and the process cannot decrease $\mu(Y)$; moreover, the linear independence of $Y \cup \{1\}$ is preserved. Thus

$$\mu(Y) = \sum r^{-\deg y} \leq \sum_1^N r^i r^{-i} = N,$$

and this proves (10). With these preparations we can state

THEOREM 5.6. *If Y is a free subset of $k\langle X \rangle$, then $\mu(Y) \leq 1$.*

Proof. As in Th. 5.5 it is enough to consider the case where Y is finite, say $\max\{\deg y \mid y \in Y\} = N$. Since Y is free, we have $\mu(Y^n) = \mu(Y)^n$ and by (10), $\mu(Y^n) \leq nN$, so as before we find $\mu(Y) \leq 1$. ■

We conclude this section by considering another monoid associated with rings. In any ring R the two-sided ideals form a monoid under the usual multiplication

$$ab = \{\sum a_i b_i \mid a_i \in a, b_i \in b\}.$$

It is easy to verify that for principal ideal domains this monoid is free abelian (cf. Ex. 14). We shall show that for a fir it is a free monoid in most other cases.

THEOREM 5.7. *Let R be a fir which has no non-unit right-invariant elements or no non-unit left invariant elements. Then the ideals of R form a free monoid.*

If, as seems plausible, every fir with a non-unit right invariant element is right principal, the conclusion of Th. 5.7 holds for all firs which are not left and right principal, but it is not known whether this is the case.

Proof. We have to verify the conditions of Th. 5.1, and we shall proceed in a number of steps. Some of the properties actually hold under weaker assumptions.

(i) *If R is a right fir, a a non-zero right ideal and b, b' any left ideals, then*

$$ab = ab' \Rightarrow b = b'.$$

For if we write $a = \oplus a_i R$, then the equation $ab = ab'$ becomes $\oplus a_i b = \oplus a_i b'$; equating the cofactors of a_i , we find $b = b'$. ■

(ii) *Let R be a right fir and a, b ideals of R satisfying $a + b = R$, $ab = ba$. Then either a or b is principal as right ideal of R . Hence if R has no non-unit right invariant elements, then*

$$a + b = R, \quad ab = ba \Rightarrow a = R \text{ or } b = R. \quad (11)$$

Proof. If $a = R$ or $b = R$, there is nothing to prove, so assume the contrary. Let $a = \bigoplus a_i R$, $b = \bigoplus b_j R$; since $Rb = b$, we have

$$ab = \bigoplus a_i b = \bigoplus a_i b_j R. \quad (12)$$

Since $a + b = R$, we can find $x \in a$ such that $1 - x \in b$. Write $x = \sum_I a_i c_i$. If a had infinite rank, I would be infinite and we could find $i_0 \in I$ such that $c_{i_0} = 0$. Now

$$\bigoplus a_i b = ab = ba \supseteq (1 - x)a_{i_0} = a_{i_0} - \sum a_i c_i a_{i_0}.$$

Equating right cofactors of a_{i_0} , we find that $1 \in b$, a contradiction. Therefore a has finite rank, p say, and similarly b has finite rank q . Consider the exact sequence

$$0 \rightarrow a \cap b \rightarrow a \oplus b \rightarrow a + b \rightarrow 0.$$

We have $a \cap b = (a \cap b)(a + b) = ba + ab = ab \subseteq a \cap b$, so the left-hand term is ab , which by (12) is free of rank pq . The right-hand term is R , which is free of rank 1 and the middle term is free of rank $p + q$, hence $pq - (p + q) + 1 = 0$, i.e. $(p - 1)(q - 1) = 0$, so $p = 1$ or $q = 1$, i.e. a or b is principal.

Now the second assertion follows because any ideal which is principal as right ideal has a right invariant generator. ■

(iii) Let R be a fir which has no non-unit right invariant elements or no non-unit left invariant elements. Then the monoid of non-zero ideals of R is rigid.

Proof. Assume that R has no non-unit right invariant elements, say. Let a, b, a', b' be non-zero ideals such that

$$ab' = ba'. \quad (13)$$

We have to find an ideal c such that $a = bc$ or $b = ac$. Suppose first that $a + b = R$ and $a' + b' = R$. Then

$$b' = ab' + bb' = ba' + bb' = b,$$

and similarly $a' = a$. Hence $ab = ba$ and R has no non-unit right invariant elements, therefore $a = R$ or $b = R$, by (ii) and the desired conclusion follows. Note that in (13) we only had to assume that a, b are right-ideals and a', b' left ideals.

Next assume merely that $a + b = R$ and let $a' + b' = \bigoplus Re_i$. Pick $i_0 \in I$ and let $\pi: \bigoplus Re_i \rightarrow R$ be the left linear functional 'cofactor of e_{i_0} '. Then $a'\pi + b'\pi = R$ and $a \cdot b'\pi = b \cdot a'\pi$, where $a'\pi, b'\pi$ are not both 0, hence neither is. By the previous case we can find an ideal c such that $a = bc$ or $b = ac$.

Finally, the general case can be reduced to the case just considered, by putting $a + b = \oplus f_j R$. ■

(iv) Let us write $a < b$ to mean: $a = cb$, where $c \neq R$. Every left fir satisfies ACC with respect to ' $<$ '.

Proof. Suppose that we have a strictly ascending chain

$$0 \neq b_0 < b_1 < b_2 < \dots, \quad (14)$$

say $b_{i-1} = a_i b_i$, then $b_0 = a_1 b_1 = a_1 a_2 b_2 = \dots$, hence

$$b_0 \subseteq \cap a_1 \dots a_n,$$

but the intersection on the right is 0, by Th. 5.11.3, while $b_0 \neq 0$. This contradiction shows that (14) must break off. ■

(v) In any fir, every non-zero ideal is a product of atoms (in the monoid of non-zero ideals).

Proof. We first show that every proper ideal has an atomic right factor. Given $a \neq R$, if a is not an atom, we can write it as $a = a_1 b_1$ where a_1, b_1 are non-units. If b_1 is not an atom, we can write $b_1 = a_2 b_2$, where a_2, b_2 are non-units. Continuing thus, we obtain a sequence (14), which must break off; but this can only happen when b_i is an atom. Thus we have found an atomic right factor of a .

By symmetry a also has an atomic left factor. Now decompose a , but splitting off an atomic left factor at each stage: $a = a_1 b_1 = a_1 a_2 b_2 = \dots$. This sequence again breaks off, by (iv), and when it does, a has been expressed as a product of atoms. ■

Exercises 6.5

1. Show that any retract of a free monoid is free.
2. Use Th. 5.2 to find a procedure for reducing a finite subset Y of a free monoid to a set Y' such that (i) Y and Y' generate the same submonoid T and (ii) if T is free, then Y' is a code.
- 3°. Let F be the free group on a finite set X and let X^* be the submonoid generated by X . Classify the submonoids between X^* and F . Does this set of submonoids satisfy ACC?

4. Show that every submonoid of a free monoid has a unique minimal generating set, consisting of its atoms.

5. Show that in a free monoid, the set of palindromes (i.e. fixed elements under the order-reversing anti-automorphism) of even length is a free submonoid, with a prefix code as generating set. What about the set of all palindromes?

6. Show that the intersection of any family of free submonoids of a free monoid is free.

7*. Let A^* be a free monoid on a set A , and for any subset X define the *free hull* of X as the minimal generating set of the intersection of all free monoids containing X . If X is any finite subset of A^+ and its free hull is Y , consider the mapping $\alpha: X \rightarrow Y$ defined as follows: for $x \in X$, x^α is the word $y \in Y$ such that $x \in yY^*$. Show that α is well-defined (using the fact that Y is a code) and surjective (take $z \in Y \setminus X^\alpha$ and show that $Z = (Y \setminus z)z^*$ is a code such that $X \subseteq Z^* \subseteq Y^*$). Show that if α is injective, then X is a code, and deduce the defect theorem: the free hull Y of a finite subset $X \subseteq A^*$, not a code, satisfies $\text{card}(Y) < \text{card}(X)$ (cf. Lothaire [83], p. 6).

8. Let S be a monoid in which the relation ' $a \leq b$ if and only if $ax = b$ for some $x \in S$ ' is a well-ordering. If S also has right cancellation ($xz = yz \Rightarrow x = y$), show that S either consists of one element or is isomorphic to $\mathbb{N}$.

9*. Let S be a monoid that is well-ordered under the relation in Ex. 8 and has left cancellation. Show that S has a presentation with generators x_α ($\alpha < \tau$) and defining relations $x_\alpha x_\beta = x_\beta$ if $\alpha < \beta$, where τ is some ordinal (Cohn [61]).

10. Assume the monomials in X to be totally ordered in some way (by increasing degree). Show that in $k\langle X \rangle$ the homogeneous elements with leading coefficient 1 form a free monoid.

11. For any set X and any $n \geq 1$, show that X^n is a maximal code.

12. Give an example of a subset of X^+ which fails to be a code, yet satisfies the Kraft-McMillan inequality. (Hint: Consider the effect of enlarging X .)

13. Let $X = \{0, 1, \dots, r-1\}$ and let $1 \leq n_1 \leq n_2 \leq \dots$ be a sequence of integers satisfying $\sum r^{-n_i} \leq 1$. Define s_i as the partial sum: $s_1 = 0$, $s_{i+1} = s_i + r^{-n_i}$ and put $p_i = r^{n_i} s_i$. Verify that p_i is an integer satisfying $0 \leq p_i < r^{n_i}$. Let $y_i \in X^*$ be p_i in the r -adic scale. Verify that y_i has at most n_i digits (so by prefixing 0's it can be taken to have exactly n_i digits), and that $\{y_1, y_2, \dots\}$ is a prefix code.

14. Let R be a principal ideal domain. Show that each non-zero ideal of R is generated by an invariant element. Deduce that $ab = ba$ for all ideals a, b . (Hint: Recall that similar invariant elements are associated, by Lemma 4.2.6.) Then conclude that the monoid of non-zero ideals is free commutative.

6.6 Subalgebras and ideals of free algebras

A subalgebra of a free associative algebra is not necessarily free. This is already clear from the commutative case, by considering the subalgebra of

the polynomial ring $k[x]$ generated by x^2 and x^3 . In the case of the polynomial ring there is a simple criterion for a subalgebra to be free.

PROPOSITION 6.1. *A subalgebra R of a polynomial ring $k[x]$ is free if and only if it is integrally closed.*

Proof. Assume that R is a free subalgebra; then either $R = k$ or $R = k[y]$ for some y transcendental over k . In either case R is clearly integrally closed.

Conversely, let R be integrally closed and denote its field of fractions by K . Since $k \subseteq R$, we have $k \subseteq K \subseteq k(x)$. If $K = k$, then $R = k$ and the result follows. Otherwise, by Lüroth's theorem, K is a simple purely transcendental extension of k . Since R is integrally closed, it is an intersection of valuation rings of K (cf. 6.1). Now any valuation ring of K is of the form $\mathfrak{o}_p \cap K$, where $\mathfrak{o}_p$ is a valuation ring of $k(x)$ (cf. e.g. Cohn [77], p. 324). If x is finite at p , then $\mathfrak{o}_p \supseteq k[x] \supseteq R$, whence $R \subseteq \mathfrak{o}_p \cap K$. Thus R is contained in all valuation rings of K over k , except at most one, namely the one obtained from the pole of x . Since $R \neq k$, R is not contained in all the valuation rings; if $\mathfrak{o}_q$ is the exceptional one, we can choose a generator y of K so that q is a pole of y , and since R is the intersection of all the other valuation rings of $K = k(y)$ over k , it follows that $R = k[y]$. ■

For free algebras on more than one generator, no such convenient criterion is known. We can of course use the characterization in terms of the weak algorithm given in 2.4. Applied to subalgebras, this yields

PROPOSITION 6.2. *Let $F = k\langle X \rangle$ be a free k -algebra. Then a subalgebra R of F is free if and only if there is a filtration on R (over k) for which R satisfies the weak algorithm. ■*

The difficulty in applying this criterion lies in the fact that the different degree-functions in a free algebra are not related in any obvious way. If F is a free algebra and R a subalgebra, then any degree-function on F will define a degree on R , and if R satisfies the weak algorithm for this function, then it must be free. However, this sufficient condition is not necessary; thus if R is a free subalgebra of F , then there is always a degree-function on R for which the weak algorithm holds (Prop. 6.2), but this function need not be defined on all of F . Let us call a subalgebra R of F *regularly embedded* if there is a degree-function on F for which both F and R satisfy the weak algorithm. Examples of irregularly embedded free subalgebras are given in Ex. 2 and 3.

There is one case where more information is available, namely when $F = k[x]$ is free of rank 1. We already know a simple test for subalgebras to

be free (Prop. 6.1) and, as we shall see, there is also a simple criterion in terms of the algorithm. For in this case (and only here) the only automorphisms are the affine transformations $x \mapsto ax + b$ ($a, b \in k$, $a \neq 0$), therefore the degree-function on F is unique up to a scalar factor. We thus obtain

PROPOSITION 6.3. *A subalgebra of $k[x]$ is free if and only if it has the division algorithm relative to the x -degree. ■*

Sometimes it is possible to obtain conditions for a homogeneous subalgebra to be free. Thus, using Cor. 2.9.14, we have

THEOREM 6.4. *Let F be a free algebra and R a homogeneous subalgebra, with augmentation ideal α . Then the following conditions are equivalent:*

- (a) *for every homogeneous subspace V of R such that $\alpha = V \oplus \alpha^2$, the natural homomorphism $\psi: k\langle V \rangle \rightarrow R$ is an isomorphism,*
- (b) *R is a free k -algebra,*
- (c) *R is right hereditary,*
- (d) *α is flat as right R -module,*
- (e) *the natural map $\alpha \otimes \alpha \rightarrow \alpha^2$ is an isomorphism.*

Proof. (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) are clear ((d) $\Rightarrow$ (e) follows by Cor. A.5), and to prove (e) $\Rightarrow$ (a) we need only show, by Cor. 2.9.14, that ψ is surjective. By hypothesis, $\alpha = V + \alpha^2$, hence $\alpha^n = (V + \alpha^2)^n \subseteq V^n + \alpha^{n+1}$, and so we have

$$\alpha^n = V^n + \alpha^{n+1} \quad \text{for all } n \geq 1. \quad (1)$$

It follows, on setting $R' = \text{im } \psi$, that

$$R = R' + \alpha^n \quad \text{for all } n \geq 1. \quad (2)$$

For this holds when $n = 1$, and since $V \subseteq R'$, it follows generally by induction, using (1). Now let u be a homogeneous element of R , take $n > \deg u$ and write $u = r' + a$, where $r' \in R'$, $a \in \alpha^n$. Since V is homogeneous, so is R' and hence $u = r' \in R'$, $a = 0$. Thus $R' = R$ and ψ is surjective, as claimed. ■

From the proof we obtain

COROLLARY 6.5. *Any homogeneous free subalgebra of F has a homogeneous set of free generators. ■*

There is another characterization of homogeneous subalgebras that is often useful.

THEOREM 6.6. *Let $F = k\langle X \rangle$ be a free k -algebra and R a homogeneous subalgebra. Then R is free, with a homogeneous free generating set Y which is right F -linearly independent if and only if F is free, or equivalently, flat as left R -module.*

Proof. Let R be free with a homogeneous free generating set Y which is right F -linearly independent, and let W be the k -space spanned by Y , so that $W \otimes_k F = YF$, $R = k\langle W \rangle$. Choose a homogeneous k -space V such that $F = V \oplus YF$; then

$$\begin{aligned} F &\cong V \oplus W \otimes F \\ &\cong V \oplus W \otimes V \oplus W \otimes W \otimes F \cong \dots \\ &\cong (k \oplus W \oplus W \otimes W \oplus \dots) \otimes V \\ &\cong R \otimes V, \end{aligned}$$

and this shows V to be spanned over k by a left R -basis of F .

Thus F is free and hence flat. Conversely, if F is flat as left R -module, take a resolution of k :

$$0 \rightarrow \mathfrak{f} \rightarrow F \rightarrow k \rightarrow 0,$$

where $\mathfrak{f}$ is the augmentation ideal of F . Similarly, if α is the augmentation ideal of R , we have an exact sequence

$$0 \rightarrow \alpha \rightarrow R \rightarrow k \rightarrow 0.$$

Since R is free as R -module, we obtain an exact sequence (by the extended Schanuel lemma):

$$0 \rightarrow \alpha \rightarrow \mathfrak{f} \oplus R \rightarrow F \rightarrow 0.$$

Here $F, R, \mathfrak{f}$ are left R -flat, hence so is α . Now Th. 6.4(d) shows that R is free, with homogeneous generating set Y . Hence Y is a right R -basis of α . But the embedding

$$0 \rightarrow R^{(Y)} = \alpha \rightarrow R$$

remains exact under the operation $\otimes_R F$, therefore Y is right F -linearly independent. ■

We now turn to examine right ideals in a free algebra. If $F = k\langle X \rangle$ is a free k -algebra and $\mathfrak{r}$ a right ideal, then $F/\mathfrak{r}$ is a right F -module, hence it is a k -space and we shall find that it has a basis derived from a prefix code in X . This will lead to another proof that $\mathfrak{r}$ is free as right R -module, but some preparations are necessary.

In every partially ordered set S with DCC there is a natural bijection between lower segments and antichains (i.e. sets of pairwise incomparable elements): with every lower segment L we associate the set L^0 of minimal elements of the complement $S \setminus L$, clearly an antichain, and with every antichain A we associate A^0 , the complement of the upper segment generated by A . It is easily checked that these mappings are mutually inverse.

Let X be a finite set and X^* the free monoid on X , partially ordered by left divisibility:

$$a \leq b \text{ if and only if } b = ac \text{ for some } c \in X^*. \quad (3)$$

It is clear that X^* satisfies DCC for this partial ordering. A lower segment in this partial ordering is called a *Schreier set*; such a set is characterized by the fact that with any word it contains all its left factors (prefixes). An antichain in X^* is just a prefix set; thus we have a natural bijection between the prefix sets and Schreier sets. Starting from a prefix set C , the corresponding Schreier set C^0 consists of all words with no prefix in C , while for a Schreier set P the corresponding antichain P^0 consists of all words not in P but with every proper prefix in P .

PROPOSITION 6.7. *Let X be any set. Then there is a natural bijection between the prefix codes and the non-empty Schreier sets on X^* , and if C, P correspond in this way, then*

$$X^* = C^*P. \quad (4)$$

It only remains to establish (4). Let $w \in X^*$; either $w \in P$ or w has a maximal prefix p in P , in which case $w = pxu$, where $x \in X$. It follows that $px \in C$ and by induction on the length we have $u \in C^*P$, hence $w \in C^*P$ and (4) follows. ■

THEOREM 6.8. *Let $F = k\langle X \rangle$ be a free algebra and let $\mathfrak{r}$ be any right ideal of F . If P is any Schreier set which is maximal k -linearly independent (mod $\mathfrak{r}$) and $C = P^0$ is the corresponding prefix set, then for each $c \in C$ there exists an element of $\mathfrak{r}$:*

$$f_c = c - \sum \alpha_{c,p} p \quad (p \in P),$$

such that the f_c form a basis of $\mathfrak{r}$ as right F -module.

Proof. Let P, C be as stated. For each $c \in C$, $P \cup \{c\}$ is still a Schreier set, and by the maximality of P this set is linearly dependent (mod $\mathfrak{r}$), say

$$f_c = c - \sum \alpha_{c,p} p \in \mathfrak{r}. \quad (5)$$

We claim that any element r of F can be written in the form

$$r = \sum f_c g_c + \sum \beta_p p, \quad \text{where } g_c \in F, \beta_p \in k. \quad (6)$$

By linearity it is enough to prove this when r is a monomial w , and we may further assume that $w \notin P$. By Prop. 6.7 we can then write $w = cu$ for some $c \in C$, hence using (5) we have

$$w = f_c u + \sum \alpha_{c,p} p u.$$

Now for any $p \in P$, either $pu \in P$ or $pu = c_1 u_1$, where $|p| < |c_1|$, $|u_1| < |u|$. In the first case we have achieved the form (6), in the second case we can apply induction on $|u|$ to express $c_1 u_1$ in the same form, and (6) follows.

Now take $r \in \mathfrak{r}$ and apply the natural homomorphism $F \rightarrow F/\mathfrak{r}$. Writing the image of a as $\bar{a}$, we have

$$0 = \bar{r} = \sum \beta_p \bar{p},$$

but by hypothesis the $\bar{p}$ are linearly independent, hence $\beta_p = 0$ in this case and so $r = \sum f_c g_c$. This shows that the f_c generate $\mathfrak{r}$. To prove their independence, assume that $\sum f_c g_c = 0$, where the g_c are not all 0. Then by (5),

$$\sum c g_c = \sum \alpha_{c,p} p g_c. \quad (7)$$

Let w be a word of maximal length occurring in the g_c , say in g_{c_0} . Since C is a prefix code, $c_0 w$ occurs with a non-zero coefficient, λ say, on the left of (7). Hence we have

$$\lambda = \sum \alpha_{c,p} \mu_{c,p},$$

where $\mu_{c,p}$ is the coefficient of $c_0 w$ in $p g_c$. Now $c_0 w = pu$ can only hold when p is a proper left factor of c_0 , hence $|p| < |c_0|$, $|u| > |w|$, and this contradicts the definition of w . Hence the f_c form a free generating set of $\mathfrak{r}$, as claimed. ■

In particular, this shows again that each right ideal in F (and, by symmetry, each left ideal) is free (Cor. 2.4.3). The uniqueness of the rank is clear since F has a homomorphism to k and so has IBN. We can also use Th. 6.8 to obtain another derivation of the Schreier–Lewin formula. If $F/\mathfrak{r}$ has dimension r over k and $\mathfrak{r}$ has rank n as right F -module, then the relation $X^* = C^*P$, as an identity in $k\langle\langle X \rangle\rangle$, gives the equation $C - 1 = P(X - 1)$. Here $\text{card}(C) = n$, $\text{card}(P) = r$ and if $\text{card}(X) = d$, then we again find

$$n - 1 = r(d - 1).$$

Free subalgebras of a free algebra F have the property that they contain no non-zero ideal of F . This result, itself not hard to prove (Th. 6.12), is best studied in a slightly wider context.

Let R be an integral domain; a subring S of R is called an *anti-ideal* if, for any $a \in R$, $b, c \in S^\times$, $ab, ca \in S$ implies $a \in S$. Some equivalent conditions for this to happen are given in

PROPOSITION 6.9. *For any integral domain R and subring S , the following conditions are equivalent:*

- (a) *for all $a \in R \setminus S$, either $aS \cap S = 0$ or $Sa \cap S = 0$,*
- (b) *for all $a \in R \setminus S$, $aS \cap Sa \cap S = 0$,*
- (c) *for all $a \in R$, $b \in S^\times$, if $ab, ba \in S$, then $a \in S$,*
- (d) *for all $a \in R$, $b, c \in S^\times$, if $ab, ca \in S$, then $a \in S$, i.e. S is an anti-ideal in R .*

Proof. (a) $\Rightarrow$ (b) is clear. (b) $\Rightarrow$ (c). If $ab, ba \in S$, then $ab^2a \in aS \cap Sa \cap S = 0$, but $a \neq 0$, so $b = 0$, a contradiction. (c) $\Rightarrow$ (d). If $ab, ca \in S$, then $abc, bca \in S$ and by (c), $a \in S$. Now (d) $\Rightarrow$ (a) is clear. ■

COROLLARY 6.10. *If R is an integral domain and S is an anti-ideal in R , $S \neq R$, then S contains no non-zero ideal of R .*

For if $S \supseteq a \neq 0$, let $c \in a$, $c \neq 0$ and $a \in R \setminus S$, then $ac, ca \in a \subseteq S$, hence $a \in S$, which is a contradiction. ■

The next result provides us with a supply of anti-ideals. A subring S of a ring R is said to be *division closed* if every unit of R which lies in S is also a unit in S .

LEMMA 6.11. *Let R be an integral domain and S a subring. If S is a 2-fir, division closed in R , then S is an anti-ideal.*

Proof. Let $a \in R$ and suppose that $b, c \in S^\times$ are such that $ab, ca \in S$. Then in S we have $c \cdot ab = ca \cdot b$, hence $cS + caS$ is principal, say dS , where $d \in S$. Now $dR = cR + caR = cR$, hence $c = du$, where u is a unit in R . We have $du = c \in dS$ and so $du = dv$, where $v \in S$; but $d \neq 0$, therefore $u = v \in S$. Since $dua = ca \in dS$, it follows that $ua \in S$; now u is a unit in R , hence in S , so $a \in u^{-1}S = S$, as claimed. ■

In a free k -algebra, the units all lie in k and so are contained in every subalgebra; hence we obtain

THEOREM 6.12. *In a free algebra, every subalgebra which is a 2-fir (in particular, every free subalgebra) is an anti-ideal. ■*

Exercises 6.6

1. Let $F = k\langle x, y \rangle$; show that the subalgebra R generated by $x + y^2$ and y^3 is free on these generators but does not satisfy the weak algorithm relative to the (x, y) -degree. By a suitable change of variables show that R is regularly embedded in F .

2. (D. R. Lane) Let $F = k\langle x, y \rangle$; show that the subalgebra generated by $u = (xy)^4 + y(xy)^2$ and $v = (xy)^3$ is free, but not regularly embedded. (Hint: Verify that $d([u, v]) < d(uv) = d(vu)$ for any degree-function d ; now use Cor. 7.4, below.)

3. (W. Dicks) Let $F = k\langle x, y \rangle$; show that the subalgebra generated by $u = yxy - y$, $v = yux$ and $w = uxy$ is free on these generators, but is not regularly embedded in F .

4. (R. E. Williams) In a free algebra F , let R be the subalgebra generated by $b_1, \dots, b_n$ such that $\deg(b_1) = \dots = \deg(b_n)$, and this is the minimum degree of elements in R . Show that if the elements $b_1, \dots, b_n$ are linearly independent, they form a free generating set.

5. Let A_n be the free k -algebra of rank n . Show that if there is a surjective homomorphism $f: A_m \rightarrow A_n$, then $n \leq m$, with equality if and only if f is an isomorphism.

6°. (Bergman) With the notation of Ex. 5, let $f: A_m \rightarrow A_n$ be a surjective homomorphism. Does there exist an automorphism g of A_m such that $(\ker f)^g$ is the ideal of A_m generated by $x_{n+1}, \dots, x_m$?

7°. (W. E. Clark) Is every retract of a free algebra of rank n free? (For $n = 1, 2$ see 6.7.)

8*. (Bergman) In $F = k\langle x, y \rangle$ consider the submonoid generated by $u = xy$, $v = yx$, $z = x^2$ and y . Verify that it satisfies the relations $yz^n u = vz^n y$ ($n = 1, 2, \dots$), but that none of these equations is implied by the earlier ones. Deduce the existence of a finitely generated subalgebra of F which is not finitely presented.

9. (Kolotov [78]) Show that in any integral domain, the family of anti-ideals is closed under intersections, unions of chains, and if C is an anti-ideal in B and B in A , then C is an anti-ideal in A . Show also that an anti-ideal of R contains with any right ideal of R its idealizer in R .

10. Let R be a commutative integral domain. Show that a subring of R is an anti-ideal if and only if it is integrally closed.

11. (W. Dicks) Let $F = k\langle x, y, z \rangle$ and let R be the subalgebra generated by $a = xy$, $b = zyx + x + z$, $c = xyz + x + z$, $d = yx$. Verify that R is an anti-ideal, in fact, for any $f \in F \setminus R$, $fR \cap R = 0$, but that R is not a 2-fir. (Hint: Verify that $a+1$ and c are right comaximal in F and examine the homomorphism of the subalgebra generated by a, d, x, z mapping a, d to -1 and x, z to 0 .)

12*. (Bergman [c]) Let $F = k\langle x, y_1, y_2, y_3, z \rangle$, G_1 the subalgebra generated by $x, y_1, y_3, y_2 z, z$ and G_2 the subalgebra generated by $x, xy_1, y_1 y_2 - y_3, y_2 z$. Verify that G_1 and G_2 are each free on the given generating sets, but that $G_1 \cap G_2$ is not a 3-fir, hence not free. (Hint: The relation $x \cdot (y_1 y_2 - y_3) z + xy_1 \cdot (-y_2 z) + xy_3 \cdot z = 0$ can be trivialized in G_1 and in G_2 but not in $G_1 \cap G_2$.)

13. (Bergman [c]) Show that in a free k -algebra, the intersection of any family of free k -subalgebras is a 2-fir.

14. A prefix code on X^* is called *complete* if the right ideal generated by it meets every right ideal of X^* . Show that in the correspondence between prefix codes and Schreier sets, the finite Schreier sets correspond to the complete finite prefix codes.

15^o. Prove an analogue of Th. 6.8 for group algebras of free groups and more generally for mixed algebras. Is there an extension for firs that are augmented k -algebras or for rings with weak algorithm?

6.7 Centralizers in power series rings and in free algebras

In a free algebra, few elements commute with each other, and one would expect the centralizer of a non-scalar element to be small. This expectation is borne out, as we shall see in Th. 7.7, below. A similar question can be raised for the free power series ring, and since this is rather easier to answer, we begin with it.

THEOREM 7.1. *Let R be an inversely filtered ring which satisfies 2-term inverse weak algorithm and is complete. Then the centralizer C of any $a \in R$, not zero or a unit, is a complete principal valuation ring.*

If, moreover, R is a connected k -algebra, then the centralizer of any non-scalar element of R is a formal power series ring in one variable over k and hence is commutative.

Proof. Let $x, y \in R$ be right linearly dependent and $v(x) \geq v(y)$ say, then $x - yz$ and y are right v -dependent for any $z \in R$, hence by Lemma 2.9.2, $x = yz$ for some $z \in R$, so $xR \subseteq yR$.

If a is as stated in the theorem, then $v(a) > 0$; let $x, y \in C$, then for any $n \geq 0$, $xa^n = a^n x$, hence for sufficiently large n , $a^n R \subseteq xR$ and $a^n R \subseteq yR$. Thus $xR \cap yR \neq 0$, and so if $v(x) \geq v(y)$, then $x = yz$. Clearly $z \in C$, so $xR \subseteq yR$, i.e. C is a right chain ring. Since R is atomic, by Prop. 2.9.5, and any non-unit of C remains a non-unit in R , it follows that C is atomic and therefore a rigid UFD, by Th. 3.4.2. It contains a large non-unit, namely a , hence it is a right (and by symmetry left) principal valuation ring, by Th. 3.4.9, and it is complete, because R is.

Now assume that R is connected and let x be an element of least positive degree in C , so that every ideal in C has the form $Cx^n = x^n C$. Then every element of C has the form $x^n v$, where v is a unit. The additional hypothesis allows us to write $v = \alpha + x'v'$, ($\alpha \in k$, v' a unit), and an induction argument, using the completeness, shows that $C = k[[x]]$. ■

In particular, the free power series ring satisfies all the hypotheses and we obtain

COROLLARY 7.2. *The centralizer of a non-scalar element in the free power series ring $k\langle\langle X \rangle\rangle$ is of the form $k[[c]]$. ■*

We know from Cor. 2.9.10 that any non-scalar element in a free power series ring generates a free power series ring in one variable. Two elements need not generate a free power series algebra, e.g. the complete algebra generated by x^2 and x^3 is not free. But we have an analogue of Cor. 5.3:

PROPOSITION 7.3. *Let R be a complete inversely filtered ring with 2-term inverse weak algorithm, which is a connected k -algebra. If x, y are two non-commuting elements of R , say $x = \alpha + x_1, y = \beta + y_1$, where $\alpha, \beta \in k, x_1, y_1 \in R_1$, then the complete algebra generated by x, y is the free power series ring $k\langle\langle x_1, y_1 \rangle\rangle$.*

Proof. By hypothesis, $xy - yx \neq 0$, hence $v(xy - yx)$ is a positive integer n ; we shall use induction on n . Since $xy - yx = x_1y_1 - y_1x_1$, we may replace x, y by x_1, y_1 , respectively; for simplicity of notation we shall drop the suffix, thus we may assume that $v(x), v(y) > 0$.

If x, y are not free, we have a relation

$$xf + yg + \lambda = 0, \quad (1)$$

where f, g are expressions in x, y and $\lambda \in k$. Equating constant terms in (1) we find that $\lambda = 0$, so x, y are right linearly dependent. If $v(x) \geq v(y)$, then

$$x = yz, \quad (2)$$

for some $z \in R$. Let $z = \gamma + z_1$ ($\gamma \in k, z_1 \in R_1$); if we can show that y, z are free, then so are x, y by (2), for the elements y, yz are clearly free in $k\langle\langle y, z_1 \rangle\rangle$. Now

$$0 \neq xy - yx = yzy - yyz = y(zy - yz),$$

hence $yz - zy \neq 0$ and $v(yz - zy) < v(xy - yx)$. The result follows by induction, since $v(xy - yx) > 0$ for any $x, y \in R$. ■

The result may be expressed more strikingly by saying: x and y are free if and only if they do not commute. In particular, since any free algebra can be embedded in a power series ring, we have

COROLLARY 7.4. *Any two non-commuting elements of a free k -algebra form a free set. ■*

We now go on to consider centralizers in a free algebra $F = k\langle X \rangle$. Let $a \in F \setminus k$ and denote by C the centralizer of a in F . The embedding of F in

$k\langle X \rangle$ shows that C is commutative, by Cor. 7.2; moreover, C is finitely generated, as module over $k[a]$ or as algebra. For if $d(a) = n$, say, we choose for each integer $v = 0, 1, \dots, n-1$ such that an element of degree $\equiv v \pmod{n}$ occurs in C , an element of least degree $\equiv v \pmod{n}$ in C . Calling these elements $c_0, \dots, c_r$ ($r \leq n-1$), we see that every element in C has the same leading term as some $c_i a^j$ ($i = 0, \dots, r, j = 0, 1, \dots$), by Cor. 7.2, hence the $c_i a^j$ span C over k , so the c_i span C over $k[a]$, and together with a they generate C as k -algebra.

Our aim is to show that C is a polynomial ring over k . In order to establish this fact we shall study homomorphisms of C into polynomial rings. If we look at the leading terms of elements of C we note that all are powers of a given one and it turns out that in order to achieve a homomorphism of C into a polynomial ring we need a preordering of the free monoid which lists each word together with all its powers. This is done by introducing 'infinite' words.

Let X be a totally ordered set and W the set of all right infinite words in X , i.e. infinite sequences of letters from X . Given $u \in X^+$, we denote by u^∞ the word obtained by repeating u indefinitely: $u^\infty = uuu \dots \in W$. Thus we have a mapping $u \mapsto u^\infty$ from X^+ to W which identifies two words if and only if they have the same root†. We shall take W to be ordered lexicographically.

LEMMA 7.5. *Given $u, v \in X^+$, if $u^\infty > v^\infty$, then*

$$u^\infty > (uv)^\infty > (vu)^\infty > v^\infty, \quad (3)$$

and similarly with $>$ replaced by $<$ or $=$ throughout.

Proof. Suppose that $(uv)^\infty > (vu)^\infty$; then

$$(vu)^\infty = v(uv)^\infty > v(vu)^\infty = v^2(uv)^\infty > v^2(vu)^\infty = \dots \rightarrow v^\infty,$$

since the lexicographic order is 'continuous'. Similarly we find that $(uv)^\infty < u^\infty$, therefore (3) follows whenever $(uv)^\infty > (vu)^\infty$. Likewise $(uv)^\infty < (vu)^\infty$ implies

$$u^\infty < (uv)^\infty < (vu)^\infty < v^\infty, \quad (4)$$

while $(uv)^\infty = (vu)^\infty$ implies

$$u^\infty = (uv)^\infty = (vu)^\infty = v^\infty. \quad (5)$$

Now for any u, v exactly one of (3), (4), (5) holds and the assertion follows. ■

The monoid algebra kX^* is just the free algebra $F = k\langle X \rangle$. Given any periodic word z in W , i.e. an infinite power of a word in X , let us define A_z as

† For example, can and $cancan$.

the k -subspace of F spanned by the words u satisfying $u = 1$ or $u^\infty \leq z$, and let I_z be the k -subspace spanned by the words u such that $u \neq 1$ and $u^\infty < z$. By Lemma 7.5, A_z is a subalgebra of F in which I_z is a two-sided ideal.

The set of words u in X such that $u^\infty = z$, together with 1, will form the set of non-negative powers of an element v which we shall call again the *root* of z . It follows that $A_z/I_z \cong k[v]$.

PROPOSITION 7.6. *Let C be a finitely generated subalgebra of a free k -algebra F . If $C \neq k$, then there is a homomorphism f of C into the polynomial ring in one variable over k such that $Cf \neq k$.*

Proof. Let $F = k\langle X \rangle$, where X is totally ordered. Take a finite generating set Y for C and let z be the maximum of u^∞ as u ranges over all monomials $\neq 1$ occurring with non-zero coefficient in members of Y . Then $Y \subseteq A_z$, hence $C \subseteq A_z$ and the quotient map $f: A_z \rightarrow A_z/I_z \cong k[v]$ is non-trivial on C . ■

When C is not finitely generated, the result need no longer hold (cf. Ex. 4).

Consider now a free k -algebra $F = k\langle X \rangle$. Given a non-scalar element a , its centralizer C is a finitely generated commutative k -algebra, as we have seen. Therefore it can be mapped non-trivially into a polynomial algebra; now C as finite extension of $k[a]$ has transcendence degree 1 over k and so it must be embedded in the polynomial algebra. Being integrally closed (by Cor. 1.2), it must be free (Prop. 6.1), so we have proved Bergman's centralizer theorem:

THEOREM 7.7. *Let $F = k\langle X \rangle$ be the free k -algebra on X . Then the centralizer of any non-scalar element of F is a polynomial ring in one variable over k . ■*

Exercises 6.7

1. (Bergman [67]) Let $F = k\langle x, y \rangle$ and let C be the centralizer of an element $a \in F \setminus k$. Show that the valuation on $k[a]$ given by the degree in a is totally ramified on C .

2. (Schur [04]) Let R be the differential ring $k[x; 1, ']$, where $k = F(t)$ is a rational function field and $'$ denotes differentiation with respect to t . Show that the centralizer of any element outside F is commutative. (Hint: Use Th. 7.1 to form the completion of R by negative powers of x .)

3. (Bergman [67]) Let X^* be the free monoid on $X = \{x_1, \dots, x_{n-1}\}$ and W the set of infinite words. With each $u = a_1 a_2 \dots$ in W associate the 'decimal expansion' $\lambda(u) = \sum a_i n^{-i}$ and obtain a formula for $\lambda(u^\infty)$ in terms of $\lambda(u)$ and the length of u . Hence express $\lambda((uv)^\infty)$ as a convex linear combination of $\lambda(u^\infty)$ and $\lambda(v^\infty)$.

4. Let F be a free k -algebra of rank greater than 1. Show that the subalgebra of F

generated by the commutator ideal cannot be mapped into a polynomial ring. Use this fact to show that the hypothesis that C be finitely generated in Prop. 7.6 cannot be omitted.

5. With every ring R we can associate another ring R^{ab} , the ring R made abelian, which is obtained by dividing R by the ideal generated by all the commutators. Thus the natural mapping $R \rightarrow R^{ab}$ is universal for homomorphisms of R into commutative rings. Given any ring R , denote by S the subring generated by the kernel of the natural mapping $R \rightarrow R^{ab}$. Given $a, b \in R$, write $c = [a, b] = ab - ba$ and establish the identity:

$$c[ac, bc] = ca[c, bc] + cb[ac, c] + c^4.$$

Deduce that if $R = k\langle X \rangle$, then S^{ab} has nilpotent elements other than 0 (and so cannot be a polynomial ring).

6. In a free algebra over a field of characteristic 0, show that if $ab + ba = 0$, then either a or b is 0 (this follows from Th. 7.7, but give a direct proof).

7. (Bergman, Kolotov [78]) Let $F = k\langle X \rangle$ be a free algebra. Show that any subalgebra generated by two elements y, z of F is either free on y, z or is contained in a 1-generator subalgebra. If $F = k\langle x, y, z \rangle$, then the elements $u = xyxz + xy$, $u' = zxyx + yx$, $v = yxz$ satisfy the relation $uv = vu'$, but there is no 2-generator subalgebra containing u, v, u' . (Hint: If there is such a subalgebra, B say, it must be a 2-fir, hence an anti-ideal. Use the given relation to show that $xy, yx \in B$ and deduce in turn that $x, y, z \in B$.)

8°. (Bergman [67]) Let F be a free k -algebra and $\hat{F}$ its completion by power series. Given $a \in F$, denote by C, C' its centralizers in $F, \hat{F}$, respectively. Is C' the closure of C in $\hat{F}$?

9*. (Koshevoi [71]) Let R be a k -algebra; a subalgebra is called *pure* if it contains $u \in R$, whenever it contains a non-scalar polynomial in u . Let R be the k -algebra generated by a, b, c, d with the defining relation $ad = bc$. Show that the map $a \mapsto x$, $b \mapsto xy$, $c \mapsto z$, $d \mapsto yz$ extends to a homomorphism of R into $F = k\langle x, y, z \rangle$ which is injective. Show that the image of R is a pure subalgebra.

10*. (Dicks [74]) Let b be any element in a free k -algebra F . Show that the idealizer $I(bF)$ is pure, and hence has the centralizer property of Th. 7.7.

11. Let G be the free group on s_i ($i \in I$) and consider the map $f: kG \rightarrow k\langle\langle x_i | i \in I \rangle\rangle$ given by $s_i \mapsto 1 + x_i$. Show that any multiplicative commutator q maps to a series $1 + \bar{q} + \dots$, where $\bar{q}$ is the corresponding additive commutator and dots indicate terms of higher order. Deduce that f is injective, and hence show that any two elements of kG either commute or generate a free subalgebra.

6.8 Automorphisms of polynomial rings

Our aim in this section and the next is to describe the structure of the automorphism groups of $k\langle X \rangle$ and $k[X]$ when $|X| = 2$. For $|X| = 1$ the

corresponding question is easily answered while for $|X| \geq 3$ the answer is not known. It turns out that the case of the free algebra and that of the polynomial ring are closely related, which is why we treat them together. We shall begin by introducing some terminology and notation.

When R is $k\langle X \rangle$ or $k[X]$, every mapping $X \rightarrow R$ can be extended to a unique endomorphism of R , i.e. an endomorphism of k -algebras or more explicitly, a k -linear ring endomorphism. We shall specify particular endomorphisms as X -tuples in R , e.g. when $X = \{x, y, z\}$, then $(x + yz, y + z^2, z)$ denotes the endomorphism (actually an automorphism)

$$f(x, y, z) \mapsto f(x + yz, y + z^2, z). \quad (1)$$

An automorphism α is said to *stabilize* a subset S of R if $S\alpha = S$; it *fixes* S if $s\alpha = s$ for all $s \in S$. We shall write $\text{stab } S$ for the stabilizer of S .

We begin by defining some classes of automorphisms frequently encountered; here R may be $k\langle X \rangle$ or $k[X]$. The mapping

$$\tau_a: x \mapsto x + a_x, \quad \text{where } x \in X, a \in k^X, \quad (2)$$

defines an automorphism of R called a *translation*.

Any automorphism of the k -space kX with basis X uniquely defines an automorphism of R :

$$x \mapsto \sum \alpha_{xx'} x', \quad \alpha_{xx'} \in k, \quad (3)$$

where for any x , $\alpha_{xx'}$ is zero for almost all $x' \in X$ and the matrix $(\alpha_{xx'})$ is invertible. Such an automorphism is said to be *linear*; thus the linear automorphisms are precisely those that stabilize kX . An automorphism is said to be *affine* if it stabilizes $k + kX$. Clearly any k -linear automorphism of $k + kX$ fixing k defines an affine automorphism. An affine automorphism has the form

$$x \mapsto \sum \alpha_{xx'} x' + \beta_x. \quad (4)$$

Let x_0 be a specified element of X . For any $f \in k\langle X \setminus x_0 \rangle$ there is an automorphism of $k\langle X \rangle$ sending x_0 to $x_0 + f$ and fixing $X \setminus x_0$; it is called an $X \setminus x_0$ -based *shear*. Somewhat more generally, an automorphism of $k\langle X \rangle$ stabilizing $k\langle X \setminus x_0 \rangle$ is called an $X \setminus x_0$ -based *de Jonquières automorphism*; in such an automorphism, $x_0 \mapsto \lambda x_0 + f$, $\lambda \in k^\times$, $f \in k\langle X \setminus x_0 \rangle$. Shears and de Jonquières automorphisms of $k[X]$ are defined similarly.

Let

$$\alpha: x \mapsto f_x, \quad x \in X, \quad (5)$$

be any endomorphism of R . If $f_x(0) = 0$ for all $x \in X$, then α is said to be *augmentation preserving* (because it preserves the augmentation ideal XR). If α is given by (5) and τ_a is the translation defined by (2), then

$$\tau_a \alpha : x \mapsto f_x + a_x,$$

and for a suitable choice of $a \in k^X$, viz. $a_x = -f_x(0)$, we can ensure that $\tau_a \alpha$ is augmentation preserving. This remark is sometimes used to effect a reduction to augmentation preserving endomorphisms, as in the proof of our first result.

PROPOSITION 8.1. *Let X be a finite set. If ψ is a surjective endomorphism of $k\langle X \rangle$, then ϕ is an automorphism.*

Proof. By composing ϕ with a translation we may take it to be augmentation preserving. Suppose that ϕ is not injective and let $0 \neq w \in \ker \phi$. Put $d = d(w)$ and denote by $\mathfrak{n}$ the set of elements of $R = k\langle X \rangle$ of order greater than d . Clearly $\mathfrak{n}$ is an ideal, and $\mathfrak{n}\phi \subseteq \mathfrak{n}$, because ϕ is augmentation preserving. Therefore ϕ induces an endomorphism $\phi_1 : R/\mathfrak{n} \rightarrow R/\mathfrak{n}$, which like ϕ is surjective. Since X is finite, $R/\mathfrak{n}$ is finite-dimensional over k , and so ϕ_1 is an automorphism, but $w\phi_1 = w\phi = 0$, a contradiction, which shows ϕ to be injective, hence an automorphism. ■

A corresponding result holds for $k[X]$. We remark that any automorphism of $k\langle X \rangle$ induces an automorphism of $k[X]$ in a natural way, and the resulting mapping

$$\text{Aut } k\langle X \rangle \rightarrow \text{Aut } k[X]$$

is easily seen to be a homomorphism. Soon we shall see that for $|X| \leq 2$ it is an isomorphism. We shall need another definition. An automorphism of $k\langle X \rangle$ or of $k[X]$ is called *tame* if it can be obtained by composing affine automorphisms and shears; all other automorphisms are called *wild*. No automorphisms of $k\langle X \rangle$ or $k[X]$ are actually known to be wild, although some surely exist when $|X| \geq 3$, but the following are some examples of automorphisms not known to be tame.

1. (Anick) Let $X = \{x, y, z\}$; for any $p \in k\langle X \rangle$ the endomorphism $(x + yp, y, z + py)$ fixes $xy - yz$. Hence for $p = xy - yz$ we obtain an automorphism with inverse $(x - yp, y, z - py)$.

2. The automorphism of 1. has tame image in $k[X]$, but the following example (due to Nagata) is not known to be tame. Let $X = \{x, y, z\}$; for any $p \in k[X]$, the endomorphism $(x + zp, y + 2xp + zp^2, z)$ fixes $x^2 - yz$, so for $p = x^2 - yz$ we get an automorphism with inverse $(x - zp, y - 2xp + zp^2, z)$.

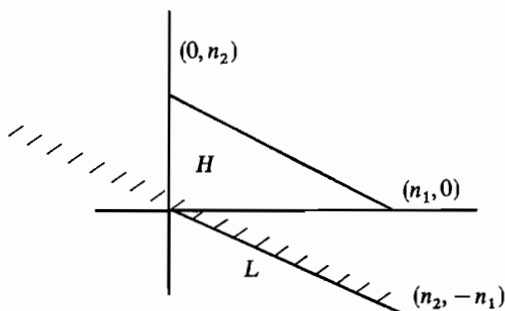
3. (Nagata, Anick) $X = \{w, x, y, z\}$, $R = k\langle X \rangle$ or $k[X]$. For any $p \in R$, the endomorphism $(w, x + pz, y + wp, z)$ fixes $wx - yz$, so it is an automorphism for $p = wx - yz$. It is wild as $k[w, z]$ -automorphism (centralizing w, z).

Our first object will be to prove that every automorphism of $k[x_1, x_2]$ is tame and to obtain a structure theorem for $\text{Aut } k[x_1, x_2]$. We shall need to use a 'weighted' degree in which x_i is of degree n_i ; some further definitions are needed.

Let n_1, n_2 be two integers not both zero. We order $\mathbb{Z}^2$ as an abelian group by putting $(a_1, a_2) \geq (0, 0)$ if either $a_1 n_1 + a_2 n_2 > 0$ or $a_1 n_1 + a_2 n_2 = 0$ and $a_1 n_2 \geq 0 \geq a_2 n_1$. Then the mapping from $\mathbb{Z}^2$ to $\mathbb{Z}$ given by

$$\lambda: (a_1, a_2) \mapsto a_1 n_1 + a_2 n_2$$

is a homomorphism of ordered groups.



Let L be the half-line in $\mathbb{R}^2$ through 0 and $(n_2, -n_1)$, and H the half-plane of points (a_1, a_2) such that $a_1 n_1 + a_2 n_2 \geq 0$. Consider any k -module V graded by $\mathbb{Z}^2$:

$$V = \bigoplus V_a, \quad a = (a_1, a_2) \in \mathbb{Z}^2.$$

Each $f \in V$ has a unique expression $f = \sum f_a$, where $f_a \in V_a$ and almost all are 0. For any $f \neq 0$ we define the (n_1, n_2) -degree $d(f)$ and the (n_1, n_2) -bidegree $D(f)$ on V by

$$d(f) = \max \{a\lambda \mid f_a \neq 0\}, \quad D(f) = \max \{a \mid f_a \neq 0\},$$

and we define the *leading form* $|f|$ and the *top component* $\|f\|$ of f as

$$|f| = \sum \{f_a \mid a\lambda = d(f)\}, \quad \|f\| = f_{D(f)}.$$

We also put $d(0) = D(0) = -\infty$, $|0| = \|0\| = 0$; then $|\cdot|, \|\cdot\|$ are multiplicative on R , while $-d, -D$ are valuations. More precisely, we have

$$d(f-g) \leq \max \{d(f), d(g)\},$$

with strict inequality if and only if $|f| = |g| \neq 0$. Further, we shall write $|S| = \{|f| \mid f \in S\}$ and similarly for $\|S\|$. We observe that $D, |\cdot|, \|\cdot\|$ depend only on the ordering of $\mathbb{Z}^2$.

If $D(f) = (b_1, b_2)$, then $d(f) = D(f)\lambda = b_1n_1 + b_2n_2$. For example, $R = k[x_1, x_2]$ has a $\mathbb{Z}^2$ -grading in which $R_a = R_{(a_1, a_2)}$ has a k -basis $x_1^{a_1}x_2^{a_2}$ for $(a_1, a_2) \in \mathbb{N}^2$, and $R_a = 0$ otherwise. It is clear how the notions of degree, leading form and top component may be extended to the field of fractions $k(x_1, x_2)$ by defining for $q = f/g$, $d(q) = d(f) - d(g)$, $D(q) = D(f) - D(g)$, $\|q\| = \|f\|/\|g\|$, $|q| = |f|/|g|$.

If $|q| = q$, we shall say that q is (n_1, n_2) -homogeneous. Let us write $n_i = hn'_i$, where h is the positive HCF of n_1 and n_2 , and choose (m_1, m_2) such that $m_1n'_1 + m_2n'_2 = 1$ (so $(m_1, m_2) > (0, 0)$), then on writing $x'_1 = x_1^{n'_1}x_2^{-n'_1}$, $x'_2 = x_1^{m_1}x_2^{m_2}$ we have $x_1 = x_1^{m_1}x_2^{n'_1}$, $x_2 = x_1^{-m_1}x_2^{n'_2}$. The set of homogeneous elements of degree hd is $k(x'_1)x_2^d$, so the homogeneous elements all lie in $k(x'_1)[x'_2]$. We shall call the $(1, 0)$ -degree the x_1 -degree, the $(0, 1)$ -degree the x_2 -degree and the $(1, 1)$ -degree the total degree. Further, we shall abbreviate $k[f, f^{-1}]$ by $k[f^{\pm 1}]$; similarly for $k[f^{\pm 1}, g]$, etc.

LEMMA 8.2. *If f_1, f_2 are non-zero (n_1, n_2) -homogeneous elements of $k(x_1, x_2)$ that are algebraically dependent over k , then*

$$f_1^{d(f_2)} = \lambda f_2^{d(f_1)} \quad \text{for some } \lambda \in k^\times.$$

For f_1, f_2 , being (n_1, n_2) -homogeneous, satisfy a homogeneous relation, so $\lambda = f_1^{d(f_2)}f_2^{-d(f_1)}$ is algebraic over k ; since k is relatively algebraically closed in $k(x_1, x_2)$, it follows that $\lambda \in k$. ■

To show that the automorphisms of $k[x_1, x_2]$ are tame we need two lemmas.

LEMMA 8.3. *Let n_1, n_2 be two integers not both 0, and consider $R = k[x_1, x_2]$ with the (n_1, n_2) -degree function. For any $p_1, p_2 \in k(x_1, x_2)$ such that $d(p_1) \neq 0$ there exist $q_1, q_2 \in k(x_1, x_2)$ such that*

$$|k[p_1, p_2]| \subseteq k[q_1^{\pm 1}, q_2] \text{ and } |p_1| = \lambda q_1^a, \quad (6)$$

for some $\lambda \in k^\times$ and $a \in \mathbb{N}_+$.

We note that $|k[p_1^{\pm 1}, p_2]| \subseteq k[q_1^{\pm 1}, q_2]$, since $|p_1^{-1}|$ is in the right-hand side, by the supplementary condition in (6).

Proof. Choose q_1 such that $|p_1| = \lambda q_1^a$, where $\lambda \in k^\times$ and where $a \in \mathbb{N}_+$ is the greatest possible. Since $ad(q_1) = d(p_1) \neq 0$, there are only finitely many possibilities for a and a greatest one exists. To find q_2 we use induction on a ; clearly $a > 0$.

If $|p_1|, |p_2|$ are algebraically independent over k , then in any polynomial in p_1, p_2 the part involving the leading forms does not vanish, so a polynomial in the p 's which is homogeneous can be expressed in terms of the leading forms alone, i.e. $|k[p_1, p_2]| \subseteq k[|p_1|, |p_2|]$. Hence we can then take $q_2 = |p_2|$. There remains the case when $|p_1|, |p_2|$ are algebraically dependent over k ; since they are (n_1, n_2) -homogeneous, $|p_1|^{d(p_2)}$ is a scalar multiple of $|p_2|^{d(p_1)}$, by Lemma 8.2.

It follows that $|p_2|^{d(p_1)}$ is a scalar multiple of $q_1^{ad(p_2)}$, and since q_1 is not a proper power, we must have $|p_2| = \mu q_1^b$, for some $\mu \in k^\times$, $b \in \mathbb{Z}$. We distinguish two cases: (i) If $a \nmid b$, say $b = ra + a'$, $0 < a' < a$, we replace p_1 by $p_2 p_1^{-r}$ and p_2 by p_1 . Now $|p_2 p_1^{-r}| = \mu' q_1^{a'}$, say, so by induction on a there exists q_2 such that $|k[p_2 p_1^{-r}, p_1]| \subseteq k[q_1^{\pm 1}, q_2]$, hence we have

$$k[q_1^{\pm 1}, q_2] \supseteq |k[p_2 p_1^{-r}, p_1^{\pm 1}, (p_2 p_1^{-r})^{-1}]| = |k[p_1^{\pm 1}, p_2^{\pm 1}]| \supseteq |k[p_1, p_2]|.$$

(ii) If $a|b$, then $|p_2| \in k[|p_1|^{\pm 1}]$. If there exists $h \in k[|p_1|^{\pm 1}]$ such that $|p_2 - h| \notin k[|p_1|^{\pm 1}]$, then we can replace p_2 by $p_2 - h$ without changing $k[p_1^{\pm 1}, p_2]$ and by the previous case there is now a suitable q_2 .

The only remaining possibility is that $|p_2 - h| \in k[|p_1|^{\pm 1}]$ for all $h \in k[|p_1|^{\pm 1}]$ (this means that p_2 can be written as a Laurent series in $|p_1|$ or $|p_1|^{-1}$, but we shall not need this fact). In particular, if $|p_2| = \mu |p_1|^b$ for $\mu \in k^\times$, $b \in \mathbb{Z}$, then $d(p_2 - \mu p_1^b) < d(p_2)$, so by repeatedly replacing p_2 by $p_2 - h$ for suitable $h \in k[|p_1|^{\pm 1}]$ we may assume that $d(p_2) < -m$ for any given $m \in \mathbb{N}$. Replacing p_1, q_1 by p_1^{-1}, q_1^{-1} if necessary, we may assume that $d(p_1) < 0$. Hence if $h \in k[p_1, p_2]$ and $d(h) > -m$, then p_2 cannot occur in h , so $|h| \in k[|p_1|]$. Since this holds for all m , we have $|k[p_1, p_2]| \subseteq k[|p_1|]$ and so we may take $q_2 = 0$. ■

LEMMA 8.4. Put $R = k[x_1, x_2]$ and let $(f_1, f_2) \in \text{Aut } R$, where d_i is the x_i -degree of f_i . Then $d_1 | d_2$ or $d_2 | d_1$ and if R has the (d_2, d_1) -degree function, then

$$|f_1| = \lambda_1 x_1^{d_1} + \dots + \lambda_2 x_2^{d_2}. \quad (7)$$

Moreover, if $d_i \neq 0$, then $\lambda_i \neq 0$ ($i = 1, 2$).

Proof. The result is clear if d_1 or d_2 is 0, so we may assume that $d_1 d_2 \neq 0$. Then clearly $d(f_1) \neq 0$ and once (7) is established, the remark following it is easily checked: if $\lambda_2 = 0$, say, then (7) must include a term $\mu x_1^{a_1} x_2^{d_2}$ with $a_1 > 0$, which is a contradiction because $d_2 a_1 + d_1 d_2 > d_1 d_2$.

Now by Lemma 8.3, there exist $q_1, q_2 \in k(x_1, x_2)$ such that $|k[f_1, f_2]| \subseteq k[q_1^{\pm 1}, q_2]$ and $|f_1| = \lambda q_1^a$ for some $\lambda \in k^\times$, $a \in \mathbb{N}_+$. In particular, $|x_i| = x_i \in k[q_1^{\pm 1}, q_2]$, therefore $D(k[q_1^{\pm 1}, q_2])$ contains $D(x_1) = (1, 0)$ and $D(x_2) = (0, 1)$ and so includes $\mathbb{N}^2$. We distinguish two cases.

Case 1. $D(q_1), D(q_2)$ are $\mathbf{Z}$ -independent. Then

$$\mathbf{N}^2 \subseteq D(k[q_1^{\pm 1}, q_2]) = \mathbf{Z} \cdot D(q_1) + \mathbf{N} \cdot D(q_2) = \mathbf{Z} \cdot (1/a)D(f_1) + \mathbf{N} \cdot D(q_2),$$

and this is a half-plane. In particular, $a^{-1}D(f_1), D(q_2)$ is a $\mathbf{Z}$ -basis of $\mathbf{Z}^2$ and $a^{-1}D(f_1)$ does not lie in the interior of the first quadrant $\mathbf{N}^2$ (because a half-plane through it contains $\mathbf{N}^2$). So it lies on one of the axes and therefore $a^{-1}D(f_1)$ is $(1, 0)$ or $(0, 1)$.

(1.a) $a^{-1}D(f_1) = (1, 0)$, so x_1^a has a non-zero coefficient in f_1 . By the construction of $D(f_1)$, $a = x_1 \cdot \deg f_1 = d_1$, so $x_1^{d_1}$ has a non-zero coefficient in f_1 and (7) holds, with $\lambda_1 \neq 0$.

(1.b) $a^{-1}D(f_1) = (0, 1)$; this can be treated similarly. Now

$$D(f_1) = \max\{D(x_1^{d_1}), D(x_2^{d_2})\} = D(x_1^{d_1})$$

by the choice of ordering, so in fact, (1.b) does not arise. Continuing with case 1, we have $D(f_1) = (d_1, 0)$, hence $a = d_1$ and $|f_1| = \lambda q_1^{d_1}$, therefore $d_2 = x_2 \cdot \deg f_1 = x_2 \cdot \deg q_1^{d_1}$, and this is divisible by d_1 .

Case 2. $D(q_1), D(q_2)$ are $\mathbf{Z}$ -dependent. Then they generate a cyclic subgroup of $\mathbf{Z}^2$, say

$$D(q_r) = j_r(i_1, i_2), \quad k_1 j_1 + k_2 j_2 = 1, \quad j_r, i_r \geq 0 \quad (r = 1, 2).$$

Hence $D(q_1^{j_1}) = j_1 j_2(i_1, i_2) = D(q_2^{j_2})$, so there exists a unique $\mu \in k^\times$ such that $D(q_1^{j_1} - \mu q_2^{j_2}) = (\tilde{i}_1, \tilde{i}_2) < j_1 j_2(i_1, i_2)$.

But $q_1^{j_1} - \mu q_2^{j_2}$ is (d_2, d_1) -homogeneous, hence

$$d(q_1^{j_1} - \mu q_2^{j_2}) = \tilde{i}_1 d_2 + \tilde{i}_2 d_1 = j_1 j_2(i_1 d_2 + i_2 d_1).$$

If $(i_1, i_2), (\tilde{i}_1, \tilde{i}_2)$ were $\mathbf{Z}$ -dependent, then this last equation would entail $(\tilde{i}_1, \tilde{i}_2) = j_1 j_2(i_1, i_2)$, which is not the case; so (i_1, i_2) and $(\tilde{i}_1, \tilde{i}_2)$ are $\mathbf{Z}$ -independent. Put $q'_1 = q_1^{k_1} q_2^{k_2}, q'_2 = q_1^{j_2} q_2^{-j_1}$, then $k[q_1^{\pm 1}, q_2^{\pm 1}] = k[q_1^{\pm 1}, q'_2^{\pm 1}]$ and so

$$\begin{aligned} \mathbf{N}^2 \subseteq D(k[q_1^{\pm 1}, q_2^{\pm 1}]) &= D(k[q_1^{\pm 1}, q'_2^{\pm 1}]) \\ &= D(k[q'_1, q'_2]) + \mathbf{Z}(i_1, i_2) \\ &= D(k[q'_1, q'_2 - \mu]) + \mathbf{Z}(i_1, i_2) \\ &= \mathbf{N}(\tilde{i}_1, \tilde{i}_2) + \mathbf{Z}(i_1, i_2) \\ &= \mathbf{Z} \cdot (j_1 a)^{-1} D(f_1) + \mathbf{N}(\tilde{i}_1, \tilde{i}_2). \end{aligned}$$

Now as in case 1 we find $|f_1| = \lambda_1 x_1^{d_1} + \dots + \lambda_2 x_2^{d_2}, \lambda_1 \lambda_2 \neq 0, D(f_1) = (d_1, 0), j_1 a = d_1$; so $D(q_r) = (j_r, 0), d(q_r) = j_r d_2 \quad (r = 1, 2)$. Since q_1, q_2 are (d_2, d_1) -homogeneous and x_2 is a (d_2, d_1) -homogeneous element of $k[q_1^{\pm 1}, q_2]$, we see

that

$$d(x_2) \in d(k[q_1^{\pm 1}, q_2]) \subseteq \mathbb{Z} \cdot d(q_1) + \mathbb{N} \cdot d(q_2) \subseteq \mathbb{Z}d_2.$$

Since $d(x_1) = d_1$, we conclude that $d_2 | d_1$, as claimed. ■

We now come to the basic Jung–van der Kulk theorem; here $\deg = d_{(1,1)}$ denotes the total degree.

THEOREM 8.5. *Every automorphism of $k[x_1, x_2]$ is tame. More precisely, if (g_1, g_2) is an automorphism of $k[x_1, x_2]$ such that $\deg g_1 \geq \deg g_2$, say, then either (g_1, g_2) is affine or there exists a unique $\mu \in k^\times$ and $d \in \mathbb{N}_+$ such that $\deg(g_1 - \mu g_2^d) < \deg g_1$.*

Proof. If $\deg g_1 = 1$, then $\deg g_2 = 1$ and (g_1, g_2) is affine; so we may assume that $\deg g_1 > 1$. Let (f_1, f_2) be the inverse automorphism; this is likewise not affine. Write $d_r = x_r - \deg f_1$ ($r = 1, 2$) and take $k[x_1, x_2]$ with the (d_2, d_1) -degree function. By Lemma 8.4 we have

$$|f_1| = \lambda_1 x_1^{d_1} + \dots + \lambda_2 x_2^{d_2} \text{ and if } \lambda_r = 0, \text{ then } d_r = 0 \quad (r = 1, 2). \quad (8)$$

Now $f_1(g_1, g_2) = x_1$. If $\deg g_1^{d_1} \neq \deg g_2^{d_2}$, then

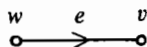
$$\deg(f_1(g_1, g_2)) = \max\{\deg g_1^{d_1}, \deg g_2^{d_2}\}$$

by (8). Since this is 1, we see that $d_1 = 0$ and f_1, g_2 are linear polynomials, so $f_1 \in k + kx_2, g_2 \in k + kx_1$.

Since (f_1, f_2) is not affine, it follows that f_2 is not linear and working with f_2 in place of f_1 we may assume that $\deg g_1^{d_1} = \deg g_2^{d_2}$, so $d_1 d_2 \neq 0, \lambda_1 \lambda_2 \neq 0$. Since $\deg g_1 \geq \deg g_2$, we have $d_1 \leq d_2$, so by Lemma 8.4, $d_1 | d_2$, say $d_2 = dd_1, d \in \mathbb{N}_+$ and $\deg g_1 = \deg g_2^d$.

Let us write $\bar{\phi}$ for the leading $(1, 1)$ -form of ϕ (to be distinguished from $|\phi|$, the leading (d_2, d_1) -form). Since $f_1(g_1, g_2) = x_1$, it follows that $\bar{g}_1$ and $\bar{g}_2$ are algebraically dependent over k , hence by Lemma 8.2, $\bar{g}_1 = \mu \bar{g}_2^d$ ($\mu \in k^\times$), and so $\deg(g_1 - \mu g_2^d) < \deg g_1$, as claimed. Now an easy induction on the degree shows every automorphism to be tame. ■

To study the structure of $\text{Aut } k[x_1, x_2]$ more closely we shall construct a graph (actually a tree) on which it acts. We take Γ to be an oriented graph whose vertices are the k -subspaces of $R = k[x_1, x_2]$, while the edges are inclusion mappings. Now $\text{Aut } R$ acts on Γ in a natural way on the right, while an action on the left is defined by $gp = pg^{-1}$ for any vertex or edge p of Γ and any $g \in \text{Aut } R$. Consider the subgraph T_0 shown in the figure, where



$w = k + kx_2$, $v = k + kx_1 + kx_2$ and e is the inclusion mapping. Let T be the subgraph generated from T_0 by $\text{Aut } R$, i.e. the union of the orbits under $\text{Aut } R$ of v , w and e . Each vertex of T has the form $k + kg_2$ or $k + kg_1 + kg_2$, where $(g_1, g_2) \in \text{Aut } R$, and any two vertices of T are neighbours, i.e. connected by an edge, only if one is 2-dimensional and the other 3-dimensional. For each vertex of T we define its *degree* by the formula

$$\deg(k + kg_2) = \deg g_2, \quad \deg(k + kg_1 + kg_2) = -\frac{1}{2} + \max\{\deg g_1, \deg g_2\}.$$

Clearly these degrees are well-defined (even though g_1, g_2 are not uniquely determined by the subspaces in which they lie), and for any vertex u we have

$$\deg u \equiv \frac{1}{2} \cdot \dim u \pmod{1}.$$

This shows in particular that neighbouring vertices cannot have the same degree. If $k + kg_2$ is any vertex of T not contained in $v = k + kx_1 + kx_2$, then there exists $g_1 \in R$ such that $(g_1, g_2) \in \text{Aut } R$ and by Th. 8.5 we may assume that $\deg g_1 < \deg g_2$, by taking g_1 of minimal degree such that $(g_1, g_2) \in \text{Aut } R$.

Let us take any vertex of T and find which of its neighbours have smaller degree. The neighbours of $k + kx_2$ are all of the form $k + k(x_1 + p) + kx_2$, where $p \in k[x_2]$, hence the neighbours of $k + kg_2$ are of the form $k + k(g_1 + p) + kg_2$, where $p \in k[g_2]$. The only neighbour with smaller degree is $k + kg_1 + kg_2$, all others have greater degree, because $p \notin k + kg_2$ unless p is linear, and then $k + k(g_1 + \lambda g_2) + kg_2 = k + kg_1 + kg_2$. Next consider a vertex $k + kg_1 + kg_2$, where $\deg g_2 < \deg g_1$, say. Its neighbours are of the form $k + k(\lambda_1 g_1 + \lambda_2 g_2)$, where λ_1, λ_2 are in k and not both zero. The only neighbour of smaller degree is $k + kg_2$, all others have greater degree. The only remaining vertex in T is $v = k + kx_1 + kx_2$ (by Th. 8.5); all its neighbours have greater degree. Thus from every vertex of T other than v there is a unique edge to a vertex of lower degree; if we proceed so as to diminish the degree at each step, we have a unique choice each time until we reach v . This shows T to be connected, non-empty and there can be no circuit, i.e. T is a *tree*. For if we had a circuit and took a vertex of highest degree, there would only be one edge of descent, which is a contradiction.

We shall use the tree T to prove a structure theorem for the automorphism group of $k[x_1, x_2]$. Let A, B be two groups with a common subgroup C , i.e. there are specified embeddings $C \rightarrow A$, $C \rightarrow B$. By the *free product* of A and B amalgamating C , denoted by $A \ast_C B$, we understand the group presented on a generating set consisting of the disjoint union of A and B , with defining relations consisting of the relations of A, B and identifying the two copies of C .

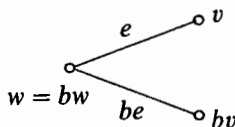
THEOREM 8.6. Every automorphism of $k[x_1, x_2]$ is tame and

$$\text{Aut } k[x_1, x_2] = A \sharp B,$$

where A is the group of affine automorphisms $\text{Aff } k[x_1, x_2]$, B is the group of x_2 -based de Jonquières automorphisms and $C = A \cap B$ consists of all automorphisms $(\alpha x_1 + \beta x_2 + \mu, \gamma x_2 + \nu)$.

Proof. Let T be the tree constructed above. We observe that the stabilizer of $v = k + kx_1 + kx_2$ is A , that of $w = k + kx_2$ is B and that of e is $A \cap B = C$.

By construction every edge of T has the form ge , $g \in \text{Aut } R$. The vertices of ge are gv, gw , so if ge has w as vertex, then $gw = w$, i.e. $g \in B$. This shows that the neighbours of w are $bw, b \in B$. Similarly, the neighbours of v are $av, a \in A$.



Let G be the subgroup of $\text{Aut } R$ generated by $A \cup B$; then $Ge \cup Gv \cup Gw$ is a subgraph of T closed under passage to neighbours, so it is all of T . Hence for any $g \in \text{Aut } R$, $gv = hv$ for some $h \in G$, therefore $g \in hA \subseteq G$, and so $G = \text{Aut } R$. This shows in particular that all automorphisms of R are tame (but, of course, the construction depended on Th. 8.5).

Consider any sequence $a_0, b_1, a_2, \dots, a_{2n}, b_{2n+1}$, where $a_i \in A$, $b_i \in B$ and write

$$g = a_0 b_1 a_2 \dots a_{2n} b_{2n+1}. \quad (9)$$

The corresponding path from v :

$$\begin{array}{ccccccc} v = a_0 v & & a_0 e & & & & a_0 b_1 e \\ \circ & \text{---} & \circ & \text{---} & \circ & \text{---} & \dots \\ & & a_0 w = a_0 b_1 w & & a_0 b_1 v & & \end{array}$$

terminates at gv . Thus if $gv = v$, some edge must be followed by its inverse, so for some i , $1 \leq i \leq 2n+1$, $a_i e = e$ or $b_i e = e$, i.e. a_i or b_i is in C . We can therefore shorten the relation (9), so any relation of this form holding in G must also hold in $A \sharp B$. A similar argument applies to $b_0 a_1 b_2 \dots b_{2n} a_{2n+1}$, and so we obtain

$$\text{Aut } k[x_1, x_2] = A \sharp B. \blacksquare$$

We note that if $\tau = (x_2, x_1)$ is the automorphism interchanging x_1 and x_2 , then A has the double coset decomposition $A = C \cup C\tau C$, therefore any element of $\text{Aut } R$ can be written as $b_0 \tau b_1 \tau \dots \tau b_n$, where $b_i \in B \setminus C$, $i =$

$1, \dots, n-1, b_0, b_n \in B$. However, this does not mean that $\text{Aut } R$ has the form $B^* \{1, \tau\}$; it is, in fact, a homomorphic image of this free product.

A connected graph Γ is said to have *width* at most n , if any two vertices of Γ can be joined by a path of length at most n ; the least such n is the *width* of Γ . A subgroup G of $\text{Aut } R$ is said to have width at most n if for each $g \in G$ the vertices $v = k + kx_1 + kx_2$ and gv are connected by a path of length at most n . Thus for a group of width n any path (corresponding to an element of the group) has length at most $2n$. It is clear that any finite subgroup of $\text{Aut } R$ has finite width; so the next result applies in particular to finite subgroups of $\text{Aut } R$.

THEOREM 8.7. *Let T be the tree constructed above for $\text{Aut } k[x_1, x_2]$ and let G be a subgroup of the automorphism group; then the following conditions are equivalent:*

- (a) G has finite width,
- (b) G stabilizes a subtree T_0 of T of finite width,
- (c) G fixes a vertex of T ,
- (d) G is conjugate to a subgroup of A or B .

Proof. (a) $\Rightarrow$ (b). Suppose that G has width n and denote by T_0 the smallest subtree of T containing Gv . Then T_0 consists of all the vertices and edges appearing in the minimal paths from v to gv , where $g \in G$. Since each such path has length at most n , T_0 has width at most $2n$, and G stabilizes Gv , hence also T_0 .

(b) $\Rightarrow$ (c). Suppose that G stabilizes a subtree T_0 of width n . If $n \leq 1$, T_0 has at most one edge and G fixes all the vertices of T_0 , so we may assume that $n \geq 2$ and use induction on n .

By an *extremity* of T_0 we shall understand a vertex of T_0 which lies on only one edge of T_0 , together with this edge. Let S_0 be the set of all extremities of T_0 ; since T_0 has more than one edge, $S_0 \neq T_0$ and hence $T_1 = T_0 \setminus S_0$ is a subtree of T_0 of width $n-2$ and G stabilizes S_0 , hence also T_1 . By induction on n it follows that G fixes a vertex of T_1 .

(c) $\Rightarrow$ (a). Suppose that G fixes a vertex u and let n be the length of a minimal path P from v to u ; then for each $g \in G$, gP is a path of length n from gv to $gu = u$, hence $P \cdot (gP)^{-1}$ is a path of length at most $2n$ from v to gv , so G has width at most $2n$.

Finally, (c) $\Leftrightarrow$ (d), because for any automorphism g , $\text{stab } gv = gAg^{-1}$, $\text{stab } gw = gBg^{-1}$. ■

COROLLARY 8.8. *Any finite subgroup of $\text{Aut } k[x_1, x_2]$ is conjugate to an affine group or a group of x_2 -based de Jonquières automorphisms. ■*

This result can be made more precise:

PROPOSITION 8.9. *Let G be a finite subgroup of $\text{Aut } k[x_1, x_2]$ such that $\text{char } k \nmid |G|$. Then G is conjugate to a group of linear automorphisms.*

Proof. The group algebra kG is semisimple by Maschke's theorem (cf. e.g. Cohn [77], p. 356), so every submodule of a kG -module is complemented, and by Th. 8.7 we may assume that $G \subseteq A$ or $G \subseteq B$. If $G \subseteq B$, then $k + kx_2$ is a kG -module with submodule k , so there exists a complement $kx'_2 \cong (k + kx_2)/k$. Now consider $k[x'_2] + kx_1$; this is a kG -module with submodule $k[x'_2]$, so there is a complement kx'_1 . Now G stabilizes $kx'_1 + kx'_2$, hence a conjugate of G stabilizes $kx_1 + kx_2$ and so is a group of linear automorphisms (in fact it acts by 'diagonal' automorphisms in this case).

If $G \subseteq A$, then $k + kx_1 + kx_2$ is a kG -module with a submodule k , so it has a complement $kx'_1 + kx'_2$, therefore a conjugate of G stabilizes $kx_1 + kx_2$ and so acts by linear automorphisms. ■

Exercises 6.8

1. (Nagata [72]) In $k[x, y, z]$ consider the endomorphism $(x + zp, y + 2xp + zp^2, z)$, for any p . Verify that it fixes $x^2 - yz$, and that it is an automorphism for $p = x^2 - yz$. Show that this is tame as $k[z, z^{-1}]$ -algebra automorphism, but not as $k[z]$ -algebra automorphism.

2. (Bergman) Do the same for $k[z] \langle x, y \rangle$, using the automorphism $(x + zp, y + xp + px + zp^2, z)$, where again $p = x^2 - yz$.

3. Let G_1, G_2 be subgroups of $\text{Aut } k[x_1, x_2]$ each fixing a vertex of T (defined as in the text), and suppose that $g_i \in G_i$ exists such that $g_i \notin G_1 \cap G_2$. Show that the subset of T fixed by G_i is a subtree T_i , and if $T_1 \cap T_2 = \emptyset$ and P is the unique shortest path from a vertex of T_1 to one of T_2 , then for each n , the path $P, (g_2P)^{-1}, g_2g_1P, (g_2g_1g_2P)^{-1}, \dots, (g_2g_1)^nP$ is minimal (i.e. no edge is followed by its inverse). Deduce that g_2g_1 has infinite order.

4. Suppose that G is a torsion subgroup of $\text{Aut } k[x_1, x_2]$ which does not lie in a conjugate of A or B . Show that G is the union of a countable chain of subgroups $G_1 \subseteq G_2 \subseteq \dots$, each conjugate to a subgroup of C . (This is usually expressed by saying that G stabilizes either a vertex of T or an *end* of T , i.e. a path to ∞ ; the ends of T may be thought of as the vertices at ∞ .) Deduce that every torsion subgroup of $\text{Aut } k[x_1, x_2]$ is locally finite.

5°. Find a wild automorphism of $k[x, y, z]$.

6.9 Automorphisms of free algebras

We now turn to the free algebra $R = k\langle x_1, x_2 \rangle$; it has again a natural $\mathbb{Z}^2$ -grading $R = \bigoplus R_a$, where R_a ($a = (a_1, a_2) \in \mathbb{N}^2$) is the subspace spanned by all

monomials in x_1, x_2 whose image in the polynomial ring is $x_1^{a_1}x_2^{a_2}$. From the definitions given earlier we see that for any integers n_1, n_2 not both zero we have an (n_1, n_2) -degree $d = d_{(n_1, n_2)}: R \rightarrow \mathbb{Z} \cup \{-\infty\}$ and a bidegree $D = D_{(n_1, n_2)}: R \rightarrow \mathbb{Z}^2 \cup \{-\infty\}$, leading form $|f|$ and top component $\|f\|$, for any $f \in R$. We note that if $\|f_1\|, \|f_2\|$ are algebraically independent, i.e. if they freely generate a free algebra, then

$$\|k\langle f_1, f_2 \rangle\| \subseteq k\langle \|f_1\|, \|f_2\| \rangle,$$

hence

$$D(k\langle f_1, f_2 \rangle) \subseteq N \cdot D(f_1) + N \cdot D(f_2). \quad (1)$$

In a free algebra any two dependent elements can be written as polynomials in a (suitably chosen) third element, by Cor. 7.4 and Th. 7.7. Here we shall need this result only for homogeneous elements, in which case it is easily proved directly:

LEMMA 9.1. *Let X be a set, $v: X \rightarrow \mathbb{N}_+$ a mapping and $R = k\langle X \rangle$ the free algebra with grading v . If $f_1, f_2 \in R$ are homogeneous in this grading, then either f_1, f_2 form a free set in R or they are both scalar multiples of powers of some homogeneous element f of R , i.e. $f_1, f_2 \in k[f]$.*

Proof. We may assume that $f_1, f_2 \notin k$. Suppose that f_1, f_2 are dependent and choose $0 \neq g \in k\langle x_1, x_2 \rangle$ such that $g(f_1, f_2) = 0$. By putting $X = 0$ we see that g has zero constant term, so we can write it as $g = x_1g_1 + x_2g_2$, therefore we have

$$f_1g_1 + f_2g_2 = 0.$$

Let $v(f_1) \geq v(f_2)$, say; by the weak algorithm there exists $f'_1 \in R$ such that $f_1 = f_2f'_1$, and f'_1, f_2 are dependent because f_1, f_2 are. By induction, $f'_1, f_2 \in k[f]$ for some f , hence $f_1, f_2 \in k[f]$. ■

We now come to the proof that all automorphisms of $k\langle x_1, x_2 \rangle$ are tame; it is convenient to prove first a somewhat technical lemma about endomorphisms which will enable us to reduce the tameness proof to that of the corresponding result for $k[x_1, x_2]$ (Th. 8.5), and which is used again later. If (f_1, f_2) and (g_1, g_2) are endomorphisms of $k\langle x_1, x_2 \rangle$ which induce the same endomorphism of $k[x_1, x_2]$, we shall write $(f_1, f_2) \equiv_{ab} (g_1, g_2)$. This amounts to requiring $f_1 - g_1, f_2 - g_2$ to lie in the ideal generated by $[x_1, x_2] = x_1x_2 - x_2x_1$.

LEMMA 9.2. *Let $f_1, f_2 \in k\langle x_1, x_2 \rangle$ be such that $(f_1, f_2) \neq (x_1, x_2)$, $(f_1, f_2) \equiv_{ab} (x_1, x_2)$. Then there exist $n_1 \leq 0 \leq n_2$, not both zero, such that if $k\langle x_1, x_2 \rangle$ is given*

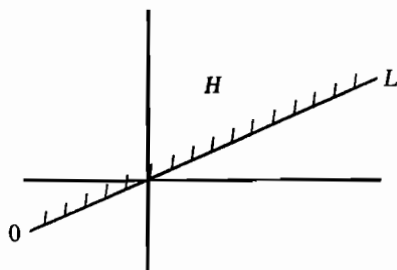
the (n_1, n_2) -degree function, then $D(f_1) + D(f_2) \neq D(x_1) + D(x_2) = (1, 1)$ and $(d(f_1), d(f_2)) = (d(x_1), d(x_2)) = (n_1, n_2)$, so $\|f_1\|, \|f_2\|$ are independent.

Moreover, (f_1, f_2) is not an automorphism and $[f_1, f_2] \neq \lambda[x_1, x_2]$ for all $\lambda \in k^\times$.

Proof. For each $f \in R = k\langle x_1, x_2 \rangle$ we have an expression $f = \sum f_a, f_a \in R_a$. We define the support of f as a subset of $\mathbb{Z}^2$ by

$$\text{supp}(f) = \{a \in \mathbb{Z}^2 \mid f_a \neq 0\},$$

and also put $\text{supp}(x_1^{-1}f) = \{(a_1 - 1, a_2) \mid f_a \neq 0\}$, and similarly for $\text{supp}(x_2^{-1}f)$. Let S be the union of $\text{supp}(x_1^{-1}f_1)$ and $\text{supp}(x_2^{-1}f_2)$ with the origin removed. Since $(f_1, f_2) \equiv (x_1, x_2)$, S lies in the first quadrant. We choose $(b_1, b_2) \in S$ to maximize $b_2/b_1 \in \mathbb{Q} \cup \{\infty\}$ and put $n_1 = -b_2, n_2 = b_1$, so that n_1, n_2 are



not both 0 and satisfy $n_1 \leq 0 \leq n_2$. We again have a half-line L in $\mathbb{R}^2$ from the origin through $(n_2, -n_1) = (b_1, b_2)$ and the half-plane H obtained by rotating L counter-clockwise about 0 through all $\theta, 0 \leq \theta < \pi$.

By the maximality of (b_1, b_2) , $H \setminus L$ does not meet S , so all positive elements of S lie on L (e.g. (b_1, b_2)). Moreover, $\text{supp}(x_r^{-1}f_r)$ ($r = 1, 2$) both contain the origin, which is also on L , so the maximum of each is on L and at least one is not the origin. Thus $D(f_r) - D(x_r) \in L$, not both are 0, so we have $D(f_1) + D(f_2) > D(x_1) + D(x_2)$. We also have $d(f_r) - d(x_r) \in d(L) = 0$, so $(d(f_1), d(f_2)) = (d(x_1), d(x_2))$.

If $\|f_1\|, \|f_2\|$ are dependent, they are scalar multiples of powers of some (homogeneous) f , by Lemma 9.1. It follows that $d(f_1), d(f_2)$ are positive multiples of $d(f)$, which contradicts the fact that $d(f_1) \leq 0 \leq d(f_2)$ and not both are 0. Therefore $\|f_1\|, \|f_2\|$ are independent. By the remarks preceding Lemma 9.1 we have (1) and

$$\{D(f_1), D(f_2)\} \neq \{D(x_1), D(x_2)\}, \quad (2)$$

because $D(f_1) + D(f_2) > (1, 1)$.

Suppose that (f_1, f_2) is an automorphism. Then $x_1, x_2 \in k\langle f_1, f_2 \rangle$, hence $D(x_1), D(x_2) \in D(k\langle f_1, f_2 \rangle)$ and so

$$N^2 \subseteq D(k\langle f_1, f_2 \rangle) \subseteq N \cdot D(f_1) + N \cdot D(f_2),$$

by (1), but this contradicts (2). Hence (f_1, f_2) is not an automorphism. Moreover,

$$\|f_1 f_2\| = \|f_1\| \cdot \|f_2\| \neq \|f_2\| \cdot \|f_1\| = \|f_2 f_1\|,$$

hence we have

$$D([f_1, f_2]) = \max\{D(f_1 f_2), D(f_2 f_1)\} = D(f_1) + D(f_2) > D(x_1) + D(x_2) \\ = D([x_1, x_2]),$$

therefore $[f_1, f_2] \neq \lambda[x_1, x_2]$ for any $\lambda \in k^\times$. ■

We can now prove the main result of this section, due to Makar–Limanov and Czerniakiewicz:

THEOREM 9.3. *All automorphisms of $k\langle x_1, x_2 \rangle$ are tame and the natural mapping*

$$\phi : \text{Aut } k\langle x_1, x_2 \rangle \rightarrow \text{Aut } k[x_1, x_2] \quad (3)$$

is an isomorphism.

We begin by proving the second part. Let B again be the set of x_2 -based de Jonquières automorphisms and $\tau = (x_2, x_1)$, then $B \cup \{\tau\}$ is in the image of ϕ . By Th. 8.5 it is surjective, and by Lemma 9.2, $\ker \phi = 1$, so ϕ is an isomorphism and it follows that $\text{Aut } k\langle x_1, x_2 \rangle$ is generated by $B \cup \{\tau\}$. Therefore every automorphism of $k\langle x_1, x_2 \rangle$ is tame. ■

Let us return to the polynomial ring $k[x_1, x_2]$ for a moment. For every endomorphism $\alpha = (f_1, f_2)$ we can define the *Jacobian matrix* $J(\alpha) = (\partial f_i / \partial x_j)$, and it is easily seen that $J(1) = I, J(\alpha\beta) = J(\alpha)^\beta J(\beta)$. It follows that for an automorphism α , $J(\alpha)$ is invertible, hence $\det(J(\alpha))$ is an element of $k^\times$. This suggests the following question:

Jacobian problem. Let α be an endomorphism of $k[x_1, x_2]$ such that $\det(J(\alpha)) \in k^\times$; is α necessarily an automorphism?

No complete answer is known, although there are partial results. But for the free algebra we have an analogue due to Dicks, in which the commutator takes the place of the Jacobian.

THEOREM 9.4. *An endomorphism (f_1, f_2) of $k\langle x_1, x_2 \rangle$ is an automorphism if and only if $[f_1, f_2] = \lambda[x_1, x_2]$, for some $\lambda \in k^\times$.*

Proof. For the necessity it is enough to check that each automorphism of $k\langle x_1, x_2 \rangle$ stabilizes the space kc , where $c = [x_1, x_2]$; this is clear for the generating set $B \cup \{\tau\}$, hence for the whole automorphism group.

Conversely, suppose that the condition holds; by a slight change of notation this may be written

$$\lambda[f_1, f_2] = [x_1, x_2], \quad \lambda \in k^\times.$$

Put $R = k[x_1, x_2]$ and let S be the image of $k\langle f_1, f_2 \rangle$ in R (under abelianization); we shall show that $S = R$. Let $\mathfrak{p}$ be a prime ideal of R and put $\mathfrak{q} = \mathfrak{p} \cap S$, then $\mathfrak{q}$ is a prime ideal of S . Writing $\mathcal{F}(-)$ for the field of fractions of a commutative integral domain, we can form the R -bimodule

$$M = \mathcal{F}(R/\mathfrak{p}) \otimes_{\mathcal{F}(S/\mathfrak{q})} \mathcal{F}(R/\mathfrak{p}).$$

For any $m_1, m_2 \in M$ there is a unique derivation $\partial: R \rightarrow M$ such that $\partial x_i = m_i$. By the definition of M we have $f_i u = u f_i$ for all $u \in M$, i.e. f_i centralizes M , therefore

$$\partial[f_1, f_2] = [\partial f_1, f_2] + [f_1, \partial f_2] = 0,$$

and it follows that

$$0 = \partial[x_1, x_2] = [m_1, x_2] + [x_1, m_2],$$

and here m_1, m_2 are arbitrary in M . Taking $m_1 = 1 \otimes 1, m_2 = 0$, we see that $1 \otimes x_2 = x_2 \otimes 1$ in M , so $1, x_2$ are linearly dependent over $\mathcal{F}(S/\mathfrak{q})$, i.e. $x_2 \in \mathcal{F}(S/\mathfrak{q})$ and similarly $x_1 \in \mathcal{F}(S/\mathfrak{q})$, so $\mathcal{F}(S/\mathfrak{q}) = \mathcal{F}(R/\mathfrak{p})$. It follows that

$$\text{tr. deg}_k \mathcal{F}(S/\mathfrak{q}) = \text{tr. deg}_k \mathcal{F}(R/\mathfrak{p}).$$

Taking $\mathfrak{p} = \mathfrak{q} = 0$, we see that $S = k[f_1, f_2] \cong R$. Further, when $\mathfrak{p}$ is a principal prime ideal in R , then $\mathfrak{q}$ is a principal prime ideal in S .

We have $x_2 \in \mathcal{F}(S)$, say $x_2 = a/b$, where a and b are coprime elements of S . If $b \notin k$, choose a prime factor p of b in R ; then we have $q = bx_2$, where $a, b \in S \cap pR = qS$ for some $q \in S$, and this contradicts the fact that a, b are coprime in S . Hence $b \in k^\times, x_2 \in S$ and similarly $x_1 \in S$, so $S = R$.

Thus (f_1, f_2) induces an automorphism of $k[x_1, x_2]$; by the isomorphism (3) of Th. 9.3 we can lift this back to an automorphism of $k\langle x_1, x_2 \rangle$. Composing its inverse with (f_1, f_2) , we may assume that $(f_1, f_2) \equiv_{ab} (x_1, x_2)$. By Lemma 9.2, $(f_1, f_2) = (x_1, x_2)$, so it is indeed an automorphism. ■

From Th. 9.4 we obtain a natural homomorphism

$$\det: \text{Aut } k\langle x_1, x_2 \rangle \rightarrow k^\times, \quad (4)$$

by mapping (f_1, f_2) to λ if $[f_1, f_2] = \lambda[x_1, x_2]$. The kernel is denoted by $\text{SAut } k\langle x_1, x_2 \rangle$, its elements are the *special* automorphisms. It can be shown that this group is the group of automorphisms of the Weyl algebra $A_1(k)$ (Dixmier [68], Makar-Limanov [70]).

As in the case of the polynomial ring $k[x_1, x_2]$, we can construct an oriented graph Γ for $R = k\langle x_1, x_2 \rangle$, whose vertices are the non-zero finite-dimensional subspaces of R , with inclusion maps as edges. As before we let $\text{Aut } R$ act on Γ in the natural way on the right and define the left action again by $g \cdot u = ug^{-1}$. Further, we put $w = k + kx_2$, $v = k + kx_1 + kx_2$ and let e be the edge between them, and write T for the disjoint union of the orbits of v, w, e . From Th. 8.6 and 9.3 we know that

$$\text{Aut } k\langle x_1, x_2 \rangle \cong \text{Aut } k[x_1, x_2] \cong A_C^* B,$$

where A, B, C are defined as before, and this shows that T is a tree and that $\text{stab } v = A$, $\text{stab } w = B$, $\text{stab } e = C$.

For the rest of this section we shall put $c = [x_1, x_2]$. If u is any vertex in Γ such that $u \subseteq k[c]$, then $\text{stab } u$ is 'large', in the sense that it contains all special automorphisms. We shall prove that for any vertex u of Γ , either $u \subseteq k[c]$ or $\text{stab } u$ lies in a conjugate of A or B .

For any $b \in B$, say $b = (\alpha x_1 + p(x_2), \beta x_2 + \lambda)$ we define the degree of b , $\deg b$ as the degree of p as a polynomial in x_2 ; thus $b \in C$ if and only if $\deg b \leq 1$.

We shall consider the $\mathbb{Z}^2$ -graded vector space $U = k\langle x_1, x_2 \rangle / k[c]$; since $\text{Aut } k\langle x_1, x_2 \rangle$ stabilizes $k[c]$, it acts on U . Any $f \in U$ is said to be *biased* if $f \neq 0$ and $D_{(1,1)}(f)$ is of the form (i_1, i_2) where $i_1 \geq i_2$. We begin with a lemma, due to D. Lane, which gives conditions under which an x_2 -based de Jonquieres automorphism decreases the degree:

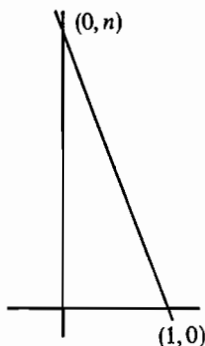
LEMMA 9.5. *Let f be a biased element of U and let $b \in B \setminus C$, $\deg b = n$. We equip U with the $(n, 1)$ -degree function, and write*

$$D_{(1,1)}(|f|) = (r_1, r_2), D_{(1,1)}(|fb|) = (r'_1, r'_2).$$

- (i) *If $f = f_{(r_1, r_2)}$, then $r'_1 \leq \min\{r_1 - 1, r_2\}$.*
- (ii) *If f is $(n, 1)$ -homogeneous (so that $f = |f|$), then $r'_1 < r'_2$.*
- (iii) *$\deg f < r'_1 + r'_2$.*

Hence $\deg f < \deg fb$.

We remark that the $(n, 1)$ -degree function represents the x_2 -degree of fb .



Proof. By the choice of n , b respects the filtration by $(n, 1)$ -degree, therefore b acts linearly on the monomials in the leading form $|f| \pmod{\text{lower terms}}$; in other words, we have $|fb| = |f|b$. We may therefore assume without loss of generality that

$$b = (\alpha x_1 + \gamma x_2^n, \beta x_2), \quad \alpha, \beta, \gamma \in k^\times. \quad (5)$$

For each $i \in \mathbb{Z}$ let b_i be the k -linear endomorphism of $k\langle x_1, x_2 \rangle$ which maps a monomial s of $(1, 0)$ -degree d to the $(1, 0)$ -component of sb , of $(1, 0)$ -degree $d - i$; for $i < 0$, $b_i = 0$. Thus $\sum s b_i$ is just the $(1, 0)$ -component of sb , $sb_i = (sb)_{d-i}$. By linearity this defines b on all of $k\langle x_1, x_2 \rangle$. We remark that b_0 is the automorphism $(\alpha x_1, \beta x_2)$, b_1 is a (b_0, b_0) -derivation and in characteristic 0 we have $b_i = (1/i!)b_1^i$ ($i \geq 1$); thus $(b_0, b_1, \dots)$ is just a higher derivation (cf. 8.7).

Since b is a homomorphism, we have

$$(s_1 s_2) b_i = \sum_{j=0}^i s_1 b_j \cdot s_2 b_{i-j} \quad (6)$$

for all monomials, and hence by linearity for all elements of $k\langle x_1, x_2 \rangle$. It can be verified that

$$b_i b_j = \binom{i+j}{i} b_{i+j} \quad \text{for all } i, j \geq 0, \quad (7)$$

where $\binom{r}{s} \in k$ is the usual binomial coefficient, defined for any $r \in \mathbb{Z}$, $s \in \mathbb{N}$ as the coefficient of t^s in $(1+t)^r$ in $\mathbb{Z}[[t]]$.

(i) We proceed by induction on $r_1 + r_2$. Here fb is $(n, 1)$ -homogeneous (by the form (5) of b) and $r_2 = x_2\text{-deg } fb$. We can lift $f \in U$ back to an element of $k\langle x_1, x_2 \rangle$, again denoted by f , such that $f = f_{(r_1, r_2)}$. If $r_1 > \min\{r_1 - 1, r_2\}$, then $fb_{r_1-j} = 0$ for all $j \leq \min\{r_1 - 1, r_2\}$, hence on writing $i = r_1 - j$, we have $fb_i = 0$ for all $i \geq \max\{r_1 - r_2, 1\}$. Clearly we may assume that $f \notin k + kx_2$,

hence $(r_1, r_2) \neq (0, 0), (0, 1)$. Therefore we can express f in the form

$$f = x_1 f_1 + x_2 x_1 f_2 + x_2^2 f_3.$$

It follows that for all $i \geq 0$,

$$\begin{aligned} f b_i &= x_1 b_0 \cdot f_1 b_i + x_1 b_1 \cdot f_1 b_{i-1} + (x_2 x_1) b_0 \cdot f_2 b_i + (x_2 x_1) b_1 \cdot f_2 b_{i-1} + x_2^2 b_0 \cdot f_3 b_i \\ &= \alpha x_1 \cdot f_1 b_i + \gamma x_2^n \cdot f_1 b_{i-1} + \alpha \beta x_2 x_1 \cdot f_2 b_i + \beta \gamma x_2^{n+1} \cdot f_2 b_{i-1} + \beta^2 x_2^2 \cdot f_3 b_i. \end{aligned}$$

Here $x_1, x_2 x_1$ and x_2^2 are right linearly independent, while for $i \geq \max\{r_1 - r_2, 1\}$, $f b_i = 0$, hence

$$f_1 b_i = 0, \quad (8)$$

$$f_2 b_i = 0, \quad (9)$$

$$\gamma x_2^{n-2} \cdot f_1 b_{i-1} + \beta \gamma x_2^{n-1} \cdot f_2 b_{i-1} + \beta^2 f_3 b_i = 0. \quad (10)$$

By the induction hypothesis and (9), since $f_2 = f_{2(r_1-1, r_2-1)}$, we see that $f_2 \in k[c]$. For $j > r_1 - r_2 + 2 = \max\{r_1 - r_2 + 2, 1\}$, (8) and (9) hold with $i = j - 1$ and (10) holds with $i = j$, so $f_3 b_j = 0$. By induction hypothesis it follows that $f_3 \in k[c]$, because $f_3 = f_{3(r_1, r_2-2)}$. But any element of $k[c]$ has equal x_1 - and x_2 -degrees and $r_1 \neq r_2 - 2$, therefore $f_3 = 0$.

If $r_1 > r_2 + 1$, then $r_1 - 1 \neq r_2 - 1$, so $f_2 = 0$ in that case. Now for all $j \geq \max\{r_1 - r_2 - 1, 1\} = r_1 - r_2 - 1 = \max\{r_1 - r_2, 1\} - 1$, (10) holds for $i = j + 1$, so by induction hypothesis, $f_1 \in k[c]$, because $f_1 = f_{1(r_1-1, r_2)}$. But $r_2 < r_1 - 1$, hence $f = 0$, which is a contradiction. It follows that $r_1 \leq r_2 + 1$ and we can take $i = 1$ in (10). Cancelling γx_2^{n-2} , we get $f_1 b_0 + \beta x_2 \cdot f_2 b_0 = 0$. But $b_0 = (\alpha x_1, \beta x_2)$, so this yields $(f_1 + x_2 f_2) b_0 = 0$, hence $f_1 + x_2 f_2 = 0$, and so $f = x_1 f_1 + x_2 x_1 f_2 = -x_1 x_2 f_2 + x_2 x_1 f_2 = -c f_2 \in k[c]$, which contradicts the fact that f was a biased (non-zero) element of U .

(ii) Let $d = r_1 n + r_2 (= d_{(n,1)}(f))$. Let us abbreviate $f_{(j, d-nj)} (\in U)$ as f_j for any $j \geq 0$, so that $f = \sum_{j \geq r_1} f_j$. Since f is biased, we have $r_1 \geq r_2 = d - r_1 n$, so $r_1 \geq d/(n+1)$. Let s be the greatest integer less than $d/(n+1)$, thus $s < d/(n+1) \leq s+1$, then $r_1 \geq s+1$. Write $r = x_1 \cdot \deg f = \max\{j | f_j \neq 0\}$ and put $D_{(1,1)}(|f, b|) = (q, d - nq)$, where $|f, b|$ is the leading $(n, 1)$ -component of f, b . By (i),

$$\begin{aligned} q &\leq \min\{r-1, d-nr\} \\ &< \frac{n-1}{n+1} \cdot r + \frac{2}{n+1} (d-nr) \\ &= -r + \frac{2d}{n+1} \\ &\leq -r + 2s + 2. \end{aligned}$$

Therefore $q \leq -r+2s+1$ and $f_r b_{r-q} \neq 0$, by the definition of q . We write $(fb)_i$ for $(fb)_{(i, d-ni)}$ and consider

$$\begin{aligned} \sum_{i=q}^s \binom{s-r}{s-i} ((fb)_i) b_{i-q} &= \sum_{i=q}^s \binom{s-r}{s-i} \left\{ \sum_{j=s+1}^r f_j b_{j-i} \right\} b_{i-q} \quad \text{since } r_1 \geq s+1 \\ &= \sum_{i=q}^s \binom{s-r}{s-i} \left\{ \sum_{j=s+1}^r \binom{j-q}{i-q} f_j b_{j-q} \right\} \\ &= \sum_{j=s+1}^r \left\{ \sum_{i=q}^s \binom{s-r}{s-i} \binom{j-q}{i-q} \right\} f_j b_{j-q} \\ &= \sum_{j=s+1}^r \binom{(s-q)-(r-j)}{s-q} f_j b_{j-q} \end{aligned} \quad (11)$$

Here the final step follows by Vandermonde's identity:

$$\sum_i \binom{\mu}{i} \binom{\nu}{k-i} = \binom{\mu+\nu}{k},$$

which is obtained by equating the coefficients of t^k in the identity $(1+t)^{\mu+\nu} = (1+t)^\mu (1+t)^\nu$.

Since $s-q \geq r-(s+1)$, the expression (11) is equal to $f_r b_{r-q}$, which is non-zero. It follows that $(fb)_i \neq 0$ for some $i \leq s < d/(n+1)$. But r'_1 was defined as $\min\{i \mid (fb)_i \neq 0\}$, hence $r'_1 < d/(n+1)$, and so $r'_1 < d-nr'_1 = r'_2$, as claimed.

(iii) Let $D_{(1,1)}(f) = (i_1, i_2)$; then

$$i_1 + i_2 \geq r_1 + r_2, \quad (12)$$

because $i_1 + i_2$ is the highest total degree occurring in f . But by definition of (r_1, r_2) , $ni_1 + i_2 \leq nr_1 + r_2$. Hence

$$\begin{aligned} (n-1)i_1 &= ni_1 - i_1 = ni_1 + i_2 - (i_1 + i_2) \\ &\leq nr_1 + r_2 - (r_1 + r_2) \\ &= (n-1)r_1. \end{aligned}$$

Therefore $i_1 \leq r_1$; by (12) $r_1 - i_1 \leq i_2 - r_2$, and $i_1 \geq i_2$ because f is biased, so we obtain

$$r_2 \leq i_2 \leq i_1 \leq r_1,$$

and it follows that $|f|$ is also biased. Since $|fb| = |f|b$, (ii) shows that $r'_1 < r'_2$ (and $nr'_1 + r'_2 = nr_1 + r_2$). Now by the contrapositive of the preceding argument we obtain $r'_1 + r'_2 > i_1 + i_2 = \deg f$, hence $\deg f < \deg fb$. ■

COROLLARY 9.6. *If $0 \neq f \in U$ and $b \in B \setminus C$, then each of the following statements implies the next:*

- (a) f is biased,
- (b) $\deg f < \deg f b \tau$, where $\tau = (x_2, x_1)$,
- (c) $\deg f \leq \deg f b \tau$,
- (d) $f b \tau$ is biased.

Proof. (a) $\Rightarrow$ (b) follows by Lemma 9.5(iii), since τ does not affect the degree. (b) $\Rightarrow$ (c) is clear; (c) $\Rightarrow$ (d): if $f b \tau$ is not biased, then $f b$ is biased, so by Lemma 9.5, $\deg f b \tau = \deg f b < \deg f b b^{-1} = \deg f$, which contradicts (c). ■

With the help of this lemma and its corollary we obtain precise conditions for stabilizers in $\text{Aut } k\langle x_1, x_2 \rangle$ to be small:

THEOREM 9.7. *If V is a non-zero finite-dimensional subspace of $U = k\langle x_1, x_2 \rangle / k[c]$, then the stabilizer of V under the action of $\text{Aut } k\langle x_1, x_2 \rangle$ lies in a conjugate of A or B .*

Proof. Put $G = \text{stab } V$, $n = \max\{\deg f \mid f \in V\}$. Given $q \in G$, let us write

$$q = b_0 \tau b_1 \tau \dots \tau b_m, \quad b_i \in B, b_i \notin C \text{ for } 0 < i < m. \quad (13)$$

We claim that $m \leq 2n$. Pick any non-zero $f \in V$ and for $-1 \leq i \leq m$ set $d_i = \deg f(b_0 \tau)(b_1 \tau) \dots (b_i \tau)$; thus $d_{-1} = \deg f$ and for $0 \leq i \leq m$,

$$d_i = \deg f(b_0 \tau) \dots (b_{i-1} \tau) b_i,$$

since τ does not change the degree. Consider the greatest integer i such that

$$d_{-1} \geq d_0 > d_1 > \dots > d_{i-1}.$$

Since $n \geq \deg f = d_{-1}$ and $d_{i-1} \geq 1$, we have a strictly descending chain, hence $i \leq n$. If $i \geq m$, then $n \geq m$ and so $m \leq 2n$ as claimed. We may therefore assume that $i < m$. Now either $i = 0$ and $d_{-1} < d_0$ or $i > 0$ and $d_{i-1} \leq d_i$, by the definition of i . In both cases $b_i \notin C$ and $d_{i-1} \leq d_i$, so by Cor. 9.6, $f b_0 \tau \dots b_i \tau$ is biased, and by repeated application of Cor. 9.6 we see that

$$1 \leq d_i < d_{i+1} < \dots < d_{m-1} \leq d_m \quad (\leq n),$$

because $f q \in V$. It follows that $m - i \leq n$, and so $m \leq n + i \leq 2n$.

Thus we see that every $q \in G$ can be expressed as in (13), with $m \leq 2n$. Hence the width of G is at most $4n + 2$. By Th. 8.7, it follows that G lies in a conjugate of A or B . ■

The following corollary is an immediate consequence of the theorem, although much weaker.

COROLLARY 9.8. *If V is a finite-dimensional subspace of $R = k\langle x_1, x_2 \rangle$ not contained in $k[c]$, then $\text{stab } V$ under the action of $\text{Aut } R$ lies in a conjugate of A or B . ■*

Exercises 6.9

1. Let X be a set, $x_0 \in X$ and $X' = X \setminus x_0$. Show that every automorphism of $k\langle X \rangle$ fixing (or even just stabilizing) $k\langle X' \rangle$ maps x_0 to $\lambda x_0 + f$, where $\lambda \in k^\times$, $f \in k\langle X' \rangle$.
2. Show that $\text{SAut } k\langle x_1, x_2 \rangle \cong SA_{\mathbb{C}}^* SB$, where $SA = A \cap \text{SAut}$, etc.
3. (Lane [76]) Prove an analogue of Cor. 9.8 for $k[x_1, x_2]$ (this is the case $c = 0$).
4. (P. Russell [80]) Show that for any finite set X , any epic k -algebra endomorphism of $k[X]$ is an automorphism (cf. Th. 9.3).
- 5*. (Dicks and Lewin [82]) Show directly (i.e. without using Th. 9.3) that an endomorphism of $k\langle x_1, x_2 \rangle$ fixing $[x_1, x_2]$ is an epimorphism.
- 6*. (Schofield [85]) Show that when X is finite, any epic endomorphism of $k\langle X \rangle$ is an automorphism (note that Ex. 5 and 6 give another proof of Th. 9.3).
7. In $k\langle x_1, x_2 \rangle$ put $p = [c, x_1]x_2 - [c, x_2]x_1$. Show that the stabilizer of kp is the group of all linear automorphisms. Find the stabilizer of kq , where $q = [c, x_2]$.

6.10 Invariants in free algebras

Let R be a k -algebra and G a group of k -algebra automorphisms of R . An element r of R is said to be an *invariant* of G if $r^g = r$ for all $g \in G$; if $r^g = \lambda_g r$ ($\lambda_g \in k$) for all $g \in G$, r is called a *relative invariant*. Thus r is a relative invariant precisely when G stabilizes kr .

The set of all invariants of G is a k -subalgebra of R , denoted by R^G and called the *fixed ring*, or *algebra of invariants* of G . The set of relative invariants is not an algebra, but we shall sometimes refer to the subalgebra generated by this set as the algebra of relative invariants. It is conjectured that if R is a free algebra and G a group of k -algebra automorphisms of R , then both these algebras are free. We shall deal with some special cases of these conjectures here. To begin with we shall show that the algebra of invariants is a 2-fir; this holds under rather wider conditions.

PROPOSITION 10.1. *Let R be a filtered ring satisfying the 2-term weak algorithm and let G be a set of endomorphisms of R . Then the fixed ring R^G is a strong E_2 -ring, and hence a 2-fir.*

Proof. Let $a, b \in R^G$ be such that a, b are right R -dependent. By Prop. 2.8.1, there exist $n \geq 1$ and a unique sequence $x_0, x_1, \dots, x_n \in R$ such that

$$a = x_0 p(x_1, \dots, x_n), \quad b = x_0 p(x_1, \dots, x_{n-1}), \quad (1)$$

where the p 's are the continuant polynomials defined in 2.7. For any $\alpha \in G$ we have

$$a = a^\alpha = x_0^\alpha p(x_1^\alpha, \dots, x_n^\alpha), \quad b = b^\alpha = x_0^\alpha p(x_1^\alpha, \dots, x_{n-1}^\alpha);$$

by uniqueness, $x_0 = x_0^\alpha, \dots, x_n = x_n^\alpha$, hence $x_i \in R^G$ ($i = 0, 1, \dots, n$). Moreover, by 2.7 there exists $U \in \mathbb{E}_2(R^G)$ such that $(a, b)U = (x_0, 0)$ hence R^G is a strong $\mathbb{E}_2$ -ring, and by Cor. 1.3.2, a 2-fir. ■

In particular, this result shows that for any free algebra $R = k\langle X \rangle$ and any group G of automorphisms of R , the fixed algebra R^G is a strong $\mathbb{E}_2$ -ring. One would like to assert that R^G satisfies the weak algorithm, but to establish this claim we shall need to assume that the automorphisms respect the grading.

PROPOSITION 10.2. *Let $R = \bigoplus R_n$ be a graded ring with weak algorithm and G a group of homogeneous automorphisms of R , i.e. $R_n G = R_n$ for all n . Then the fixed ring $R^G = \bigoplus R_n^G$ is a graded ring with weak algorithm.*

Proof. It is clear that $R^G = \bigoplus R_n^G$ is a graded ring. Consider a well-arranged homogeneous relation

$$a_1 b_1 + \dots + a_n b_n = 0$$

in R^G , where $v(a_1) \leq \dots \leq v(a_n)$, $v(a_i) + v(b_i) = d$. By the weak algorithm in R there exists i , $1 \leq i \leq n$, such that a_i is right linearly dependent on $a_1, \dots, a_{i-1}$:

$$a_i = a_1 c_1 + \dots + a_{i-1} c_{i-1}, \quad (c_j \in R).$$

If we choose the least such i , then $a_1, \dots, a_{i-1}$ are right R -independent. Let $g \in G$, then $a_i^g = a_j$, hence

$$a_1(c_1 - c_1^g) + \dots + a_{i-1}(c_{i-1} - c_{i-1}^g) = 0.$$

By R -independence we find that $c_j^g = c_j$, so $c_j \in R^G$ ($j = 1, \dots, i-1$), and this establishes the weak algorithm for R^G . ■

In particular, we may take $R = k\langle X \rangle$, where X is graded in any way, to obtain a result of Kharchenko and Lane:

THEOREM 10.3. *Let $R = k\langle X \rangle$ be a free algebra and G a group of k -automorphisms of R . If G is homogeneous with respect to the grading on R induced by some function $d: X \rightarrow \mathbb{N}_+$, then the algebra of invariants of G is free, on a set which is homogeneous with respect to d . ■*

As a special case we have the standard grading: $d(x) = 1$ for all $x \in X$. By way of example, let us calculate the algebra of relative invariants of the group of all linear automorphisms of $R = k\langle X \rangle$, where $X = \{x_1, \dots, x_d\}$ and $\text{char } k = 0$. We have $R = \bigoplus R_n$ and $G = \text{GL}_d(k)$, viewed as the group of all linear automorphisms of R . Denote by R' the algebra of relative invariants. For each $n \geq 1$, Sym_n and G both act on R_n , where for $\sigma \in \text{Sym}_n$, $y_1, \dots, y_n \in X$ we define

$$(y_1 \dots y_n)\sigma = y_{1\sigma} \dots y_{n\sigma}$$

and extend this action by linearity; we shall also say that Sym_n acts by *place-permutations* on R_n . It is easy to see that the actions of Sym_n and G commute, so R_n is a $G \times \text{Sym}_n$ -module, while R' is a Sym_n -module. Now consider the standard polynomial of degree d (cf. e.g. Cohn [77], p. 452):

$$\delta = S(x_1, \dots, x_d) = \sum_{\sigma \in \text{Sym}_d} \text{sgn } \sigma \cdot x_{1\sigma} \dots x_{d\sigma} \in R_d \quad (2)$$

For each $g \in G$, $\delta^g = \det g \cdot \delta$, so δ is a relative invariant of G . Therefore R' contains the k -algebra spanned by

$$\bigcup_n \{ \delta^n \sigma \mid \sigma \in \text{Sym}_{nd} \}. \quad (3)$$

By the representation theory of the symmetric group (cf. e.g. James and Kerber [81], Ch. 3) this is all of R' , at least when $k = \mathbb{Q}$. To find a k -basis, we restrict the σ 's in (3) to range over the standard Young tableaux with d rows, all of depth n . By Nagata's formula the number of these tableaux is

$$\frac{(nd)!}{\prod_{i=1}^d \prod_{j=1}^n (i+j-1)}.$$

In fact, R' is a free algebra, by Th. 10.3, as the fixed ring of $\text{SL}_d(k)$. To obtain a free generating set we restrict σ in (3) further by deleting the standard Young tableaux constructed by juxtaposing Young tableaux with fewer columns, e.g. where σ fixes jd , for some $j < n$. For example, when $d = 2$, then $\delta = c = [x_1, x_2]$ and the free generating set has the Hilbert series

$$\sum_{n \geq 0} \frac{1}{n+1} \binom{2n}{n} t^{2n+2} = \frac{1}{2} (1 - \sqrt{1-4t^2}) = t^2 + t^4 + 2t^6 + \dots$$

So there is one generator in degree 2, viz. c , one of degree 4, $x_1cx_2 - x_2cx_1$, etc. (for $d > 2$ the series is probably not algebraic).

The results obtained so far do not indicate when R^G is finitely generated; for finite groups of linear automorphisms we have a precise result:

THEOREM 10.4. *Let G be a finite group of linear automorphisms of a free algebra of finite rank, $R = k\langle x_1, \dots, x_d \rangle$. Then R^G is finitely generated as a k -algebra if and only if G is scalar. In this case G is cyclic, generated by an automorphism*

$$(\zeta x_1, \zeta x_2, \dots, \zeta x_d), \quad (4)$$

for some primitive m th root of 1, ζ , and R^G is generated by the d^m monomials of length m .

Proof. Assume that G is scalar; it is isomorphic to a subgroup of $k^\times$, hence cyclic, of order m , say, with generator given by (4). It is clear that R^G is freely generated by the d^m monomials of length m .

Conversely, assume that R^G is finitely generated and let $g \in G$; we have to show that G is scalar. For $d = 1$ there is nothing to prove, so we may take $d \geq 2$. We first make a reduction to the case where k is algebraically closed. Let $\bar{k}$ be an algebraic closure of k and put $\bar{R} = R \otimes_k \bar{k}$, then $G \subseteq \text{GL}_d(\bar{k})$; since $\bar{k}$ has a k -basis, it follows that $\bar{R}^G = R^G \otimes_k \bar{k}$ and this is finitely generated as $\bar{k}$ -algebra. We may therefore take k to be algebraically closed.

On conjugating G by a suitable $U \in \text{GL}_d(k)$ we may take g to be in Jordan normal form, say

$$(\lambda_1 x_1 + \varepsilon_1 x_2, \lambda_2 x_2 + \varepsilon_2 x_3, \dots, \lambda_{d-1} x_{d-1} + \varepsilon_{d-1} x_d, \lambda_d x_d), \quad (5)$$

where $\lambda_i \in k^\times$ and for $1 \leq i \leq d-1$, either $\varepsilon_i = 0$ or $\varepsilon_i = 1$ and $\lambda_i = \lambda_{i+1}$. Let $R = \bigcup R_n$ be filtered by degree, so that $R^G = \bigcup R_n^G$ (because G acts linearly). Since R^G is finitely generated, it must be generated by R_N^G for some N , and on increasing N if necessary, we may assume that $|G|$ divides N .

Let A be the set of all monomials $\neq 1$ which occur in the support of some element of R_N^G . Then R_N^G lies in the k -space spanned by $A \cup \{1\}$, hence R^G lies in the k -algebra generated by A . Let $\delta = S(x_1, \dots, x_d)$ again be the standard polynomial as in (2) and take $\sigma \in \text{Sym}_d$ to act by place-permutations. Then $\delta^N \sigma$ is fixed by G , because for any $h \in \text{GL}_d(k)$,

$$(\delta^N \sigma)h = (\det h)^N \delta^N \sigma.$$

When $h \in G$, $(\det h)^N = 1$ (by the choice of N), therefore $\delta^N \sigma \in R^G$. This shows that every monomial of length Nd occurring in the support of $\delta^N \sigma$ is a product of elements of A , and these elements are of length at most N . Since

σ was arbitrary, it follows that every place-permutation of $(x_1 \dots x_d)^N$ is a product of elements of A (of length at most N), therefore every monomial of length N has a left factor in A . By considering x_1^N we see that

$$x_1^m \in A \quad \text{for some } m \geq 1. \quad (6)$$

Moreover,

$$\text{there is a monomial } w \text{ such that } wx_1, \dots, wx_d \text{ all lie in } A; \quad (7)$$

for if this were not so, then we could for every $r \geq 1$ construct a monomial of length r with no left factor in A , which is a contradiction.

By (6) there exists an element r of R^G whose support contains x_1^m , and we can take r to be homogeneous and the coefficient of x_1^m to be 1. Thus for some $\mu \in k$ we have

$$r \in x_1^m + \mu x_1^{m-1} x_2 + V,$$

where V is the k -space spanned by all monomials of length at most m , other than $x_1^m, x_1^{m-1} x_2$. Since g has the form (5), it follows that $Vg \subseteq V$, and so

$$\begin{aligned} r &= r^g \in (\lambda_1 x_1 + \varepsilon_1 x_2)^m + \mu (\lambda_1 x_1 + \varepsilon_1 x_2)^{m-1} (\lambda_2 x_2 + \varepsilon_2 x_3) + V, \\ &= \lambda_1^m x_1^m + \lambda_1^{m-1} \varepsilon_1 x_1^{m-1} x_2 + \mu \lambda_1^{m-1} \lambda_2 x_1^{m-1} x_2 + V, \end{aligned}$$

It follows that $\lambda_1^m = 1$, $\mu = \lambda_1^{m-1} \varepsilon_1 + \mu \lambda_1^{m-1} \lambda_2$, so $\lambda_1 \mu = \varepsilon_1 + \mu \lambda_2$, $\varepsilon_1 = \mu(\lambda_1 - \lambda_2)$. But we saw that $\varepsilon_1 = 0$ unless $\lambda_1 = \lambda_2$, so $\varepsilon_1 = 0$ in all cases. Since the order of the Jordan block was arbitrary, g must be diagonal, say $g = (\lambda_1 x_1, \dots, \lambda_d x_d)$.

If $r \in R^G$, then r is fixed by g . Consider the monomial w as in (7) and suppose that $wg = \lambda w$ ($\lambda \in k^\times$); then $(wx_i)g = wx_i$ because $wx_i \in A$. But $(wx_i)g = \lambda \lambda_i wx_i$, hence $\lambda_i = \lambda^{-1}$ for $i = 1, \dots, d$, and so g is a scalar. ■

We can also describe the fixed algebra R^G by its Hilbert series:

PROPOSITION 10.5. *Let k be a field of characteristic 0 and G a finite group of linear automorphisms of $R = k\langle x_1, \dots, x_d \rangle$. Then the Hilbert series of the fixed ring R^G is given by*

$$H(R^G : k) = |G|^{-1} \sum_{g \in G} (1 - \text{Tr}(g)t)^{-1},$$

where $\text{Tr} : \text{GL}_d(k) \rightarrow k$ is the trace map.

Proof. Let $R = \bigoplus R_n$ be graded by total degree in the x 's. Each R_n is a kG -module and the Hilbert series of R^G is

$$H(R^G : k) = \sum \dim_k R_n^G \cdot t^n.$$

Write $e = |G|^{-1} \sum g \in kG$; then e is the idempotent such that $R^G = Re$, therefore $\dim_k R_n^G = \text{Tr}(e: R_n \rightarrow R_n)$. Now $R_n = R_1^{\otimes n}$, where $\otimes n$ denotes the n th tensor power and g acts on R_n as $g^{\otimes n}$, while $\text{Tr}(g^{\otimes n}) = (\text{Tr } g)^n$. Therefore

$$\begin{aligned} \text{Tr}(e: R_n \rightarrow R_n) &= |G|^{-1} \sum_{g \in G} \text{Tr}(g: R_n \rightarrow R_n) \\ &= |G|^{-1} \sum_{g \in G} (\text{Tr } g)^n. \end{aligned}$$

Hence

$$\begin{aligned} H(R^G: k) &= \sum \dim_k R_n^G \cdot t^n \\ &= \sum |G|^{-1} \sum (\text{Tr } g)^n \cdot t^n \\ &= |G|^{-1} \sum (1 - \text{Tr}(g)t)^{-1}. \blacksquare \end{aligned}$$

COROLLARY 10.6. *If a subgroup G of Sym_d acts on $R = k\langle x_1, \dots, x_d \rangle$ by variable permutations, then*

$$H(R^G: k) = |G|^{-1} \sum_m \frac{\beta_m}{1 - mt},$$

where β_m is the number of elements of G fixing m of the x_i .

Proof. G acts by permuting the monomials of R , so the cardinal of the basis of R_n^G does not depend on the characteristic of k . In fact a basis of R_n^G is given by

$$\Sigma\{wg \mid g \in G\}, \text{ where } w \text{ runs over all monomials of length } n.$$

We may therefore take $\text{char } k$ to be 0. By Th. 10.5

$$H(R^G: k) = |G|^{-1} \sum (1 - \text{Tr}(g)t)^{-1}$$

and here $\text{Tr}(g)$ is the number of x 's fixed by g , so there are β_m summands $(1 - mt)^{-1}$, for $m = 0, 1, \dots$, hence the result. $\blacksquare$

We remark that by Th. 10.3 the fixed algebra is free, and a free generating set may be written down explicitly.

For free algebras of rank 2 there is an analogue for finite groups of automorphisms that are not necessarily linear. We begin by working out a special case. Consider again the group B of x_2 -based de Jonquières automorphisms of $R = k\langle x_1, x_2 \rangle$:

$$B = \{(\alpha x_1 + p, \beta x_2 + \lambda) \mid \alpha, \beta \in k^\times, \lambda \in k, p \in k[x_2]\}.$$

We shall compute the algebra R' of relative invariants of B . Put

$$B_0 = \{(x_1 + p, x_2) \mid p \in k[x_2]\};$$

we claim that the algebra of relative invariants of B_0 , R'' say, is $k\langle c, x_2 \rangle$, where again $c = [x_1, x_2]$.

It is clear that $R'' \cong k\langle c, x_2 \rangle$. Conversely, let $r \in R''$, $\deg r = d$, say. Choose p of degree $d+2$. The ideal of R generated by x_1 and p has as right v -independent basis

$$x_1, p, c, x_2c, x_2^2c, \dots, x_2^dc,$$

as is easily checked. Hence there is a unique expression

$$r = q + \sum x_2^i cr_i + x_1 r_{d+1} + p r_{d+2}, \quad r_j \in R, q \in k[x_2], \deg q < d+2.$$

By comparing degrees in x_2 we see that $r_{d+2} = 0$. Now take $\sigma = (x_1 + p, x_2)$, then for some $\lambda \in k^\times$, $r\sigma = \lambda r$; we have

$$\begin{aligned} r\sigma &= q + \sum x_2^i c(r_i\sigma) + (x_1 + p)(r_{d+1}\sigma), \\ \lambda r &= \lambda q + \sum x_2^i c(\lambda r_i) + x_1(\lambda r_{d+1}), \end{aligned}$$

and by uniqueness, $\lambda q = q$, $r_i\sigma = \lambda r_i$, $r_{d+1}\sigma = \lambda r_{d+1}$, $r_{d+1}\sigma = 0$, hence

$$r - q = \sum x_2^i cr_i,$$

where the r_i are in R'' but of lower degree than r ; so by induction $r \in k\langle c, x_2 \rangle$. Thus $R'' = k\langle c, x_2 \rangle$.

Now B acts on $R'' = k\langle c, x_2 \rangle$ as

$$\{(\alpha\beta c, \beta x_2 + \lambda) \mid \alpha, \beta \in k^\times, \lambda \in k\} = \{(\alpha c, \beta x_2 + \lambda) \mid \alpha, \beta \in k^\times, \lambda \in k\}.$$

For $i \geq 0$ put $c_i = (\text{ad } x_2)^i c$, where $(\text{ad } x)r = [x, r]$. We shall show that R' is generated by the c_i ($i \geq 0$) and if k is a finite field, $x_2^{|k|} - x_2$. Clearly all these elements lie in R' . Now consider any relative B -invariant r . Using the relation $x_2c_i = c_ix_2 + c_{i+1}$, we can easily verify that the c_i form a right R'' -basis for the ideal $R''cR''$, so there is a unique expression

$$r = q + \sum c_i r_i, \quad r_i \in R'', q \in k[x_2].$$

Given $b \in B$, we have $rb = \lambda r$ for some $\lambda \in k^\times$, hence

$$rb = qb + \sum (c_i \lambda_i)(r_i b) = \lambda q + \sum c_i (\lambda r_i).$$

By uniqueness, $qb = \lambda q$, $r_i b = \lambda_i^{-1} \lambda r_i$, so q, r_i are relative B -invariants. The r_i are of lower degree than r , so by induction they lie in the subalgebra generated by the c_i and $x_2^{|k|} - x_2$. Finally, $q(x_2 + \lambda) = q(x_2)$ for all $\lambda \in k$, so the set of zeros of q (in a splitting field of k) is closed under addition by elements of k . Hence either $q \in k$ or k is finite and q is a polynomial in $x_2^{|k|} - x_2$.

So R' is generated by $c_0, c_1, \dots$, and $x_2^{|k|} - x_2$ if k is finite. Suppose that k is infinite, so that $R' = k\langle c_0, c_1, \dots \rangle$; clearly the c_i form a free generating set. If k is finite, with q elements say, then

$$[x_2^q - x_2, c_i] = [x_2^q, c_i] - [x_2, c_i] = c_{i+q} - c_{i+1}.$$

Therefore R' is generated by $c_0, c_1, \dots, c_{q-1}, x_2^q - x_2$, and these elements are easily seen to form a free generating set.

We shall also need a lemma on the weak algorithm in R^G :

LEMMA 10.7. *Let R be a filtered ring satisfying the weak algorithm and let G be a group of automorphisms of R preserving the filtration. Then $K = R_0^G$ is a field, and if each R_n is finite-dimensional as a right K -space and G is locally finite, with no elements of order divisible by $\text{char } K$, then the filtered ring R^G again satisfies the weak algorithm.*

Proof. Consider a v -dependence relation in R^G ,

$$v(\sum a_i b_i) < \max\{v(a_i) + v(b_i)\},$$

where $v(a_1) \leq v(a_2) \leq \dots$, say. By the weak algorithm there is some j and $c_1, \dots, c_{j-1} \in R$ such that if $v(a_j) = n$, say, then

$$a_j - \sum a_i c_i \in R_{n-1}, \quad a_i c_i \in R_n \quad (i = 1, \dots, j-1).$$

For each finite subgroup H of G , R_n^H is a subspace of the finite-dimensional right K -space R_n , and $R_n^H \supseteq R_n^G$; we can choose H so as to minimize the right K -dimension of R_n^H , and then $R_n^H = R_n^G$. Further, we have $R_{n-1}^H = R_n^H \cap R_{n-1} = R_n^G \cap R_{n-1} = R_{n-1}^G$. We put

$$e = |H|^{-1} \sum_{h \in H} h \in KH;$$

then e is an idempotent such that right multiplication by e in R_n is a left R^H -linear map projecting R_n on $R_n^H = R_n^G$. Each $c_i e$ lies in $R_n e = R_n^H = R_n^G$ and since $a_i \in R_n^G$, we have

$$a_j - \sum a_i (c_i e) = (a_j - \sum a_i c_i) e \in R_{n-1} e = R_{n-1}^H = R_{n-1}^G,$$

while $a_i (c_i e) = (a_i c_i) e \in R_n e = R_n^H = R_n^G$ ($i = 1, \dots, j-1$). Thus a_j is right v -dependent on $a_1, \dots, a_{j-1}$ in R^G and the weak algorithm holds in R^G . ■

THEOREM 10.8. *Let G be a group of automorphisms of $k\langle x_1, x_2 \rangle$. If G is locally finite with no elements of order divisible by $\text{char } k$, then the algebra of invariants of G is free.*

Proof. Write $R = k\langle x_1, x_2 \rangle$ and let A, B be the subgroups of $\text{Aut } R$ defined as in 6.8. If G does not lie in a conjugate of A or B , then by Th. 9.7, $R^G \subseteq k[c]$, where $c = [x_1, x_2]$. Since G acts on $k[c]$, it is clear that R^G is free.

We may therefore assume that G lies in a conjugate of A or B , and by

passing to a conjugate, that G itself is contained in A or B . If $G \subseteq A$, then R^G is free by Lemma 10.7, since R satisfies the hypotheses of that lemma, with the usual filtration. Thus we may assume that $G \subseteq B$. We recall that we have a map $\deg: B \rightarrow \mathbb{N} \cup \{\infty\}$. If $\{\deg g \mid g \in G\}$ is unbounded, then the argument for the above worked example shows that $R^G \subseteq k\langle c, x_2 \rangle$ and G acts on $k\langle c, x_2 \rangle$. By Lemma 10.7, with the usual filtration on $k\langle c, x_2 \rangle$ we conclude that R^G is free. There remains the case where $\deg g$ ($g \in G$) is bounded, by d , say. If we take the $(d, 1)$ -filtration on R , we find that G preserves the filtration and $\dim_k R_n$ is finite, hence R^G is free, by Lemma 10.7. ■

Exercises 6.10

1. Find the invariants and relative invariants of $k[x]$.

2. Let R be a graded ring with weak algorithm and G a group of homogeneous automorphisms of R . Find all the relative invariants of R .

3. Let R be a free algebra of finite rank and G a group of homogeneous automorphisms of R . Show that R has a basis of elements $\sum \{mg \mid g \in G\}$, where m ranges over all monomials that (relative to the lexicographical ordering) are maximal in their orbits. Calling these monomials G -maximal, show that the set of all G -maximal monomials is a free monoid, and the indecomposable ones form a prefix code, which is a free generating set of R^G .

4*. (Bergman and Cohn [69]) Let $R = k\langle X \rangle$ and let G be a group acting on X with finite orbits and without fixed points. Show that R^G is a free k -algebra. If $|X| = d$, and G is precisely doubly transitive on X , so that $|G| = d(d-1)$ and each element of G (apart from 1) fixes exactly one point of X , show that the number of orbits of length n in X^* is $(d-1)^{n-2}$ ($n \geq 2$), while there is one orbit of length 1. Show also that if γ_n denotes the number of generators of length n , then $\gamma_n = (d-2)(\gamma_{n-1} + \gamma_{n-2})$.

5. (Dicks and Formanek [82]) Find $H(R^G : k)$ when $|G| = p = \text{char } k$.

6. (P. Webb) Find $H(R^G : k)$ when $|G| = p^n$, $p = \text{char } k$.

7°. Is $H(R^G : k)$ always rational, for $G \subseteq \text{GL}_d(k)$? (From the text and Ex. 6 this is so when $\text{char } k = 0$ or when $\text{char } k = p \neq 0$ and G is a p -group.)

8. (Almkvist, Dicks and Formanek [85]) Show that $H(R^G : k)$ is rational whenever G has a cyclic Sylow p -subgroup, where $p = \text{char } k$.

9. (Almkvist, Dicks and Formanek [85]) Let G be a compact subgroup of $\text{GL}_d(\mathbb{C})$. Show that if $R = \mathbb{C}\langle x_1, \dots, x_d \rangle$, then

$$H(R^G : \mathbb{C}) = \int_G (1 - t \cdot \text{Tr})^{-1} \quad \text{for } |t| < d^{-1},$$

where $\int_G$ is the normalized Haar measure.

10. (Lane) Let R be $k\langle x_1, x_2 \rangle$ and G the cyclic group generated by an automorphism of the form $(x_1 + p, x_2)$, where $p \in k[x_2]^*$. Show that for all $a \in R^G$, $pax_1 - x_1ap \in R^G$ and that R^G is the smallest k -algebra with this property containing $k\langle c, x_2 \rangle$. Find a free generating set for R^G .

11. (Lane) In Th. 10.8 show that the algebra of relative invariants is free.

12. If G is a finite group of automorphisms of $R = k\langle x_1, x_2 \rangle$ and $\text{char } k \nmid |G|$, then R^G is free. Use Prop. 8.9, Th. 9.3 and Th. 10.3 to give an alternative proof of this fact.

13. (Bergman) Let R be a ring, S a subring of R and X a subset of S such that the right ideal XR is R -projective, with dual R -basis $\{\alpha_x : XR \rightarrow R \mid x \in X\}$. Show that if there is an S -bimodule map $\pi : R \rightarrow S$ fixing S , then $\{\alpha_x \pi : XS \rightarrow S \mid x \in X\}$ is a dual S -basis for the right ideal XS . Deduce that if R is right (semi-)hereditary, then so is S .

14. (Bergman) Let R be a ring and G a group of automorphisms of R such that the orbit rG of any $r \in R$ is finite, of order invertible in R . Show that the averaging map $\pi : R \rightarrow R^G$ given by

$$r\pi = |rG|^{-1} \sum rg$$

is an R^G -bimodule map fixing R^G . Deduce that if R is right (semi-)hereditary, then so is R^G .

15. (Schofield) Let $R = k\langle x_1, x_2 \rangle$, where $\text{char } k = 0$. If $\sigma = (x_1 + p, x_2 + \alpha)$, where $p \in k[x_2]^*$, $\alpha \in k^*$, show that there exists $q \in k[x_2]$ such that $q(x) - q(x + \alpha) = p(x)$ and deduce that σ is conjugate to $(x_1, x_2 + \alpha)$.

16^o. Show that for any group of automorphisms of the free algebra $k\langle X \rangle$, the ring of invariants is free.

6.11 The Galois correspondence

Our aim in this section is to establish a Galois correspondence for free algebras. However, it will be instructive to prove the result in a more general setting, to bring out more clearly what special properties of $k\langle X \rangle$ are needed.

Throughout this section R will be an integral domain and $\text{End}(R)$ the ring of all endomorphisms of R , as abelian group, and these endomorphisms will be thought of as maps of R , written on the right, with composition $x.y$. For each $r \in R$ we identify r with the element of $\text{End}(R)$ given by right multiplication by r and thus view R as a subring of $\text{End}(R)$. Similarly we regard the opposite ring R^o as a subring of $\text{End}(R)$ by identifying the element r^o corresponding to $r \in R$ with the map consisting of left multiplication by r . Thus for any $r, s \in R$, we have

$$r.s = rs, \quad r^o.s^o = (sr)^o, \quad r.s^o = s^o.r, \quad rs^o = sr. \quad (1)$$

Moreover, $r = r^o$ if and only if r lies in the centre of R . The group $\text{Aut}(R)$ of

ring automorphisms is a subgroup of $U(\text{End}(R))$, and it will now be convenient to denote the effect of g on r by rg rather than r^g . We shall be concerned with elements of the form

$$f = \sum g_i \cdot a_i^0 \cdot b_i, \quad \text{where } g_i \in \text{Aut}(R), a_i, b_i \in R.$$

Such an element will be called an R -trace form; its effect on R is

$$rf = \sum a_i(rg_i)b_i.$$

We also note the rules

$$r \cdot g = g \cdot rg, \quad r^0 \cdot g = g \cdot (rg)^0 \quad (r \in R, g \in \text{Aut}(R)),$$

which follow from the formulae $(sr)g = (sg)(rg)$, $(rs)g = (rg)(sg)$.

An element g of $\text{Aut}(R)$ is said to be X -inner, after Kharchenko (= *XapyeHKo*) on a subring S of R , if there exist $a, b \in R^\times$ such that $a \cdot g \cdot b^0 - a^0 \cdot b$ vanishes on S , i.e.

$$b(sg)(ag) = asb \quad \text{for all } s \in S. \quad (2)$$

For $s = 1$ this yields $b \cdot ag = ab$, hence if T is an extension ring of R containing $c = a^{-1}b$, we have $sg = c^{-1}sc$ in T , so that g is actually induced by an inner automorphism of a certain extension ring; however, the definition does not presuppose the existence of T . If g is X -inner on all of R , we call it X -inner. From $b \cdot ag = ab$ we obtain $b^0 \cdot a^0 = (ag)^0 \cdot b^0$, and by (1)

$$(ag)^0 \cdot b \cdot b^0 = b^0 \cdot a^0 \cdot b = b^0 \cdot a \cdot g \cdot b^0 = b^0 \cdot g \cdot ag \cdot b^0.$$

Since R is a domain, we can cancel b^0 on the right, and replacing a by cg^{-1} , we obtain

$$b^0 \cdot g \cdot c = c^0 \cdot b, \quad \text{where } b, c \in R^\times.$$

This shows that the property of being X -inner is left-right symmetric. An automorphism which is not X -inner is called X -outer, and a subgroup G of $\text{Aut}(R)$ is called X -outer if the only X -inner element of G is 1.

We shall recall the results of Kharchenko's Galois theory, essentially following Montgomery and Passman. We remark that most of these results were actually proved for prime rings; the restriction to integral domains (which is all we shall need) simplifies many of the arguments.

We leave to the reader the verification that the X -inner automorphisms form a subgroup of $\text{Aut}(R)$. Our first object is an analogue of Dedekind's lemma on the independence of automorphisms; this is prefaced by two lemmas.

LEMMA 11.1. Let R be an integral domain and S a subring; further, let $\alpha = \sum g_i \cdot a_i^0 \cdot b_i$ be an R -trace form and β an element of $R \cdot \alpha \cdot S$, say $\beta =$

$\sum_{j=1}^m r_j \cdot \alpha \cdot s_j$ ($r_j \in R, s_j \in S$). Then we have

$$R\beta \subseteq (R\alpha)S,$$

and

$$\beta = \sum g_i \cdot a_i^0 \cdot c_i, \quad \text{where } c_i = \sum_{j=1}^m (r_j g_i) b_i s_j.$$

Proof. $R\beta \subseteq R(R \cdot \alpha \cdot S) \subseteq (R\alpha)S$; now the rest is straightforward. ■

LEMMA 11.2. Let R be an integral domain and S a subring such that $RaS \cap S \neq 0$ for all $a \in R^\times$. Given an R -trace form $\alpha = \sum g_i \cdot a_i^0 \cdot b_i$ in which all the b_i are non-zero, suppose that for any $c_i \in R$, if

$$\sum g_i \cdot a_i^0 \cdot c_i \in R \cdot \alpha \cdot S,$$

then either all the c_i vanish or none do. Then $g_j^{-1}g_i$ is X -inner on S , for all $i, j = 1, \dots, n$.

Proof. Since $Rb_1S \cap S \neq 0$, we may replace α by a suitable element of $R \cdot \alpha \cdot S$ so as to ensure that $b_1 \in S^\times$. Fix $s \in S^\times$; we have

$$\alpha \cdot sb_1 - (b_1s)g_1^{-1} \cdot \alpha = \sum g_i \cdot a_i^0 \cdot c_i,$$

where $c_i = b_1sb_1 - ((b_1s)g_1^{-1}g_i)b_i$, by Lemma 11.1. Then $c_1 = 0$, therefore $c_i = 0$ for all i , i.e. $b_1^0 \cdot b_1 - b_1^0 \cdot g_1^{-1}g_i \cdot b_i$ vanishes on S , so $g_1^{-1}g_i$ is X -inner on S , hence so is $g_j^{-1}g_i = (g_1^{-1}g_j)^{-1}g_1^{-1}g_i$. ■

We shall apply these results as follows.

PROPOSITION 11.3. Let R be an integral domain, G an X -outer subgroup of $\text{Aut}(R)$ and let $\alpha = \sum g_i \cdot a_i^0 \cdot b_i$ be an R -trace form, where $g_i \in G$. Then

- (i) $R \cdot \alpha \cdot R$ contains an element of the form $g \cdot (\sum a_i^0 \cdot c_i)$, where the c_i are not all 0 and the sum is over those i for which $g_i = g$,
- (ii) if $\alpha = 0$ and $g_1 \cdot a_1^0, \dots, g_n \cdot a_n^0$ is a minimal right R -dependent family, then all the g_i are equal and
- (iii) if the g_i are distinct and the a_i, b_i are non-zero, then $a \cdot \alpha \neq 0$ for all $a \in R^\times$.

Proof. We may assume that all the b_i are non-zero and $n \geq 1$. By Lemma 11.1 (with $S = R$), any $\beta \in R \cdot \alpha \cdot R$ is of the form

$$\beta = \sum g_i \cdot a_i^0 \cdot c_i$$

If we choose a β with the least positive number of non-zero c_i and apply Lemma 11.2, we find that β has the form required in (i), because G is X -outer.

If $\alpha = 0$, then $\beta = 0$ and (ii) follows by minimality. Finally, suppose that $g_1, \dots, g_n$ are distinct and the a_i are non-zero. By (ii) a minimal dependent subfamily of $g_1 \cdot a_1^0, \dots, g_n \cdot a_n^0$ must consist of a single element, but $g_i \cdot a_i^0 \cdot b_i \neq 0$ if $b_i \neq 0$, so any 1-element subfamily is right R -independent. It follows that the $g_1 \cdot a_1^0, \dots, g_n \cdot a_n^0$ are right R -independent, hence for any $a, b_1, \dots, b_n \in R^\times$,

$$a \cdot \alpha = \sum g_i \cdot a_i^0 \cdot (a g_i b_i) \neq 0. \blacksquare$$

PROPOSITION 11.4. *Let R be an integral domain, G a finite X -outer subgroup of $\text{Aut}(R)$ and $a \in R^\times$. Then $R^G a R$ and $R a R^G$ contain non-zero ideals of R which meet R^G non-trivially.*

Proof. Put $\alpha = \sum g$; by Prop. 11.3(i), $R \cdot (\alpha \cdot a) \cdot R$ contains an element of the form $\beta = g \cdot c$ for some $c \in R^\times$, and $R\alpha \subseteq R^G$, because $\alpha h = \alpha$ for all $h \in H$; therefore

$$Rc = (Rg)c = R\beta \subseteq (R\alpha)aR \subseteq R^G a R. \quad (3)$$

It follows that $I = RcR \subseteq R^G a R$ is a non-zero ideal of R in $R^G a R$. Now choose $0 \neq b \in \bigcap_{g \in G} Ig$, e.g. $b = (cg_1) \dots (cg_n)$, where $G = \{g_1, \dots, g_n\}$, then for all $g \in G$, $bg \in I \subseteq R^G a R$. By Prop. 11.3(iii), $b \cdot \alpha \neq 0$, hence

$$0 \neq (R)b \cdot \alpha = Rb \cdot \alpha \subseteq R^G \cap \sum Rbg \subseteq R^G \cap I.$$

Thus I meets R^G ; by symmetry the same result holds for $R a R^G$. ■

We shall want to know under what circumstances an X -inner map reduces to the identity. In fact, it is convenient to state the conditions more generally for inclusion maps.

LEMMA 11.5. *Let R be an integral domain, G a finite X -outer subgroup of $\text{Aut}(R)$, S an R^G -subring of R and $\sigma: S \rightarrow R$ an R^G -ring homomorphism. If there exist $a, b, c \in R^\times$ such that*

$$asb = b(\sigma)c \quad \text{for all } s \in S,$$

then σ is the inclusion map.

Proof. Given $r \in R$, we have $s = \sum rg \in R^G \subseteq S$, so we have $\overset{\sigma}{s} = s$ (because σ fixes R^G) and hence $asb = bsc$. Thus for all $r \in R$,

$$\sum a(rg)b = \sum b(rg)c,$$

i.e.

$$\sum (g \cdot a^0 \cdot b - g \cdot b^0 \cdot c) = 0.$$

By Prop. 11.3(ii), $g \cdot a^\circ$ and $g \cdot b^\circ$ are right R -dependent, for some $g \in G$, say

$$g \cdot a^\circ \cdot d = g \cdot b^\circ \cdot e, \quad \text{where } d, e \in R^\times.$$

Since g is a unit in $\text{End}(R)$, we have $a^\circ \cdot d = b^\circ \cdot e$, and so

$$\begin{aligned} b \cdot a^\circ \cdot d &= b \cdot b^\circ \cdot e \\ &= b^\circ \cdot be = b^\circ \cdot ad \\ &= b^\circ \cdot a \cdot d. \end{aligned}$$

Cancelling d , we obtain $b \cdot a^\circ = b^\circ \cdot a$. Hence, for all $s \in S$, we have $b(s\sigma)c = asb = bsa$. Putting $s = 1$ we find that $a = c$ and on cancelling a , $\sigma = 1$, so σ is induced by the inclusion map. ■

These results may be used to derive analogues of the fundamental theorem of Galois theory for fields; we shall limit ourselves here to what is actually needed later.

PROPOSITION 11.6. *Let R be an integral domain, G a finite X -outer subgroup of $\text{Aut}(R)$ and S an R^G -subring of R .*

(i) *If H is the subgroup of G fixing S elementwise, then*

$$R^G \subseteq S \subseteq R^H$$

and S contains a non-zero ideal of R^H .

(ii) *Any R^G -monomorphism $\sigma: S \rightarrow R$ can be extended to an element of G .*

Proof. Let $g_1, \dots, g_n$ be a left transversal of H in G , so that G is the disjoint union of the cosets $g_i H$. Put $\alpha = \sum g_i$, $\beta = \sum h$, then $\alpha \cdot \beta = \sum g$. For any $i \neq j$, $g_i^{-1} g_j \notin H$, so $g_i^{-1} g_j$ does not restrict to the inclusion map on S , by the choice of H , so by Lemma 11.5, $g_i^{-1} g_j$ is not X -inner on S . By Prop. 11.4, we have

$$RaS \cap S \supseteq RaR^G \cap R^G \neq 0 \quad \text{for all } a \in R^\times,$$

so we can apply Lemma 11.2. The (R, S) -bimodule $R \cdot \alpha \cdot S$ contains an element

$$\sum r_j \cdot \alpha \cdot s_j = g \cdot c, \quad \text{where } c \in R^\times, \quad g = g_i \text{ for some } i.$$

Since $RcS \cap S \neq 0$, we may assume that $c \in S^\times$, and on replacing α by $g^{-1}\alpha$, we may take $g = 1$. By Prop. 11.3(iii), $c \cdot \beta \neq 0$ and β centralizes R^H , so

$$\begin{aligned} 0 \neq c \cdot \beta &= \sum r_j \cdot \alpha \cdot s_j \cdot \beta \\ &= \sum r_j \cdot \alpha \beta \cdot s_j. \end{aligned} \tag{4}$$

Further, $c \cdot \beta = \beta \cdot c$, therefore

$$0 \neq R(\beta \cdot c) \subseteq (R(\alpha\beta))S \subseteq R^G S \subseteq S \subseteq R^H.$$

But β, c centralize $(R^H)^o$, hence $R(\beta \cdot c)$ is a non-zero left ideal of R^H contained in S . By symmetry S also contains a non-zero right ideal ϕR and so $S \supseteq R(\beta \cdot c)\phi R$, i.e. S contains a non-zero ideal of R^H and (i) is proved.

Let $\sigma: S \rightarrow R$ be an R^G -ring monomorphism. We may view $\text{Hom}_Z(R, S)$ as a subset of $\text{End}(R)$, and (4) takes place in this subset because $R\alpha\beta \subseteq R^G \subseteq S$, so $\beta \in \text{Hom}_Z(R, S)$. Therefore (4) may be right multiplied by $s \cdot \sigma$ for any $s \in S$, and this yields

$$\begin{aligned} \beta \cdot c \cdot s \cdot \sigma &= c \cdot \beta \cdot s \cdot \sigma \\ &= \sum r_j \cdot \alpha\beta \cdot s_j \cdot s \cdot \sigma \\ &= \sum r_j \cdot \alpha\beta \cdot (s_j\sigma) \cdot (s\sigma), \end{aligned}$$

because $R(\alpha\beta) \subseteq R^G$ and σ is left R^G -linear. Since $\alpha\beta = \Sigma g$, we have

$$\beta \cdot c \cdot s \cdot \sigma = \Sigma g \cdot b_g \cdot s\sigma, \quad (5)$$

where $b_g = \Sigma(r_j g)(s_j\sigma)$ is independent of s . Since $\beta \neq 0$ and σ is injective, the expression (5) is non-zero, hence $b_{g_0} \neq 0$ for some $g_0 \in G$. Replacing s by sc , we see that

$$\begin{aligned} \Sigma g \cdot b_g \cdot (sc)\sigma &= \beta \cdot c \cdot sc \cdot \sigma \\ &= cs \cdot \beta \cdot c \cdot \sigma \\ &= cs \cdot \Sigma g \cdot b_g \\ &= \Sigma g \cdot (cs)g \cdot b_g. \end{aligned}$$

By Prop. 11.3(iii), $b_g(sc)\sigma = (cs)gb_g$ for all $g \in G$. Applying g_0^{-1} , we find

$$(b_{g_0}g_0^{-1})(s\sigma g_0^{-1})(c\sigma g_0^{-1}) = cs(b_{g_0}g_0^{-1}).$$

By Lemma 11.5, σg_0^{-1} is the inclusion map on S , and this means that σ extends to g_0 . ■

We can now describe Kharchenko's form of the Galois correspondence between groups and subrings.

THEOREM 11.7. *Let R be an integral domain and G a finite X -outer subgroup of $\text{Aut}(R)$. Then there is a natural bijection between the subgroups H of G and the R^G -subrings of R that are anti-ideals in R , given by*

$$H \leftrightarrow R^H. \quad (6)$$

Moreover, each R^G -automorphism of an R^G -subring of R extends to an element of G .

Proof. Let H be a subgroup of G . Given $r \in R$, if $ar \in R^H$ for some non-zero $a \in R^H$, then for all $h \in H$, $ar = (ar)h = a(rh)$, so $rh = r$ and $r \in R^H$. This shows R^H to be an anti-ideal in R (in fact, it satisfies a stronger, one-sided condition). If H, K are subgroups of G and $R^H = R^K$, then each $h \in H$ fixes R^K , so by Prop. 11.6(ii) (with $S = R$, $G = K$) h extends to an element of K , i.e. $h \in K$, and thus $H \subseteq K$; by symmetry, $H = K$, so the correspondence (6) is injective.

Now consider any R^G -subring S of R which is an anti-ideal in R . Let H be the subgroup of G fixing S elementwise, so that $S \subseteq R^H$; we shall prove that equality holds. By Prop. 11.6(i), S contains a non-zero ideal I of R^H . Given any $r \in R^H$ and $0 \neq a \in I$, we have $ra, ar \in I \subseteq S$, hence $r \in S$ and so $R^H = S$ as claimed. This establishes the bijection and now the rest follows by Prop. 11.6(ii). ■

To apply this result we need to determine the X -outer automorphisms and anti-ideals in 2-firs.

LEMMA 11.8. *If R is a 2-fir all of whose invariant elements are central, then $\text{Aut}(R)$ is X -outer.*

Proof. Let $g \in \text{Aut}(R)$ be X -inner, say

$$a^\circ \cdot b = a \cdot g \cdot b^\circ = g \cdot ag \cdot b^\circ. \quad (7)$$

Applying this relation to 1, we get $0 \neq ab = b(ag) \in aR \cap bR$, hence

$$aR + bR = dR, \quad a = da_1, \quad b = db_1.$$

Now (7) becomes

$$a_1^\circ \cdot d^\circ \cdot d \cdot b_1 = g \cdot (da_1)g \cdot b_1^\circ \cdot d^\circ, \quad (8)$$

and the result of applying this to 1 is

$$da_1db_1 = db_1[(da_1)g]. \quad (9)$$

Cancelling d° on the right in (8) and multiplying by a_1 on the left, we obtain

$$a_1^\circ \cdot a_1db_1 = a_1 \cdot g \cdot (da_1)g \cdot b_1^\circ;$$

by (9) this becomes $a_1^\circ \cdot b_1 \cdot (da_1)g = a_1 \cdot g \cdot b_1^\circ \cdot (da_1)g$, hence

$$a_1^\circ \cdot b_1 = a_1 \cdot g \cdot b_1^\circ.$$

A comparison with (7) shows that we may take $d = 1$. We now have $aR + bR = R$ and hence

$$Rb = aRb + bRb = b(Rg)(ag) + bRb = bR(ag) + bRb = b(R(ag) + Rb) = bR,$$

because $1 \in R(ag) + Rb$. Thus b is invariant, and so by hypothesis, $b^\circ = b$; by symmetry, $a^\circ = a$. Cancelling b from (7) we have $a = a^\circ = a.g$, hence $g = 1$. ■

We also recall from 6.6 that if S is a 2-fir, division closed in R (i.e. $U(R) \cap S = U(S)$), then S is an anti-ideal (Lemma 6.11). As a consequence we obtain

THEOREM 11.9. *Let R be a filtered ring with 2-term weak algorithm, whose invariant elements are central. If G is a finite group of R_0 -automorphisms of R , then there is a natural bijection between subgroups of G and R^G -subrings that are 2-firs.*

Proof. By Lemma 11.8, G is X -outer, the above remark shows that each R_0 -subring of R which is a 2-fir is an anti-ideal and by Prop. 10.1 each fixed ring is a 2-fir. ■

For free algebras the Galois correspondence takes the following form:

THEOREM 11.10. *Let R be a free k -algebra and G a finite group of k -algebra automorphisms of R which are homogeneous with respect to some grading of R for which the weak algorithm is satisfied (e.g. linear automorphisms). Then there exists a natural bijection between the subgroups H of G and the free subalgebras of R containing R^H .*

This is an immediate consequence of Th. 11.9 and Th. 10.3. ■

Exercises 6.11

1. Prove the results of this section for prime rings.
2. Show that for an integral domain R the X -inner automorphisms form a normal subgroup of $\text{Aut}(R)$.

Notes and comments on Chapter 6

Bergman's centralizer theorem (Th. 7.7, cf. Bergman [67, 69]) was one of the main results of his thesis. It had been conjectured by the author in the early 1960's (cf. Cohn

[62] for special cases of Th. 7.1 and Cohn [63] for the fact that the centralizer is commutative). The proof of the centralizer theorem depends on the results in 6.1 on the integral closure of commutative subrings. These results are proved by Bergman [67] for Ore subrings; this form (also given in the first edition) is outlined in the exercises.

The notions of bounded and invariant element are treated for principal ideal domains by Jacobson [43]; our account in 6.1, 6.4 follows Bowtell and Cohn [71]; for Th. 4.13 see Carcanague [71]. Prop. 1.6 and Th. 1.8 are new in this edition; Th. 1.9 as well as 3.3 and 4, and the developments in 6.2, are taken from Bergman and Cohn [71], with simplified proofs. Ischebeck [71] has extended Th. 1.9 by showing that the centre of a directly indecomposable hereditary ring is a Krull ring. The conclusion of Th. 3.3 (every invariant element is a unit) has been established for (non-Ore) 2-firs with right ACC_2 by M. L. Roberts (unpublished), but there are examples showing it to be false for 2-firs with right ACC_1 (Bergman and Cohn [71]); see also Ex. 3. Th. 3.3 also ceases to hold when we require c to be merely right invariant (Ex. 6) or left invariant (Ex. 7). The main problem that remains is whether a non-Ore fir can have non-unit *right* invariant elements (Ex. 9⁰). Its solution could simplify the statement of Th. 5.7.

Codes, i.e. free subsets of free monoids (also called 'comma-free codes') form a flourishing subject in their own right, cf. Lothaire [83], Berstel, Perrin and Schützenberger [a]. Th. 5.1 occurs essentially in Levi [44], cf. also Dubreil-Jacotin [47]; Th. 5.2(a) was obtained by Schützenberger [59], for the rest of Th. 5.2 see Cohn [62]. The presentation of 6.5 and 6 follows Cohn [62], but Th. 5.5 and 6 have been added in this edition. The Kraft-McMillan inequality was shown by Kraft [49] to be necessary and sufficient for a prefix code to exist; McMillan [56] proved that the inequality holds for every code. The analogue for free algebras, Th. 5.6, is due to Kolotov [78]. Th. 5.7 is due to Bergman and Lewin [75]. Prop. 6.1 first appeared in Cohn [64]; Th. 6.4 (and Cor. 2.9.14, on which it is based) is taken from Knus [68], the proof given here is due to Dicks. The sufficiency of Th. 6.6 is in Moran [73], the necessity was proved by Dicks (unpublished). Th. 6.8 goes back to Lewin [68]; the present proof, with Prop. 6.7, is due to Berstel and Reutenauer [84]. The notion of anti-ideal was introduced by Kolotov [78], who also proved that free subalgebras of free algebras are anti-ideals (Th. 6.12), generalizing a result in the first edition, that the subalgebra generated by a non-trivial ideal is not free. The more general form in Lemma 6.11 is due to Bergman (unpublished). Prop. 7.3, appeared originally in the first edition (it was proved by the author in response to a question by G. Baumslag).

Sections 6.8–11 are new in this edition and follow closely W. Dicks' exposition in the Bedford study group 1982–84. The first result in this field was obtained by Jung [42], who showed that every automorphism of $\mathbb{C}[x_1, x_2]$ is tame; v. d. Kulk [53] extended the result to arbitrary ground fields. This is Th. 8.5; the original proof used methods from algebraic geometry, the proof given here is essentially that of Makar-Limanov [70], with simplifications by Dicks [83]. The de Jonquières automorphisms correspond to certain transformations of the plane, with multiple base point, which were first described by E. de Jonquières (1820–1901). Th. 8.6 is implicit in the work of v. d. Kulk and was first made explicit by Shafarevich [66], see also Nagata [72]; our proof follows Alperin [79]. Prop. 8.9 has frequently been observed, the present proof is due to Dicks.

The fact that all automorphisms of $k\langle x_1, x_2 \rangle$ are tame (Th. 9.3) was proved by Makar-Limanov [70] and independently by Czerniakiewicz [71]. The proof here

essentially follows that of Czerniakiewicz [71]; the present form of Lemma 9.2 is due to Dicks [82]. An up-to-date account of the Jacobian problem (in the commutative case) can be found in Bass, Connell and Wright [82]. The non-commutative form, Th. 9.4, was conjectured in Cohn [62'] and first proved by Dicks [82]. The results of 9.5–8 are implicit in Lane [76]; the presentation here is that of Dicks.

Prop. 10.1 is due to Bergman [71], Th. 10.3 was proved by Kharchenko [78], and independently by Lane [76]. Th. 10.4–6 are taken from Dicks and Formanek [82], but 10.4 was also found by Kharchenko [84], on whose proof the one in the text is based (with simplifications by Dicks). This type of argument is used in the theory of reductive algebraic groups, cf. Springer [77]. Th. 10.8 is due to Lane [76], the presentation is Dicks'. S. Donkin (unpublished) has shown that the results of the example after Th. 10.3 still hold for any infinite field. For a finite field k , $\text{GL}_d(k)$ is a finite group and there are many relative invariants, e.g. $\sum v^n$ may not lie in the subalgebra generated by δ and its transforms by place-permutations (take $k = \mathbb{F}_2$, $n = 3$, $d = 2$).

It may be of interest to compare the algebra of invariants of a finite group G of linear automorphisms of $k\langle x_1, \dots, x_d \rangle$ with that of $k[x_1, \dots, x_d]$. We list the results without proof and refer for details (and references) to Stanley [79]:

1. T. Molien showed in 1898 that in characteristic 0,

$$H(k[x_1, \dots, x_d]^G : k) = |G|^{-1} \sum (\det(I - gt))^{-1}.$$

(Note that $\det(I - gT) = 1 - t \cdot \text{Tr } g + \dots \pm t^n$.)

2. In 1916, E. Noether proved that $k[x_1, \dots, x_d]^G$ is finitely generated; this holds even if G is not linear (but still finite).

3. Coxeter, Shephard and Todd (1954), Chevalley (1955) and Serre (1958) showed that if $k[x_1, \dots, x_d]^G$ is a polynomial ring, then G can be generated by pseudo-reflexions, i.e. transformations g such that $g - I$ has rank 1. The converse holds if $\text{char } k$ does not divide the order of G , but not in general; see also Nakajima [83]. For example, in $G = \text{Sym}_d \subseteq \text{GL}_d(k)$ the transpositions are pseudo-reflexions, they generate G and $k[x_1, \dots, x_d]^G$ is a polynomial ring, in any characteristic (by the fundamental theorem for symmetric functions).

Galois theory underwent several generalizations (cf. Jacobson [64]). Kharchenko [77] presented a version for prime rings; we follow the trace-form presentation of this theory by Montgomery and Passman [84], adapted here to integral domains. Kharchenko's key result is Th. 11.7; the application to free algebras, Th. 11.10, appears in Kharchenko [78]. Lemma 11.8 was proved for free algebras by Kharchenko [78] and independently by Martindale and Montgomery [83]; the present form is taken from Bergman and Lewin [75].

7. Skew fields of fractions

This chapter studies ways of embedding rings in fields or, more generally, the homomorphisms of rings into fields. For a commutative ring such homomorphisms can be described completely in terms of prime ideals, and we shall see that the same type of description applies to quite general rings.

After some generalities on the rings of fractions obtained by inverting matrices (7.1) and on R -fields and their specializations (7.2), we introduce in 7.3 the notion of a matrix ideal. This corresponds to the concept of an ideal in a commutative ring, but has no direct interpretation. The analogue of a prime ideal, the prime matrix ideal defined in 7.4, has properties corresponding closely to those of prime ideals; what is more, the prime matrix ideals can be used to describe the homomorphisms of general rings into fields, just as prime ideals are used in the commutative case. This follows from Th. 4.8, which characterizes the prime matrix ideals as 'singular kernels', i.e. the sets of matrices which become singular under a homomorphism into some field.

This characterization is applied in 7.5 to derive criteria for a general ring to be embeddable in a field, or to have a universal field of fractions. These results are used to show that every Sylvester domain (in particular every semifir) has a universal field of fractions.

In the rest of this chapter these ideas are used to obtain localization theorems (7.10), a description of centralizers (7.8) and of the multiplicative group of the universal field of fractions of a semifir (7.9). This requires a comparison of the different representations of a given element of this field (7.6), and it involves a numerical invariant, the depth (7.7), which has no counterpart in the commutative case.

An alternative approach to the results of 7.2–4 is given in 7.11; this also leads to a localization theorem for n -firs (Th. 11.22).

7.1 The rational closure of a homomorphism

Let R be a ring; our basic problem will be to study the possible ways of embedding R in a field. Of course, there may be no such embedding, and it is more natural to treat the wider problem of finding homomorphisms of R into a field. Even this problem may have no solution, e.g. if $R = A_n$ is a full matrix

ring over a ring A , where $n > 1$, then any image of R is again an $n \times n$ matrix ring and so cannot be a subring of a field.

As a step towards the solution we may take a subset M of R and consider M -inverting homomorphisms. In the commutative case, once we have an $R^\times$ -inverting homomorphism to a non-zero ring, we have achieved the embedding in a field, but in general this need not be the case. Thus, for a general non-commutative ring the M -inverting homomorphisms are not very good approximations to homomorphisms into a field. We shall remedy this defect by inverting, instead of a set of elements, a set of square matrices. For a commutative ring this gives nothing new, since we can invert any square matrix A simply by adjoining an inverse of $\det A$. In the general case this is no longer possible, for even the Dieudonné determinant turns out, on closer examination, to be a *rational* function of the matrix entries.

Let R be a ring and Σ a set of matrices over R . A homomorphism $f: R \rightarrow S$ is said to be Σ -inverting if every matrix in Σ is mapped by f to an invertible matrix over S . Given any Σ -inverting homomorphism $f: R \rightarrow S$, we define the Σ -rational closure of R in S as the set $R_\Sigma(S)$ of all entries of inverses of matrices in Σ^f , the image of Σ under f ; the elements of $R_\Sigma(S)$ are also said to be Σ -rational over R . When Σ is the set of all matrices whose images under f have an inverse in S , we also write $R^f(S)$ instead of $R_\Sigma(S)$ and speak of the f -rational or simply the rational closure.

As we shall see, the f -rational closure of a ring R under a homomorphism f is always a subring containing $\text{im } f$. For general sets Σ the Σ -rational closure need not be a subring, as we know from the commutative case. If M is a multiplicative subset of a commutative ring R , then as we have seen in 0.8, the localization R_M is a ring. Let us call a set Σ of matrices *upper multiplicative*

if $1 \in \Sigma$, and whenever $A, B \in \Sigma$, then $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \in \Sigma$ for any matrix C of appropriate size; *lower* multiplicative sets are defined similarly (with C in the lower corner). If Σ is upper multiplicative and any matrix in Σ still lies in Σ after any permutation of rows and columns, then Σ is said to be *multiplicative*; clearly such a set is also lower multiplicative.

We first check that the set of all matrices inverted in a homomorphism is multiplicative.

PROPOSITION 1.1. *Given any homomorphism $f: R \rightarrow S$, let Σ be the set of all matrices over R whose image under f is invertible over S . Then Σ is multiplicative.*

Proof. Clearly 1^f is invertible, and if A, B are invertible matrices over S , then for any matrix C of suitable size,

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \text{ has the inverse } \begin{pmatrix} A^{-1} & -A^{-1}CB^{-1} \\ 0 & B^{-1} \end{pmatrix},$$

and invertibility is unaffected by permuting rows or columns. ■

The Σ -rational closure can be characterized in various ways. In stating the result we shall use the notation e_i for the column vector with 1 in the i th place and 0's elsewhere.

THEOREM 1.2. *Let R be a ring and Σ an upper multiplicative set of matrices over R . Given any Σ -inverting homomorphism $f: R \rightarrow S$, the Σ -rational closure $R_\Sigma(S)$ is a subring of S containing $\text{im} f$, and for any $x \in S$ the following conditions are equivalent:*

(a) $x \in R_\Sigma(S)$,

(b) x is a component of the solution u of a matrix equation

$$Au - e_j = 0, \quad \text{where } A \in \Sigma^f, \quad (1)$$

(c) x is a component of the solution of a matrix equation

$$Au + a = 0, \quad \text{where } A \in \Sigma^f, \quad (2)$$

and a is a column with entries in $\text{im} f$,

(d) $x = bA^{-1}c$, where $A \in \Sigma^f$, b is a row and c is a column with entries in $\text{im} f$.

Proof. We first prove the equivalence of the four conditions. (a) $\Rightarrow$ (b). By definition $R_\Sigma(S)$ consists of the entries of the inverses of matrices in Σ^f . If x occurs as (i, j) -entry of A^{-1} , then it is the i th component of the solution of (1), so (b) holds.

Next (c) is a special case of (b) and (c) $\Rightarrow$ (d) holds because when (2) holds, then $u_i = -e_i^T A^{-1}a$. To show (d) $\Rightarrow$ (a), let $x = bA^{-1}c$, then we have

$$\begin{pmatrix} 1 & b & 0 \\ 0 & A & c \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -bA^{-1} & bA^{-1}c \\ 0 & A^{-1} & -A^{-1}c \\ 0 & 0 & 1 \end{pmatrix},$$

where the matrix whose inverse is taken is again in Σ .

To prove that the Σ -rational closure $R_\Sigma(S)$ is a ring containing $\text{im} f$ we shall use property (c). Let $c \in \text{im} f$, then c satisfies the equation $1 \cdot u - c = 0$, which is of the form (2), hence $R_\Sigma(S) \supseteq \text{im} f$. Next if u_i is the i th component of the solution of (2) and v_j is the j th component of the solution of

$$Bv + b = 0,$$

then $u_i - v_j$ is the i th component of the solution of

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} w + \begin{pmatrix} a \\ b \end{pmatrix} = 0,$$

where C has for its j th column the i th column of A and the rest 0. Next $u_i v_j$ is the i th component of the solution of

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} w + \begin{pmatrix} 0 \\ b \end{pmatrix} = 0,$$

where C has as its j th column a and the rest 0. This shows that $R_\Sigma(S)$ is closed under subtraction and multiplication, and we have already seen that it contains 1, therefore it is a subring, as claimed. ■

Often it is convenient to use a different notation for the system (2), by taking as our basic matrix the augmented matrix (a, A) . Thus, omitting the reference to f , for simplicity, we shall write our system in the form

$$Au = 0, \quad A \in {}^m R^{m+1}. \quad (3)$$

The columns of A will be indicated by suffixes, thus

$$A = (A_0, A_1, \dots, A_m) = (A_0, A_*, A_\infty),$$

where $A_\infty = A_m$ is the last column and A_* is the $m \times (m-1)$ matrix $(A_1, \dots, A_{m-1})$. We shall call (3) an *admissible system* and A an *admissible matrix* of order m for the element p if (3) has a unique solution $u \in {}^{m+1}S$, normalized by the condition $u_0 = 1$, and $p = u_m$. Thus an $m \times (m+1)$ matrix over R is admissible if the image of the matrix formed by the last m columns is invertible over S . The last m columns, (A_*, A_∞) form the *denominator*, the first m columns, (A_0, A_*) form the *numerator* and A_* is the *core* of p in the representation (3) and we write $u = (1, u_*, p)^T$. Of course, these matrices depend not merely on p , but on the form of (3). We note the following analogue of Cramer's Rule:

PROPOSITION 1.3. *Let $f: R \rightarrow S$ be a homomorphism of rings. Given $p \in S$, if A is a matrix over $\text{im } f$ admissible for p , then*

$$(A_*, -A_0) = (A_*, A_\infty) \begin{pmatrix} I & u_* \\ 0 & p \end{pmatrix}; \quad (4)$$

thus over S , p is stably associated to the numerator of the system (3).

This is immediate since (A_*, A_∞) is invertible over S and the left-hand side of (4) is associated to the numerator. ■

For reference we note that if A, B are admissible matrices for p, q , respectively, then

$$\begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 \\ A_0 & 0 & A_\infty & A_* & A_\infty \end{pmatrix} \text{ and } \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 \\ 0 & 0 & A_0 & A_* & A_\infty \end{pmatrix} \quad (5)$$

are admissible matrices for $p-q$ and pq , respectively. As we see from (5), with the present conventions we need to take our set of matrices to be lower multiplicative.

We note that matrices over the Σ -rational closure can be obtained from solutions of matrix equations in precisely the same way:

PROPOSITION 1.4. *Let R be a ring, Σ a lower multiplicative set of matrices over R and $f: R \rightarrow S$ a Σ -inverting homomorphism. Then for any $m \times n$ matrix P over $R_\Sigma(S)$ there exists $r \geq 0$ and $A \in {}^{r+m}(\text{im } f)^{n+r+m}$, $u = (I, U, P)^T \in {}^{n+r+m}S^n$ such that*

$$Au = 0, \quad (A_*, A_\infty) \in \Sigma^f. \quad (6)$$

Proof. We have to show that every matrix P is determined by an equation (6). Suppose that P', P'' are determined by matrices A', A'' , respectively, then $P = P' + P''$ is determined by the system

$$\begin{pmatrix} A'_0 & A'_* & A'_\infty & 0 & 0 \\ A''_0 & 0 & -A''_\infty & A''_* & A''_\infty \end{pmatrix} \begin{pmatrix} I \\ U' \\ P' \\ U'' \\ P \end{pmatrix} = 0.$$

Therefore it is enough (by induction) to consider a matrix P with a single non-zero entry, say $P = \begin{pmatrix} p & 0 \\ 0 & 0 \end{pmatrix}$. If $Cu = 0$ is an admissible system for p , then we have

$$\begin{pmatrix} C_0 & 0 & C_* & C_\infty & 0 \\ 0 & 0 & 0 & 0 & I_{m-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & I_{n-1} \\ u_* & 0 \\ p & 0 \\ 0 & 0 \end{pmatrix} = 0. \quad \blacksquare$$

The admissible system (3) is more accurately described as a *left admissible* system; it follows by symmetry that the elements of $R_\Sigma(S)$ can equally well

be determined in terms of a *right* admissible system $vB = 0$, where B is a matrix of index -1 . The next result describes the relation between these two types.

PROPOSITION 1.5. *In the situation of Prop. 1.4, let P be an $m \times n$ matrix with a left admissible $(r+m) \times (n+r+m)$ matrix A . Then*

$$\begin{pmatrix} 0 & 0 & -I \\ A_0 & A_* & A_\infty \\ I & 0 & 0 \end{pmatrix}$$

is right admissible for P .

Proof. By hypothesis we have an equation

$$(A_0, A_*, A_\infty) \begin{pmatrix} I \\ U \\ P \end{pmatrix} = 0.$$

Hence we have

$$\begin{pmatrix} A_0 & A_* & A_\infty \\ I_n & 0 & 0 \end{pmatrix} \begin{pmatrix} I_n \\ U \\ P \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ I \end{pmatrix}.$$

The matrix on the left is square, and it has an inverse over S , so this means that

$$\begin{pmatrix} A_0 & A_* & A_\infty \\ I & 0 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} X & I \\ Y & U \\ Z & P \end{pmatrix}, \text{ say.}$$

Therefore

$$(I_m, Z, P) \begin{pmatrix} 0 & 0 & -I_m \\ A_0 & A_* & A_\infty \\ I & 0 & 0 \end{pmatrix} = 0,$$

and this is an equation of the required form. ■

Let R be a ring, Σ a multiplicative set of matrices over R and $f: R \rightarrow S$ a Σ -inverting homomorphism. Then for any $p \in R_\Sigma(S)$ we define the *left depth* $d_l(p)$ as the least n such that there is an $n \times n$ matrix A in Σ^f and a column $c \in {}^n(Rf)$ such that p occurs among the entries of the column $A^{-1}c$; the *right depth* $d_r(p)$ is defined similarly, using rows bA^{-1} . It is also possible to define

an upper depth $\bar{d}(p)$, using matrices A^{-1} and a lower depth $\underline{d}(p)$ using elements $bA^{-1}c$, but they will not be needed in what follows. It is easily seen that

$$\underline{d}(p) \leq d_l(p) \leq \bar{d}(p), \quad \underline{d}(p) \leq d_r(p) \leq \bar{d}(p), \quad (7)$$

and d_l, d_r cannot differ by more than 1 from each other and the other depths. This follows from Th. 1.2 and may be left to the reader to verify.

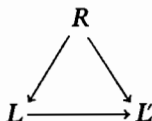
Exercises 7.1

1. Let R be a commutative ring and M a subset of R . Find conditions on M for the set $\{a^f(s^f)^{-1} \mid a \in R, s \in M\}$, under any M -inverting homomorphism f , to be a subring.
2. Show that if in a homomorphism $f: R \rightarrow S$, S is the rational closure of f , then f is epimorphic. Give an example to show that the converse is false.
3. Let K be any field. Show that if the transpose of every invertible matrix is invertible, then K is commutative. (Hint: Try a 2×2 matrix with $(1, 1)$ -entry 1.)
4. For any homomorphism $f: R \rightarrow S$ define the *division closure* of R in S as the least subring of S containing $\text{im } f$ and closed under forming inverses of elements, when they exist in S . Show that the division closure is contained in the rational closure and give examples to show that in general these two closures are distinct.
5. Prove the inequalities (7) for the depths of an element.
6. Let R be a ring, Σ a multiplicative set of matrices over R and $f: R \rightarrow S$ a Σ -inverting homomorphism. Show that for any $p, q \in R_\Sigma(S)$ and $a \in \text{im } f$, $a \neq 0$, $d(p-q) \leq d(p) + d(q)$, $d(pq) \leq d(p) + d(q)$, $d(a) = 1$, $d(p) - 1 \leq d(pa) \leq d(p)$, $d(p) \leq d(ap) \leq d(p) + 1$, where d is the left depth. What are the corresponding inequalities for the right depth?
7. Given a homomorphism $f: R \rightarrow S$, if every square matrix from R maps either to a left and right zerodivisor (or 0) or to an invertible matrix over S , show that $R^f(S)$ is such that every non-zero element is a left and right zerodivisor or invertible in S . If, further, S is an integral domain, deduce that the rational closure of R in S is a field. Show that under a homomorphism of R into a field, the rational closure is a subfield.
8. Given a homomorphism $f: R \rightarrow S$, where S is weakly finite, let Σ be the set of all matrices inverted under f . Show that for any square matrices A, B over R , if $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \in \Sigma$ for some C , then $A, B \in \Sigma$. Is weak finiteness necessary?
9. Let R and S be algebras over an infinite field. Given a homomorphism $f: R \rightarrow S$, show that for any finite set of elements in the rational closure there exist a matrix A and columns $c_1, \dots, c_r$ such that the given elements are the first components of the solutions of equations $Ax - c_i = 0$, $i = 1, \dots, r$.

10. Show that a matrix over $k\langle\langle X \rangle\rangle$ is invertible if and only if its image under the augmentation map is invertible over k . Deduce that the division closure of $k\langle X \rangle$ in $k\langle\langle X \rangle\rangle$ is equal to its rational closure.

7.2 The category of R -fields and specializations

Given a ring R , by an R -ring we understand a ring L with a homomorphism $R \rightarrow L$. For fixed R , the R -rings form a category in which the maps are the ring-homomorphisms $L \rightarrow L'$ such that the triangle shown is commutative.

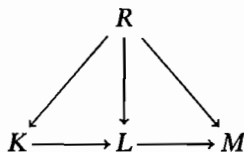


An R -ring which is a field is called an R -field. By an *epic* R -field we understand an R -field which is generated as field by the image of R . If K is an epic R -field for which the canonical map $R \rightarrow K$ is injective, K is called a *field of fractions* of R .

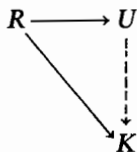
The only R -ring homomorphism possible between epic R -fields is an isomorphism. For any homomorphism between fields must be injective, and in this case the image will be a field containing the image of R , hence we have a surjection, and so an isomorphism. This shows the need to consider more general maps. Let us define a *local homomorphism* between R -fields K, L as an R -ring homomorphism $f: K_0 \rightarrow L$ from an R -subring K_0 of K to L such that any element of K_0 not in the kernel of f has an inverse in K_0 . The definition shows that K_0 is a local ring with maximal ideal $\ker f$, hence $K_0/\ker f$ is a field, isomorphic to a subfield of L , namely $\text{im } f$. The latter is a subfield of L containing the image of R in L , hence if L is an epic R -field, then $\text{im } f = L$. Thus, any local homomorphism to an epic R -field is surjective. Two local homomorphisms from an R -field K to another one, L , are considered equivalent if they agree on a subring K_0 of K and the common restriction to K_0 is again a local homomorphism. It is clear that this is, indeed, an equivalence relation; now a *specialization* between two R -fields K and L is defined as an equivalence class of local homomorphisms from K to L . $\spadesuit$

The R -fields and specializations form a category $\mathcal{F}_R$, say. Here it is only necessary to check that the composition of maps is defined. Given specializations $f: K \rightarrow L, g: L \rightarrow M$, let K_0, L_0 be the domains of f and g , respectively, and put $K_1 = \{x \in K_0 \mid xf \in L_0\}$, $f_1 = f|_{K_1}$. We assert that $f_1g: K_1 \rightarrow M$ is a local homomorphism and so defines a specialization. Let us denote the canonical mapping $R \rightarrow K$ by μ_K ; then we have $\mu_K f = \mu_L$, hence $R\mu_K \subseteq K_1$,

so that K_1 is an R -ring. Moreover, if $x \in K_1$ and $xf_1g \neq 0$, then $xf = xf_1 \neq 0$, so $x^{-1} \in K_0$ and $(x^{-1})f = (xf)^{-1} \in L_0$, hence $x^{-1} \in K_1$. This shows that f_1g defines in fact a specialization.



Let $\mathcal{E}_R$ be the full subcategory of $\mathcal{F}_R$ whose objects are the epic R -fields. An initial object in $\mathcal{E}_R$ is called a *universal R -field*. Explicitly a universal R -field is an epic R -field U such that for any epic R -field K there is a unique specialization $U \rightarrow K$. Clearly a universal R -field, if it exists at all, is unique up to isomorphism.



In general a ring R need not have a universal R -field, even when it has R -fields, e.g. a commutative ring R has a universal R -field if and only if its nil radical is prime, and we shall obtain an analogous condition for general rings in 7.5. Suppose that R has a universal R -field U ; then R has a field of fractions if and only if U is its field of fractions, as a glance at the above triangle shows. In that case we call U the *universal field of fractions* of R .

Let us illustrate these definitions by taking R to be a commutative ring. Then the epic R -fields correspond precisely to the prime ideals of R . Thus given any epic R -field K , the kernel of the canonical map $\mu_K: R \rightarrow K$ is a prime ideal and, conversely, if $\mathfrak{p}$ is a prime ideal of R , then the mapping $R \rightarrow \mathcal{F}(R/\mathfrak{p})$, where $\mathcal{F}(A)$ is the field of fractions of the domain A , gives us an epic R -field. The category $\mathcal{E}_R$ in this case is equivalent to the set of all prime ideals of R , with inclusion maps as morphisms. There is a universal R -field if and only if there is a least prime ideal, i.e. the nil radical is prime, and when this is 0 (i.e. when R is an integral domain), we have a universal field of fractions. A similar correspondence exists in the general case, and will be described in 7.5, below, once we have identified the objects to be used in place of prime ideals.

To find what we need, let us consider a general epic R -field K with homomorphism $\mu_K: R \rightarrow K$. Writing $\ker \mu_K = \mathfrak{p}$, we have a commutative diagram

$$\begin{array}{ccc}
 R & \longrightarrow & R_p \\
 \downarrow & \searrow \mu_K & \downarrow \\
 R/p & \longrightarrow & K
 \end{array}$$

When R is commutative, then so is K and then $\mathfrak{p}$ is a prime ideal of R which determines K up to isomorphism: we can either pass to $R/\mathfrak{p}$, an integral domain, and obtain K as its field of fractions, or we can localize at $\mathfrak{p}$ to obtain a local ring $R_{\mathfrak{p}}$, whose residue class field will be isomorphic to K . In the general case $\mathfrak{p}$ is no longer sufficient to determine K , since there are rings with several non-isomorphic fields of fractions (cf. Ex. 9).

Our aim will be to study the epic R -field K by means of the set of all matrices inverted over K ; before we can do so, we need the obvious but important remark that for any set Σ of matrices there always exists a *universal Σ -inverting homomorphism*: by this term we understand a homomorphism $\lambda: R \rightarrow R_{\Sigma}$ which is Σ -inverting and such that any Σ -inverting homomorphism f can be factored uniquely by λ , i.e. given $f: R \rightarrow S$ such that Σ^f consists of invertible matrices, there is a unique homomorphism $\bar{f}: R_{\Sigma} \rightarrow S$ such that the accompanying triangle commutes. The ring R_{Σ} is clearly determined up to isomorphism by these conditions; it is called the *universal Σ -inverting ring* or also a *universal localization* of R .

$$\begin{array}{ccc}
 R & \xrightarrow{\lambda} & R_{\Sigma} \\
 & \searrow f & \downarrow \bar{f} \\
 & & S
 \end{array}$$

Such a ring always exists (for any choice of R and Σ) and it may be constructed as follows. For each $m \times n$ matrix $A = (a_{ij})$ in Σ we take a set of mn symbols, arranged as an $n \times m$ matrix $A' = (a'_{ji})$ and take a ring presentation of R_{Σ} consisting of all the elements of R , as well as all the a'_{ji} as generators, and as defining relations take all the relations holding in R , together with the relations, in matrix form,

$$AA' = A'A = I \quad \text{for each } A \in \Sigma. \quad (1)$$

The mapping taking each element of R to the corresponding element of R_{Σ} is clearly a homomorphism $\lambda: R \rightarrow R_{\Sigma}$, which is Σ -inverting, by construction. If $f: R \rightarrow S$ is any Σ -inverting homomorphism, we define a homomorphism $\bar{f}: R_{\Sigma} \rightarrow S$ by putting $x\bar{f} = xf$ for all $x \in R$ and for any matrix $A \in \Sigma$ defining $\bar{f}$ on $(A\lambda)^{-1}$ by putting $(A\lambda)^{-1}\bar{f} = (Af)^{-1}$. This gives a well-defined homomorphism $\bar{f}$, because any relation in R_{Σ} is a consequence of the defining

relations in R and the relations (1), and all these relations also hold in S . Since we are mainly concerned with embeddings in fields, our sets Σ will usually consist of square matrices, but this is not essential in the general situation.

Of course the canonical homomorphism $\lambda: R \rightarrow R_\Sigma$ need not be injective and may, in fact, be zero; this is so if and only if $R_\Sigma = 0$, and we shall meet explicit criteria for this to happen later in this chapter. However, from the commutative triangle above we already see that if there is a Σ -inverting homomorphism f which is injective, then λ must be injective. We sum up these results in

THEOREM 2.1. *Let R be any ring and Σ any set of matrices over R . Then there is a universal localization R_Σ , unique up to isomorphism, with a universal Σ -inverting homomorphism*

$$\lambda: R \rightarrow R_\Sigma. \quad (2)$$

Moreover, λ is injective if and only if R can be embedded in a ring over which all the matrices of Σ have inverses. ■

Let us now consider, for an epic R -field K , in place of $\mathfrak{p} = \ker \mu_K$ the set $\mathcal{P}$ of all square matrices over R which become singular over K . This set $\mathcal{P}$ is called the *singular kernel* of μ_K and is written $\text{Ker } \mu_K$. There is no obvious construction for $R/\mathcal{P}$, 'the ring obtained by making the matrices in $\mathcal{P}$ singular', but we can form $R_\mathcal{P}$ as the universal Σ -inverting ring R_Σ , where Σ is the complement of $\mathcal{P}$ in the set $\mathfrak{M}(R)$ of all square matrices over R . By abuse of notation we sometimes write this as $R_\mathcal{P}$, as in the commutative case. Our next result describes any epic R -field in terms of its singular kernel, or rather, its complement in $\mathfrak{M}(R)$.

THEOREM 2.2. *Let R be any ring. Then*

- (i) *if Σ is a set of matrices such that the universal localization R_Σ is a local ring, then the residue-class field of R_Σ is an epic R -field and*
- (ii) *if K is an epic R -field and Σ the set of all matrices over R whose images in K are invertible, then Σ is multiplicative and R_Σ is a local ring with residue-class field isomorphic to K .*

Proof. Let Σ be a set of matrices such that R_Σ is a local ring, and denote its residue-class field by K . By composing the natural mappings we get a homomorphism $R \rightarrow R_\Sigma \rightarrow K$, and K is generated by the inverses of images of matrices in Σ , hence it is an epic R -field.

Conversely, let K be any epic R -field and Σ the set of all matrices over R whose images in K are invertible. Then Σ is multiplicative, by Prop. 1.1.

Further, by the definition of Σ and by Th. 2.1, we have an R -ring homomorphism $\alpha: R_\Sigma \rightarrow K$, and it will be enough to prove that any element of R_Σ not in $\ker \alpha$ is invertible.

Let $p \in R_\Sigma$ have the admissible system $Au = 0$, where $(A_*, A_\infty) \in \Sigma$, and hence, by Cramer's rule, p is stably associated to (A_0, A_*) . If $p^\alpha \neq 0$, then p^α is invertible, hence so is $(A_0, A_*)^\alpha$. Therefore $(A_0, A_*) \in \Sigma$ and applying Cramer's rule once more we find that p is invertible in R_Σ ; thus every element of R_Σ not in $\ker \alpha$ has an inverse. It follows that R_Σ is a local ring with maximal ideal $\ker \alpha$, and its residue-class field is therefore isomorphic to K , as claimed. ■

The sets Σ for which R_Σ is a local ring may be described as follows.

PROPOSITION 2.3. *Let R be a ring and Σ a multiplicative set of matrices such that $R_\Sigma \neq 0$ and for $A \in \Sigma$, if the $(1, 1)$ -minor of A is not in Σ , then $A - e_{11}$ is invertible over R_Σ . Then R_Σ is a local ring.*

Proof. Let u_1 be the first component of the solution of

$$Au - e_1 = 0, \quad (3)$$

and suppose that u_1 does not have a left inverse in R_Σ ; by Cramer's rule the numerator of u_1 in (3) cannot then be in Σ . This numerator is just the $(1, 1)$ -minor of A , hence $A - e_{11}$ is then invertible over R_Σ . Now consider the system

$$(A - e_{11})v - e_1 = 0. \quad (4)$$

This may be rewritten as

$$Av - e_1(1 + v_1) = 0. \quad (5)$$

By (3) and (5), $Av = e_1(1 + v_1) = Au(1 + v_1)$, so $v = u(1 + v_1)$. By equating first components we find $v_1 = u_1(1 + v_1)$, hence

$$(1 - u_1)(1 + v_1) = 1.$$

Thus for any $x \in R_\Sigma$, either x has a left inverse or $1 - x$ has a right inverse, hence R_Σ is a local ring, by Prop. 0.5.4. ■

In Th. 2.2 we saw that any R -field may be described entirely in terms of matrices over R ; we now show how to express specializations in terms of the sets of matrices inverted:

THEOREM 2.4. *Let R be any ring, K_1, K_2 any epic R -fields, Σ_i the set of all matrices over R inverted in K_i and R_i the universal localization R_{Σ_i} , with*

maximal ideal $\mathfrak{m}_i$ ($i = 1, 2$). Then the following are equivalent:

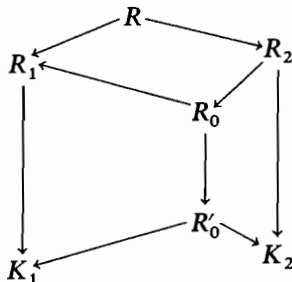
- (a) there is a specialization $\alpha: K_1 \rightarrow K_2$,
- (b) $\Sigma_1 \supseteq \Sigma_2$,
- (c) there is an R -ring homomorphism $R_2 \rightarrow R_1$.

If there is a specialization from K_1 to K_2 and one from K_2 to K_1 , then $K_1 \cong K_2$.

We note the reversal of direction in (c) compared with (a).

Proof. (a) $\Rightarrow$ (b). Let $\mu_i: R \rightarrow K_i$ be the canonical homomorphism. Take $A \in \Sigma_2$, then A^{μ_2} has an inverse which is the image of a matrix B over K_1 : $(A^{\mu_2})(B^{\mu_1}) = I$, hence $A^{\mu_1}B = I + C$, where $C^{\mu_2} = 0$. Hence $I + C$ has an inverse, and therefore so does A^{μ_1} , i.e. $A \in \Sigma_1$.

(b) $\Rightarrow$ (c) is clear, for when (b) holds, then $\lambda_1: R \rightarrow R_1$ is Σ_2 -inverting and so may be factored by λ_2 .



(c) $\Rightarrow$ (a). Let R_0 be the image of R_2 in R_1 . Then the natural homomorphism $R_1 \rightarrow K_1$ maps R_0 to $R'_0 = R_0/(R_0 \cap \mathfrak{m}_1)$. Now R_0 is a local ring (as homomorphic image of R_2) and $R_0 \cap \mathfrak{m}_1$ is a proper ideal, therefore the natural homomorphism $R_2 \rightarrow K_2$ can be taken via R'_0 , giving a homomorphism from a local subring (namely R'_0) of K_1 onto K_2 ; this is the required specialization. Now the last point follows using (b). ■

In Th. 2.4 suppose that K_2 is itself a universal localization; this means that $\mathfrak{m}_2 = 0$ and $R_2 = K_2$. Then the homomorphism $R_2 \rightarrow R_1$ is an isomorphism. Let us call a specialization of R -fields *proper* if it is not an isomorphism. Then we can express the result as

COROLLARY 2.5. *If an epic R -field K is itself a universal localization, then K cannot be obtained by a proper specialization from another R -field.* ■

In particular, this shows that if R has a universal R -field U , then U is the only epic R -field which can be a universal localization.

For example, the field of fractions of a commutative integral domain R is a universal R -field (even a field of fractions) and a universal localization of R . By contrast, if E and F are commutative fields, then $R = E \times F$ has two epic R -fields E and F , both universal localizations, but neither of them a universal R -field. The ring $S = E[x]/(x^2)$ has a unique epic S -field, viz. $S/(x)$, so this is a universal S -field, but not a universal localization of S (because S is itself a local ring). The ring $T = S \times F$ also has E and F as epic T -fields, but only one of them, viz. F , is a universal localization and neither is a universal T -field. Other examples are given in the exercises.

Exercises 7.2

1. Let R be a commutative ring and Σ a set of square matrices over R . Show that $R_\Sigma \cong R_S$, where S is the set consisting of all determinants of elements of Σ . Deduce that R_Σ is commutative.
2. Show that a commutative ring R has a universal R -field if and only if R has a unique least prime ideal (i.e. the nil radical of R is prime). What is the corresponding condition for left Noetherian rings?
3. Show that the $n \times n$ upper triangular matrix ring R over a field K has n epic R -fields isomorphic to K as fields, but distinct as R -fields. Deduce that for $n > 1$, R has no universal R -field. (Hint: Note that all nilpotent matrices map to 0 in any R -field.)
4. Show that if R is a left Bezout domain, then for any multiplicative set Σ of matrices, $R_\Sigma \cong R_S$ for a suitable subset S of R . (Hint: Use Ex. 0.8.13.)
5. Let R be a ring, Φ the set of all full matrices over R and suppose that there exists a Φ -inverting homomorphism into a non-zero weakly finite ring. Show that the universal localization R_Φ is a local ring.
6. Show that a homomorphism $R \rightarrow K$ of a ring into a field is an epimorphism (in the category of rings) if and only if K is an epic R -field.
7. For which values of n does the ring R of integers mod n have (i) a field of fractions, (ii) a universal localization which is a field and (iii) a universal R -field?
8. A set Σ of square matrices over a ring R is said to be *saturated* if any square matrix over R which becomes invertible over R_Σ lies in Σ . Show that any saturated set is multiplicative. Given any set Σ of square matrices over R , let Σ' be the set of all square matrices inverted in R_Σ . Show that Σ' is saturated and that $R_\Sigma \cong R_{\Sigma'}$ (Σ' is called the *saturation* of Σ).
- 9*. (Fisher [71]) Let k be a commutative field and $P = k[t]$ the polynomial ring with the endomorphism α_n induced by $t \mapsto t^n$. Show that if $n > 1$, the k -subalgebra of

$P[x; \alpha_n]$ generated by x and $y = xt$ is free on x and y . Hence obtain an embedding of the free algebra $R = k\langle x, y \rangle$ in a field, viz. $K(x; \alpha_n)$, where $K = k(t)$. Show that the fields so obtained for $n = 2, 3, \dots$ are 'minimal' objects in the category of fields of fractions of R and are non-isomorphic, as R -fields, for different n .

10*. (G. M. Bergman) Let R be a ring with two non-isomorphic fields of fractions K, L ; say L is not a specialization of K . Denote by $\bar{R} = R + tK[t]$ the subring of $K[t]$ consisting of polynomials with constant term in R . Show that $\bar{R}$ has a field of fractions, viz. $K(t)$, but no universal R -field.

11°. For which specializations of R -fields is the domain a valuation ring?

12°. Find conditions for the universal R -field to be a universal localization (cf. 7.6, below, for a special case).

7.3 Matrix ideals

We have seen in 7.2 that just as the epic R -fields for a commutative ring R can be described in terms of their kernels, so in general they are determined by their singular kernels. Now the kernels in the commutative case are well known as prime ideals, and it remains to give a description of singular kernels. This will be done in 7.4, where we shall describe them by a set of axioms reminiscent of the axioms for prime ideals. The resulting concept of a 'prime matrix ideal' is best studied as a special case of a matrix ideal; this is a set of matrices with properties similar to those of an ideal in a ring. In order to define it we need to describe the appropriate operations on matrices; instead of sum and product we have determinantal sum and diagonal sum. The latter we have already met; we recall that for any two matrices A, B the diagonal sum $A \oplus B$ is always defined, and if A, B are square of orders r, s , respectively, then $A \oplus B$ is square of order $r + s$. We shall use notations such as $A_1 \oplus A_2 \oplus \dots \oplus A_m$ for repeated diagonal sums; if in the last sum all the A_i are equal to A , we also write $\overset{m}{\oplus} A$.

We recall that a square matrix A is non-full if it can be written in the form PQ , where P is $n \times r$, Q is $r \times n$ and $r < n$. It follows that a non-full matrix maps to a singular matrix under any homomorphism into a field. If A is non-full, then its diagonal sum with any square matrix is again non-full. For if $A = PQ$ as above, then $A \oplus B = (P \oplus B)(Q \oplus I)$. On the other hand, if A is full, its diagonal sum with another full matrix need not be full. Thus a ring fails to have UGN precisely if the unit matrix I_n of some order is not full, even though 1 is full in any non-zero ring. We also note that if A is a matrix (full or not) such that $\overset{r}{\oplus} A$ is not full, for some r , then $\overset{r}{\oplus} A$ and hence A itself maps to a singular matrix under any homomorphism into a field.

The second operation to be described is only defined for certain pairs of matrices. Let A and B be two $n \times n$ matrices which differ at most in the first column, say $A = (A_1, A_2, \dots, A_n)$, $B = (B_1, A_2, \dots, A_n)$; then we define the *determinantal sum* of A and B with respect to the first column as the matrix $C = (A_1 + B_1, A_2, \dots, A_n)$. The determinantal sum with respect to another column (or row) is defined similarly, when it exists. We write

$$C = A \nabla B, \quad (1)$$

indicating, if necessary, in words the row or column whose elements are being added.

When R is commutative, so that determinants are defined, we have $\det C = \det A + \det B$, whenever A, B, C are related as in (1). On the other hand, over a field, even skew, if two of A, B, C are singular, then so is the third, as is easily seen. Further, any ring homomorphism preserves determinantal sums, therefore if two of A, B, C map to singular matrices under a homomorphism into a field, then so does the third.

Repeated determinantal sums need to be used with care, since the operation is not everywhere defined and *a fortiori* not associative. Thus to say that C is a determinantal sum of matrices $A_1, \dots, A_m$ means that we can replace a pair of neighbouring matrices among $A_1, \dots, A_m$ by their determinantal sum (with respect to some row or column) and repeat this process on another pair in the resulting set, and so on until we are only left with one matrix, namely C . We shall indicate this by writing

$$C = A_1 \nabla \dots \nabla A_m, \quad (2)$$

where it is understood that the sum has to be bracketed appropriately to be evaluated. We also note the identity

$$(A \nabla B) \oplus P = (A \oplus P) \nabla (B \oplus P), \quad (3)$$

which holds whenever $A \nabla B$ is defined.

Let R be a ring and $\mathfrak{M}(R)$ the set of all square matrices over R . We define a *matrix pre-ideal* in R as a subset $\mathcal{P}$ of $\mathfrak{M}(R)$ satisfying the following three conditions.

M.1. $\mathcal{P}$ includes all non-full matrices.

M.2. If $A, B \in \mathcal{P}$ and their determinantal sum C (with respect to some row or column) exists, then $C \in \mathcal{P}$.

M.3. If $A \in \mathcal{P}$, then $A \oplus B \in \mathcal{P}$ for all $B \in \mathfrak{M}(R)$.

If, further, we have

M.4. $A \oplus 1 \in \mathcal{P}$ implies $A \in \mathcal{P}$,

we call $\mathcal{P}$ a *matrix ideal*. A matrix pre-ideal is said to be *proper* if it does not contain the unit matrix of any size; clearly a matrix ideal is proper precisely when it does not contain the element 1.

We note the following simple consequences of the definitions. Let $\mathcal{P}$ be any matrix pre-ideal.

(a) A *hollow* matrix, i.e. $n \times n$ with an $r \times s$ block of zeros, where $r + s > n$, is not full, and hence is in $\mathcal{P}$; clearly this still holds if the zero block only arises after certain permutations of rows or columns. In particular, any square matrix with a zero row or column is hollow and so is in $\mathcal{P}$.

(b) In an equation $C = A \nabla B$, if B is non-full, then A lies in $\mathcal{P}$ if and only if C does.

If $A \in \mathcal{P}$, then $C \in \mathcal{P}$ by M.2, while if $C \in \mathcal{P}$, we can bring B over to the other side, changing the sign of the row or column to be added; clearly this leaves B non-full.

(c) Let $A \in \mathcal{P}$. Then the result of adding any right multiple of one column of A (or any left multiple of a row) to another again lies in $\mathcal{P}$.

Writing $A = (A_1, A_2, \dots, A_n)$, we have

$$(A_1 + A_2 c, A_2, \dots, A_n) = A \nabla (A_2, \dots, A_n) \begin{pmatrix} c \\ 0 \\ I_{n-1} \end{pmatrix};$$

on the right we have the determinantal sum (with respect to the first column) of A and a non-full matrix, hence the result is in $\mathcal{P}$.

(d) If $A \in \mathcal{P}$, then the result of interchanging any two columns (or rows) of A and changing the sign of one of them again lies in $\mathcal{P}$. In particular, if $A \oplus B \in \mathcal{P}$, then $B \oplus A \in \mathcal{P}$.

This follows in familiar fashion from (c). Writing only the two columns in question, we have, by repeated application of (c),

$$(A_1, A_2) \rightarrow (A_1 + A_2, A_2) \rightarrow (A_1 + A_2, -A_1) \rightarrow (A_2, -A_1).$$

(e) Let $A \in R_m$, $B \in R_n$, $C \in {}^n R^m$. Then

$$\begin{pmatrix} A & 0 \\ C & B \end{pmatrix} \in \mathcal{P} \Leftrightarrow \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in \mathcal{P}. \quad (4)$$

For given A, B, C let A_1, C_1 be the first columns of A, C , respectively, and write $A = (A_1, A')$, $C = (C_1, C')$. We have

$$\begin{pmatrix} A & 0 \\ C & B \end{pmatrix} = \begin{pmatrix} A_1 & A' & 0 \\ 0 & C' & B \end{pmatrix} \nabla \begin{pmatrix} 0 & A' & 0 \\ C_1 & C' & B \end{pmatrix}, \quad (5)$$

where the determinantal sum is with respect to the first column. By (a) the second matrix on the right is not full, hence by (b),

$$\begin{pmatrix} A & 0 \\ C & B \end{pmatrix} \in \mathcal{P} \Leftrightarrow \begin{pmatrix} A_1 & A' & 0 \\ 0 & C' & B \end{pmatrix} \in \mathcal{P}.$$

In a similar way we can vary the other columns of C and so prove the assertion. An analogous argument, using B , shows that for any $m \times n$ matrix C ,

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \in \mathcal{P} \Leftrightarrow \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in \mathcal{P}.$$

(f) If $\mathcal{P}$ is actually a matrix ideal, then for any two square matrices A, B of the same size, $AB \in \mathcal{P}$ whenever either A or B belongs to $\mathcal{P}$.

Assume that $A \in \mathcal{P}$, say. Then $A \oplus B \in \mathcal{P}$ by M.3; using (e) and (c) several times, we obtain in turn

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \rightarrow \begin{pmatrix} A & 0 \\ I & B \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -AB \\ I & B \end{pmatrix} \rightarrow \begin{pmatrix} AB & 0 \\ -B & I \end{pmatrix} \rightarrow \begin{pmatrix} AB & 0 \\ 0 & I \end{pmatrix}.$$

Now an application of M.4 shows that $AB \in \mathcal{P}$.

(g) If A belongs to a matrix ideal $\mathcal{P}$, then the result of permuting the rows or columns of A in any way again belongs to $\mathcal{P}$.

For we can achieve any permutation by multiplying by an appropriate permutation matrix on the left or right.

(h) A matrix ideal $\mathcal{P}$ is proper if and only if $\mathcal{P} \neq \mathfrak{M}(R)$.

If $\mathcal{P}$ is proper, then $1 \notin \mathcal{P}$, by definition; conversely, if $\mathcal{P}$ is improper, then $1 \in \mathcal{P}$ and hence by (f), $A = AI \in \mathcal{P}$ for all $A \in \mathfrak{M}(R)$.

There is an operation on matrix ideals analogous to residual division of ideals. Let $\mathcal{P}$ be a matrix pre-ideal and Σ a subset of $\mathfrak{M}(R)$ which is *directed*: $1_0 \in \Sigma$ and if $P_1, P_2 \in \Sigma$, then there exists $P \in \Sigma$ which is conjugate by permutation matrices, to $P_1 \oplus P'_1$ and $P_2 \oplus P'_2$, for suitable P'_i . We shall define

$$\mathcal{P}/\Sigma = \{A \in \mathfrak{M}(R) \mid A \oplus P \in \mathcal{P} \text{ for some } P \in \Sigma\}. \quad (6)$$

It is easily seen that $\mathcal{P}/\Sigma$ is again a matrix pre-ideal; M.1 is clear and M.2 follows by (3): if $A \oplus X, B \oplus Y \in \mathcal{P}$, then $A \oplus Z, B \oplus Z \in \mathcal{P}$, where Z is conjugate to $X \oplus X', Y \oplus Y'$, and hence $(A \oplus Z) \nabla (B \oplus Z) \in \mathcal{P}$. Now M.3 follows by (d), and when $\mathcal{P}$ is a matrix ideal, we can verify M.4 in the same way. This is still true if $\mathcal{P}$ is merely a matrix pre-ideal but all the unit matrices lie in Σ .

We observe that $\mathcal{P}/\Sigma \supseteq \mathcal{P}$. Suppose that $\mathcal{P}$ is a proper matrix ideal and that $\mathcal{P} \cap \Sigma = \emptyset$; we claim that $\mathcal{P}/\Sigma$ is then proper. For if not, then $A = 1_0 \oplus A \in \mathcal{P}$ for some $A \in \Sigma$, which is a contradiction. Conversely, when

$\mathcal{P} \cap \Sigma \neq \emptyset$, then $\mathcal{P}/\Sigma$ is improper, as we see by reversing the argument. Summing up, we have

PROPOSITION 3.1. *Let R be a ring, $\mathcal{P}$ a matrix pre-ideal and Σ a directed subset of $\mathfrak{M}(R)$. Then the set $\mathcal{P}/\Sigma$ defined by (6) is a matrix pre-ideal; it is a matrix ideal provided that either $\mathcal{P}$ is a matrix ideal or Σ contains all unit matrices. If $\mathcal{P}$ is a matrix ideal, then $\mathcal{P}/\Sigma$ is proper if and only if $\mathcal{P} \cap \Sigma = \emptyset$. ■*

Taking Σ to consist only of unit matrices, we obtain by the same argument

COROLLARY 3.2. *Let $\mathcal{P}$ be any matrix pre-ideal in a ring R and $\mathcal{I}$ the set of all unit matrices over R . Then $\mathcal{P}/\mathcal{I}$ is the least matrix ideal containing $\mathcal{P}$. ■*

Let $(\mathcal{P}_\lambda)$ be any family of matrix ideals; then it is clear that $\mathcal{P} = \bigcap \mathcal{P}_\lambda$ is again a matrix ideal. We can therefore speak of the 'least' matrix ideal containing a given subset $\mathcal{X}$ of $\mathfrak{M}(R)$. This least matrix ideal is also called the matrix ideal generated by $\mathcal{X}$. Similarly, we can define the matrix pre-ideal generated by $\mathcal{X}$. Explicitly this is obtained by forming the determinantal sums, where possible, of all matrices $X \oplus A$ ($X \in \mathcal{X}$, $A \in \mathfrak{M}(R)$) and non-full matrices. For this set is contained in any matrix pre-ideal containing $\mathcal{X}$, and it satisfies M.1 and 2; let us show that it also satisfies M.3. If the set contains C , then (for suitable bracketing)

$$C = B_1 \nabla \dots \nabla B_r, \quad (B_i = X \oplus A \text{ or non-full}), \quad (7)$$

hence for any $P \in \mathfrak{M}(R)$, by (3),

$$C \oplus P = (B_1 \oplus P) \nabla \dots \nabla (B_r \oplus P),$$

with the same bracketing.

Thus the matrix pre-ideal generated by a set $\mathcal{X}$ consists precisely of all determinantal sums of the form (7). Bearing in mind Cor. 3.2, we thus obtain

PROPOSITION 3.3. *The matrix ideal generated by a set $\mathcal{X}$ is proper if and only if the unit matrix (of any size) cannot be expressed as a determinantal sum of non-full matrices and matrices of the form $X \oplus A$, where $X \in \mathcal{X}$ and $A \in \mathfrak{M}(R)$. ■*

Let $\mathcal{E}$ be the matrix pre-ideal generated by the empty set. Clearly this is the least matrix pre-ideal and it consists precisely of all determinantal sums of non-full matrices. By Prop. 3.3 we obtain

COROLLARY 3.4. *A given ring has proper matrix ideals if and only if no unit matrix can be written as a determinantal sum of non-full matrices. ■*

For example, any ring without UGN (and *a fortiori*, any ring without IBN) has no proper matrix ideals, because the unit matrix of a certain size is actually non-full.

Exercises 7.3

1. Give an example to show that property (f) need not hold for matrix pre-ideals.
2. Let α be an ideal in a commutative ring R . Show that the set α^* of matrices A such that $\det A \in \alpha$ is a matrix ideal. If $\mathcal{P}$ is a matrix ideal, show that the set $\mathcal{P}^*$ of elements $\det A$, $A \in \mathcal{P}$, if closed under addition, is an ideal. Verify that $\alpha^{**} = \alpha$, $\mathcal{P}^{**} \supseteq \mathcal{P}$, and that equality holds here when R is Euclidean. What is the relation between $\mathcal{P}^*$ and the 1×1 matrices in $\mathcal{P}$?
3. Show that a field has precisely one proper matrix ideal.
4. A matrix is called *degenerate* if two rows are left (or two columns are right) linearly dependent. Show that if A is non-full, then $A \oplus I$ (for suitable I) is a determinantal sum of hollow and degenerate matrices.
- 5°. Find conditions for the determinantal sum of two nilpotent matrices to be nilpotent.
6. If $C = A \vee B$, where the determinantal sum is with respect to a column, show that $PC = PA \vee PB$, for any P of appropriate size, but that in general $CP \neq AP \vee BP$. However, show that $CP = AP \vee BP$ when P is diagonal.
- 7°. Investigate the notion of a 'sum' or 'join' of matrix ideals; is the lattice of matrix ideals modular? For what rings is the lattice of matrix ideals distributive?
- 8°. Develop a theory of one-sided matrix ideals, admitting determinantal sums only with respect to rows (or only columns), and find an application.

7.4 Prime matrix ideals

We now come to the definition of prime matrix ideals, which stand in the same relation to matrix ideals as do prime ideals to ideals. It will be convenient to begin by defining a multiplication of matrix ideals; this is an associative and commutative operation, like the multiplication of ideals in commutative rings.

Definition. Given two matrix ideals $\mathcal{P}_1, \mathcal{P}_2$ in a ring R , their *product*, denoted by $\mathcal{P}_1 \mathcal{P}_2$, is defined as the matrix ideal generated by all $A_1 \oplus A_2$ with $A_i \in \mathcal{P}_i$ ($i = 1, 2$).

The product so defined is easily seen to be associative. We write $\mathcal{P}_1 \mathcal{P}_2 \mathcal{P}_3$,

etc., for repeated products, and abbreviate $\mathcal{P}\mathcal{P}, \mathcal{P}\mathcal{P}\mathcal{P}, \dots$ as $\mathcal{P}^2, \mathcal{P}^3, \dots$. From property (g) it follows that the product is commutative, and by M.3 it is contained in the intersection of the factors:

$$\mathcal{P}_1\mathcal{P}_2 = \mathcal{P}_2\mathcal{P}_1 \subseteq \mathcal{P}_1 \cap \mathcal{P}_2.$$

The following lemma is often useful in constructing products. In writing a quotient $\mathcal{P}/\Sigma$, if $\Sigma = \{A\}$, we shall simply write $\mathcal{P}/A$.

LEMMA 4.1. In any ring R , let $\mathcal{X}_i \subseteq \mathfrak{M}(R)$ ($i = 1, 2$), $\mathcal{X}$ the set of matrices $A_1 \oplus A_2$ ($A_i \in \mathcal{X}_i$) and $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}$ the matrix ideals generated by $\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}$, respectively. Then $\mathcal{P} = \mathcal{P}_1\mathcal{P}_2$.

In particular, writing (A) for the matrix ideal generated by A , we have $(A)(B) = (A \oplus B)$, for any $A, B \in \mathfrak{M}(R)$.

Proof. Clearly $\mathcal{X} \subseteq \mathcal{P}_1\mathcal{P}_2$, hence $\mathcal{P} \subseteq \mathcal{P}_1\mathcal{P}_2$. To establish equality, let $A_i \in \mathcal{X}_i$, then $A_1 \oplus A_2 \in \mathcal{P}$ by definition, hence $\mathcal{X}_1 \subseteq \mathcal{P}/A_2$ and so $\mathcal{P}_1 \subseteq \mathcal{P}/A_2$. It follows that $B_1 \oplus A_2 \in \mathcal{P}$ for all $B_1 \in \mathcal{P}_1$, so fixing B_1 we have $\mathcal{X}_2 \subseteq \mathcal{P}/B_1$, hence $\mathcal{P}_2 \subseteq \mathcal{P}/B_1$ and so $B_1 \oplus B_2 \in \mathcal{P}$ for all $B_i \in \mathcal{P}_i$, and it follows that $\mathcal{P}_1\mathcal{P}_2 \subseteq \mathcal{P}$. ■

A matrix ideal $\mathcal{P}$ is said to be *prime* if it is proper and

$$A \oplus B \in \mathcal{P} \Rightarrow A \in \mathcal{P} \text{ or } B \in \mathcal{P}.$$

An alternative description is given in

PROPOSITION 4.2. For any matrix ideal $\mathcal{P}$ in a ring R the following three properties are equivalent:

- (a) $\mathcal{P}$ is prime,
- (b) $\mathcal{P}$ is proper and for any matrix ideals $\mathcal{P}_1, \mathcal{P}_2$ we have $\mathcal{P}_1\mathcal{P}_2 \subseteq \mathcal{P} \Rightarrow \mathcal{P}_1 \subseteq \mathcal{P} \text{ or } \mathcal{P}_2 \subseteq \mathcal{P}$,
- (c) $\mathcal{P}$ is proper and for any matrix ideals $\mathcal{P}_1, \mathcal{P}_2$ such that $\mathcal{P}_i \supseteq \mathcal{P}$ we have $\mathcal{P}_1\mathcal{P}_2 \subseteq \mathcal{P} \Rightarrow \mathcal{P}_1 = \mathcal{P} \text{ or } \mathcal{P}_2 = \mathcal{P}$.

Proof. (a) $\Rightarrow$ (b). Let $\mathcal{P}$ be prime and $\mathcal{P}_1\mathcal{P}_2 \subseteq \mathcal{P}$ but $\mathcal{P}_i \not\subseteq \mathcal{P}$ ($i = 1, 2$). Then there exists $A_i \in \mathcal{P}_i$ but $A_i \notin \mathcal{P}$. Since $\mathcal{P}$ is prime, $A_1 \oplus A_2 \notin \mathcal{P}$, but $A_1 \oplus A_2 \in \mathcal{P}_1\mathcal{P}_2 \subseteq \mathcal{P}$, a contradiction. Clearly $\mathcal{P}$ is also proper.

(b) $\Rightarrow$ (c) is clear; to prove (c) $\Rightarrow$ (a), suppose that $A_1 \oplus A_2 \in \mathcal{P}$. Consider the matrix ideal $\mathcal{P}_i$ generated by $\mathcal{P}$ and A_i ; for any $B_i \in \mathcal{P} \cup \{A_i\}$, $B_1 \oplus B_2 \in \mathcal{P}$, hence $\mathcal{P}_1\mathcal{P}_2 \subseteq \mathcal{P}$. By hypothesis, $\mathcal{P}_1$ or $\mathcal{P}_2$ must equal $\mathcal{P}$, so A_1 or A_2 lies in $\mathcal{P}$, i.e. (a). ■

Let $\mathcal{P}$ be any matrix ideal and define its *radical* as the set

$$\sqrt{\mathcal{P}} = \{A \in \mathfrak{M}(R) \mid \bigoplus^r A \in \mathcal{P} \text{ for some } r \geq 1\}.$$

This set is again a matrix ideal: M.1, 3 and 4 clearly hold, and to prove M.2, we note that $\bigoplus^n (A \nabla B)$ is a determinantal sum of terms $C_1 \oplus \dots \oplus C_n$, where each C_i is A or B . Hence if $\bigoplus^r A$ and $\bigoplus^s B$ lie in $\mathcal{P}$, then $\bigoplus^{r+s-1} (A \nabla B) \in \mathcal{P}$; therefore $\sqrt{\mathcal{P}}$ also satisfies M.2 and so is a matrix ideal.

A matrix ideal $\mathcal{P}$ is said to be *semiprime* if $\sqrt{\mathcal{P}} = \mathcal{P}$; e.g. $\sqrt{\mathcal{P}}$ is always semiprime (it is the least semiprime matrix ideal containing $\mathcal{P}$, cf. Ex. 8).

The usual method of constructing prime ideals also works for prime matrix ideals:

THEOREM 4.3. *Let R be any ring, Σ a non-empty subset of $\mathfrak{M}(R)$ closed under diagonal sums and $\mathcal{A}$ any matrix ideal such that $\mathcal{A} \cap \Sigma = \emptyset$. Then there exists a matrix ideal $\mathcal{P}$ which is maximal subject to the conditions $\mathcal{P} \supseteq \mathcal{A}$, $\mathcal{P} \cap \Sigma = \emptyset$, and any such matrix ideal is prime.*

Proof. The collection $\mathcal{C}$ of all matrix ideals containing $\mathcal{A}$ and disjoint from Σ is clearly inductive, so by Zorn's lemma it has a maximal member $\mathcal{P}$, and this satisfies the conditions of the theorem. Any such $\mathcal{P}$ is proper, because $\Sigma \neq \emptyset$; now let $\mathcal{P}_i \supseteq \mathcal{P}$ be matrix ideals such that $\mathcal{P}_1 \mathcal{P}_2 \subseteq \mathcal{P}$. If $\mathcal{P}_i \neq \mathcal{P}$ for $i = 1, 2$, then $\mathcal{P}_i \supset \mathcal{P}$, so by maximality of $\mathcal{P}$, $\mathcal{P}_i \cap \Sigma \neq \emptyset$. Take $A_i \in \mathcal{P}_i \cap \Sigma$, then $A_1 \oplus A_2 \in \mathcal{P} \cap \Sigma$, which is a contradiction. Hence $\mathcal{P}_1$ or $\mathcal{P}_2$ equals $\mathcal{P}$, and so $\mathcal{P}$ is prime, by Prop. 4.2. ■

This theorem shows for example that every maximal proper matrix ideal is prime; we need only take $\Sigma = \mathcal{I} = \{I_n \mid n = 0, 1, \dots\}$. Of course, its main use in the commutative case is the representation of semiprime ideals as the intersection of prime ideals; a corresponding result holds in the general case.

THEOREM 4.4. *If R is any ring and $\mathcal{A}$ a matrix ideal in R , then*

$$\sqrt{\mathcal{A}} = \bigcap \{\mathcal{P} \mid \mathcal{P} \text{ prime } \supseteq \mathcal{A}\}. \quad (1)$$

Thus $\sqrt{\mathcal{A}}$ is the intersection of all prime matrix ideals containing $\mathcal{A}$.

Proof. If $\mathcal{P}$ is a prime matrix ideal and $\mathcal{P} \supseteq \mathcal{A}$, then $\sqrt{\mathcal{A}} \subseteq \sqrt{\mathcal{P}} = \mathcal{P}$, hence $\sqrt{\mathcal{A}} \subseteq \bigcap \mathcal{P}$ where the intersection is taken over all prime matrix ideals $\mathcal{P}$ containing $\mathcal{A}$. To establish equality, let $A \notin \sqrt{\mathcal{A}}$ be given and consider the set Σ_A of all diagonal sums of copies of A . By definition, $\mathcal{A} \cap \Sigma_A = \emptyset$,

hence by Th. 4.3 there is a maximal matrix ideal $\mathcal{P}_0$ containing $\mathcal{A}$ and disjoint from Σ_A , and $\mathcal{P}_0$ is prime. Thus $A \notin \bigcap \mathcal{P}$ and this proves equality in (1). ■

Taking $\mathcal{A}$ to be semiprime, we obtain

COROLLARY 4.5. *In any ring a matrix ideal is semiprime if and only if it is an intersection of prime matrix ideals.* ■

Let R be a ring and K an R -field. It is straightforward to show that the singular kernel of K is a prime matrix ideal. Our object will be to show that conversely, every prime matrix ideal occurs as singular kernel of some R -field. For the proof we shall need a construction of groups from semigroups and an axiomatic description of fields, which we bring as separate lemmas. We recall that semigroups, in contrast to monoids, need not possess a unit element.

LEMMA 4.6. *Let M be a non-empty semigroup with a unary operation $x \mapsto x^\circ$ and a subsemigroup M_0 such that (i) $x^{\circ\circ} = x$, (ii) $x \in M_0 \Rightarrow x^\circ \in M_0$, (iii) $xy \in M_0 \Rightarrow yx \in M_0$ and (iv) $xx^\circ \in M_0$ for all $x \in M$. Then the relation defined on M by*

$$x \sim y \text{ whenever } xzy^\circ \in M_0 \text{ for some } z \in M_0, \quad (2)$$

is a congruence on M whose quotient is a group. Moreover, $x \sim e$, where e is the neutral element, if and only if $xu \in M_0$ for some $u \in M_0$.

Proof. By (iv) M_0 is non-empty. If $xy^\circ \in M_0$ and $u \in M_0$, then by (iii), $y^\circ x \in M_0$, hence $y^\circ xu \in M_0$ and so $xuy^\circ \in M_0$; therefore $x \sim y$ whenever $xy^\circ \in M_0$.

To prove that (2) is an equivalence, we have for any $x \in M$, $xx^\circ \in M_0$, hence $x \sim x$ by the remark just made. Next if $x \sim y$, $y \sim z$, say $u, xuy^\circ, v, yvz^\circ \in M_0$, then by (iii) and (iv) $y^\circ y \in M_0$ and so $uy^\circ yv \in M_0$. Now $x \cdot uy^\circ yv \cdot z^\circ \in M_0$, therefore $x \sim z$, i.e. ' $\sim$ ' is transitive. To establish symmetry, assume that $x \sim y$, say $u, xuy^\circ \in M_0$. Then $y^\circ xu, uy^\circ x \in M_0$ and hence $y^\circ xu^2 y^\circ x \in M_0$. Further, $y^\circ xu \cdot u \in M_0$ hence $xu^2 y^\circ \in M_0$ and so $yy^\circ \cdot xu^2 y^\circ \cdot xx^\circ \in M_0$. But $y^\circ xu^2 y^\circ x \in M_0$, therefore $y \sim x$ and the symmetry of ' $\sim$ ' is proved.

Next we observe that

$$x \sim y \Rightarrow x^\circ \sim y^\circ. \quad (3)$$

For if $u, xuy^\circ \in M_0$, then $x^\circ xuy^\circ y \in M_0$, hence $x^\circ \sim y^\circ$.

Given any $x, y \in M$, we have $yy^\circ x^\circ x \in M_0$, hence $y^\circ x^\circ xy \in M_0$, and so by the

initial remark,

$$y^o x^o \sim (xy)^o. \quad (4)$$

Now let $x \sim x_1$, $y \sim y_1$, say $u, xux_1^o, v, yvy_1^o \in M_0$. Then $x^o x \cdot yvy_1^o \in M_0$ hence $xyvy_1^o x^o \in M_0$ and $xux_1^o \in M_0$, so

$$xy \cdot yvy_1^o x^o \cdot xux_1^o \cdot x_1 y_1 \cdot (x_1 y_1)^o \in M_0. \quad (5)$$

On the other hand, $w = x^o xux_1^o x_1 \in M_0$ and $vy_1^o y_1 \in M_0$, hence $y_1 vy_1^o \in M_0$, therefore $y_1 vy_1^o w \in M_0$, and $vy_1^o wy_1 \in M_0$. But this is just the central factor in (5), so $xy \sim x_1 y_1$ and we have established that ' $\sim$ ' is a congruence.

Let $G = M/\sim$ be the quotient semigroup and denote the class containing M_0 by e . Then in G , $ex = xe = x$, $x^o x = xx^o = e$ for all $x \in G$ (where x^o is the class of all a^o , for a in x ; this makes sense by (3)). This shows G to be a group. Now if $xu \in M_0$ for some $u \in M_0$, then clearly $x \sim e$. Conversely, if $x \sim e$, then $x \sim u$ for some $u \in M_0$, so $xvu^o \in M_0$ and here $vu^o \in M_0$. ■

LEMMA 4.7. *Let K be an additive group, not necessarily abelian, with neutral element 0. Suppose that $K^\times = K \setminus \{0\}$ is a group under multiplication, and this is extended to K by putting $a0 = 0a = 0$. If, further,*

$$(a+1)b = ab+b, \quad b(a+1) = ba+b, \quad \text{for all } a, b \in K, \quad (6)$$

then K is a field.

Proof. We need only check the distributive laws and the commutativity of addition. To prove

$$(x+y)z = xz+yz,$$

we note that for $y = 0$ both sides reduce to xz . When $y \neq 0$ we can write $a = xy^{-1}$ and then $(x+y)z = (ay+y)z = (a+1)yz = ayz+yz = xz+yz$. The other distributive law follows by symmetry and the commutative law of addition follows by a comparison of

$$(a+1)(b+1) = (a+1)b+a+1 = ab+b+a+1,$$

with

$$(a+1)(b+1) = a(b+1)+b+1 = ab+a+b+1. \quad \blacksquare$$

We now come to the main result of this section.

THEOREM 4.8. *Let R be any ring and $\mathcal{P}$ a prime matrix ideal. Then there is an epic R -field K with singular kernel $\mathcal{P}$; in fact, K is the residue class field of the local ring $R_{\mathcal{P}}$.*

Proof. We shall construct the field K as set of equivalence classes of admissible matrices. Consider the set M of all matrices A of index 1, say $A = (A_0, A_*, A_\infty)$ such that the denominator (A_*, A_∞) is not in $\mathcal{P}$. It will be convenient not to distinguish between matrices related by the following *trivial* operations: (i) left multiplication by an invertible matrix and (ii) right multiplication by an invertible matrix of the form

$$Q^* = \begin{pmatrix} 1 & 0 & 0 \\ Q_0 & Q_* & Q_\infty \\ 0 & 0 & 1 \end{pmatrix}. \quad (7)$$

A matrix of the form (7) is said to be *bordered*; thus a bordered matrix agrees with the unit matrix in the first and last row. We note that (ii) consists of the operations (a) right multiplication of the core A_* by an invertible matrix and (b) the addition of a right multiple of A_* to A_0 and to A_∞ . We shall use an arrow to indicate the application of (i) or (ii).

On M we define a binary operation by

$$A \boxplus B = \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 \\ A_0 & 0 & -A_\infty & A_* & A_\infty \end{pmatrix}. \quad (8)$$

This is unaffected by trivial operations on A, B and hence is well-defined. Further, it is clear that $A \boxplus B \in M$. Let M_0 be the set of all $A \in M$ such that the numerator (A_0, A_*) lies in $\mathcal{P}$, and put $A^0 = (A_0, A_*, -A_\infty)$. We shall verify that Lemma 4.6 applies. It is clear that $A^{00} = A$ and $A \in M_0 \Rightarrow A^0 \in M_0$. Next we have

$$\begin{aligned} B \boxplus A &= \begin{pmatrix} A_0 & A_* & A_\infty & 0 & 0 \\ B_0 & 0 & -B_\infty & B_* & B_\infty \end{pmatrix} \rightarrow \begin{pmatrix} B_0 & 0 & -B_\infty & B_* & B_\infty \\ A_0 & A_* & A_\infty & 0 & 0 \end{pmatrix} \\ &\rightarrow \begin{pmatrix} B_0 & B_* & -B_\infty & 0 & 0 \\ A_0 & 0 & A_\infty & A_* & A_\infty \end{pmatrix}, \end{aligned}$$

and a comparison with (8) shows that $A \boxplus B \in M_0 \Leftrightarrow B \boxplus A \in M_0$, thus (iii) holds. Finally

$$A \boxplus A^0 = \begin{pmatrix} A_0 & A_* & -A_\infty & 0 & 0 \\ A_0 & 0 & -A_\infty & A_* & A_\infty \end{pmatrix} \rightarrow \begin{pmatrix} A_0 & A_* & -A_\infty & 0 & 0 \\ 0 & 0 & 0 & A_* & A_\infty \end{pmatrix},$$

and the numerator in the last expression is hollow; hence it lies in $\mathcal{P}$ and $A \boxplus A^0 \in M_0$.

Thus the hypotheses of Lemma 4.6 hold and we obtain a group consisting of equivalence classes of admissible matrices. We shall write $[A]$ for the class of A and note the following rules.

If P is a square matrix not in $\mathcal{P}$ and A is admissible, then

$$\begin{pmatrix} A_0 & Q & A_* & A_\infty \\ 0 & P & 0 & 0 \end{pmatrix} \sim (A_0 \quad A_* \quad A_\infty) \quad (9)$$

and

$$\begin{pmatrix} A_0 & 0 & A_* & A_\infty \\ B_0 & P & B_* & B_\infty \end{pmatrix} \sim (A_0 \quad A_* \quad A_\infty). \quad (10)$$

For we have

$$\begin{pmatrix} A_0 & A_* & -A_\infty & 0 & 0 & 0 \\ A_0 & 0 & -A_\infty & Q & A_* & A_\infty \\ 0 & 0 & 0 & P & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} A_0 & A_* & -A_\infty & 0 & 0 & 0 \\ 0 & 0 & 0 & Q & A_* & A_\infty \\ 0 & 0 & 0 & P & 0 & 0 \end{pmatrix},$$

and the numerator of the last expression is hollow and therefore lies in $\mathcal{P}$. Similarly for (10) we have

$$\begin{pmatrix} A_0 & A_* & -A_\infty & 0 & 0 & 0 \\ A_0 & 0 & -A_\infty & 0 & A_* & A_\infty \\ B_0 & 0 & -B_\infty & P & B_* & B_\infty \end{pmatrix} \rightarrow \begin{pmatrix} A_0 & A_* & -A_\infty & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_* & A_\infty \\ B_0 & B_* & -B_\infty & P & B_* & B_\infty \end{pmatrix},$$

and this expression again has a hollow numerator (this is seen most easily by observing that the core A_* has index -1).

Next we note that

$$[A] = 0 \text{ if and only if } A \in M_0. \quad (11)$$

For if $A \in M_0$, then $[A] = 0$ by definition. Conversely, if $[A] = 0$, then $A \boxplus B \in M_0$ for some $B \in M_0$, hence the numerator of $A \boxplus B$ is in $\mathcal{P}$. But this numerator has the form

$$\begin{pmatrix} B_0 & B_* & B_\infty & 0 \\ A_0 & 0 & -A_\infty & A_* \end{pmatrix} = \begin{pmatrix} B_0 & B_* & 0 & 0 \\ A_0 & 0 & -A_\infty & A_* \end{pmatrix} \nabla \begin{pmatrix} B_0 & B_* & B_\infty & 0 \\ A_0 & 0 & 0 & A_* \end{pmatrix}.$$

Here the first term on the right is in $\mathcal{P}$ because $B \in M_0$. Therefore the second term is in $\mathcal{P}$, hence by applying a trivial operation and property (e) of matrix ideals, $(B_*, B_\infty) \oplus (A_0, A_*) \in \mathcal{P}$. Since B is admissible, $(B_*, B_\infty) \notin \mathcal{P}$, and so $(A_0, A_*) \in \mathcal{P}$, i.e. $A \in M_0$.

We also note that

$$[A] = 1 \text{ if and only if } (A_0 + A_\infty, A_*) \in \mathcal{P}. \quad (12)$$

For by (11), $[A] = 1$ holds if and only if

$$\begin{pmatrix} 1 & -1 & 0 & 0 \\ A_0 & A_\infty & A_* & A_\infty \end{pmatrix} \in M_0,$$

i.e. $(A_0 + A_\infty, A_*, A_\infty) \in M_0$, by an application of a trivial operation and (9). Now (12) is clear.

We define a multiplication on M by the rule

$$A \cdot B = \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 \\ 0 & 0 & A_0 & A_* & A_\infty \end{pmatrix}. \quad (13)$$

It is clear that this is associative and compatible with the trivial operations. To show that it is compatible with ' $\sim$ ', let $A \sim A'$; by a series of trivial operations we have

$$\begin{aligned} (A \cdot B)^0 \boxplus A' \cdot B &= \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & A'_0 & A'_* & A'_\infty & 0 & 0 & 0 & 0 \\ B_0 & 0 & 0 & 0 & 0 & B_* & B_\infty & 0 & 0 \\ 0 & 0 & 0 & 0 & A_\infty & 0 & A_0 & A_* & -A_\infty \end{pmatrix} \\ &\rightarrow \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & A'_0 & A'_* & A'_\infty & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & B_* & B_\infty & 0 & 0 \\ 0 & 0 & A_0 & 0 & A_\infty & 0 & A_0 & A_* & -A_\infty \end{pmatrix}; \end{aligned}$$

the numerator of this matrix is a block triangular sum of (B_0, B_*) , (B_*, B_∞) and the numerator of $A^0 \boxplus A'$, which is in $\mathcal{P}$ by hypothesis. Hence $A \cdot B \sim A' \cdot B$ and similarly $B \sim B'$ implies $A \cdot B \sim A \cdot B'$. Thus we have a monoid on $M/\sim$ with multiplication (13). Next we define, for any $A \notin M_0$, $\bar{A} = (A_\infty, A_*, A_0)$; this is again admissible and $\bar{\bar{A}} = A$. Moreover, we have

$$A \cdot \bar{A} = \begin{pmatrix} A_\infty & A_* & A_0 & 0 & 0 \\ 0 & 0 & A_0 & A_* & A_\infty \end{pmatrix} \rightarrow \begin{pmatrix} A_\infty & A_* & A_0 & 0 & 0 \\ -A_\infty & 0 & 0 & A_* & A_\infty \end{pmatrix}$$

and this represents 1, by (12), because

$$\begin{pmatrix} A_\infty & A_* & A_0 & 0 \\ 0 & 0 & 0 & A_* \end{pmatrix} \in \mathcal{P}.$$

Hence $K \setminus \{0\} = K^\times$ is a group under the multiplication (13). The operation (13) makes sense even if $[A]$ or $[B]$ is 0, and in that case it is clear that $A \cdot B \in M_0$, so that we have $x0 = 0x = 0$. It remains to verify the distributive

laws in the form (6). We have

$$\begin{aligned}
 (A \boxplus 1) \cdot B &= \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & A_0 & -A_\infty & A_* & A_\infty \end{pmatrix} \\
 &\rightarrow \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & A_0 - A_\infty & -A_\infty & A_* & A_\infty \end{pmatrix} \\
 &\sim \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 \\ 0 & 0 & A_0 - A_\infty & A_* & A_\infty \end{pmatrix};
 \end{aligned}$$

$$\begin{aligned}
 AB \boxplus B &= \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 & 0 & 0 \\ B_0 & 0 & 0 & B_* & B_\infty & 0 & 0 \\ 0 & 0 & -A_\infty & 0 & A_0 & A_* & A_\infty \end{pmatrix} \\
 &\rightarrow \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & B_* & B_\infty & 0 & 0 \\ 0 & 0 & A_0 - A_\infty & 0 & A_0 & A_* & A_\infty \end{pmatrix},
 \end{aligned}$$

and this reduces to the same matrix as before. Similarly,

$$\begin{aligned}
 B(A \boxplus 1) &= \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ A_0 & -A_\infty & A_* & A_\infty & 0 & 0 \\ 0 & 0 & 0 & B_0 & B_* & B_\infty \end{pmatrix} \\
 &\rightarrow \begin{pmatrix} 0 & -1 & 0 & 0 & 0 & 0 \\ A_0 - A_\infty & -A_\infty & A_* & A_\infty & 0 & 0 \\ 0 & 0 & 0 & B_0 & B_* & B_\infty \end{pmatrix} \\
 &\sim \begin{pmatrix} A_0 - A_\infty & A_* & A_\infty & 0 & 0 \\ 0 & 0 & B_0 & B_* & B_\infty \end{pmatrix};
 \end{aligned}$$

$$\begin{aligned}
 BA \boxplus B &= \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 & 0 & 0 \\ A_0 & 0 & 0 & A_* & A_\infty & 0 & 0 \\ 0 & 0 & -B_\infty & 0 & B_0 & B_* & B_\infty \end{pmatrix} \\
 &\rightarrow \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 & 0 & 0 \\ A_0 - A_\infty & 0 & 0 & A_* & A_\infty & 0 & 0 \\ 0 & 0 & 0 & 0 & B_0 & B_* & B_\infty \end{pmatrix},
 \end{aligned}$$

and this again reduces to the same matrix as before. Thus we have indeed a field, by Lemma 4.7. We denote this field by K and define a map $f: R \rightarrow K$ by

$a \mapsto [a, -1]$. This is a homomorphism since $1f = [1, -1]$,

$$af + bf = \begin{bmatrix} b & -1 & 0 \\ a & 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ a+b & 1 & -1 \end{bmatrix} = [a+b, -1] = (a+b)f,$$

$$af \cdot bf = \begin{bmatrix} b & -1 & 0 \\ 0 & a & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ ab & a & -1 \end{bmatrix} = [ab, -1] = (ab)f.$$

We claim that for any $p \in K$, if p is represented by the admissible matrix A , then there is an equation

$$Au = 0, \quad \text{where } u_0 = 1, u_\infty = p. \quad (14)$$

To verify (14) we need to work out the left-hand side in K , bearing in mind that u_∞ is represented by A , while u_i is represented by the matrix obtained from A by interchanging the columns A_i and A_∞ . This is a straightforward though lengthy calculation, best illustrated by a case of low order, say $m = 2$. Using (8) and (13), we find the matrix for the i th row on the left of (14) to be

$$\begin{pmatrix} A_0 & A_1 & A_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{i2} & -1 & 0 & 0 & 0 & 0 \\ A_0 & 0 & 0 & 0 & A_2 & A_1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & a_{i1} & -1 & 0 \\ a_{i0} & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

If we subtract the first row block from the third, and then add the 5th and 6th column blocks to the 3rd and 2nd respectively, we obtain, after an operation (9),

$$\begin{pmatrix} A_0 & A_1 & A_2 & 0 & 0 & 0 \\ 0 & 0 & a_{i2} & -1 & 0 & 0 \\ 0 & a_{i1} & 0 & 1 & -1 & 0 \\ a_{i0} & 0 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

This matrix lies in M_0 since its numerator is non-full, as we see by adding the last 3 rows to the i th row. Hence the matrix represents zero by (11); the general case is entirely analogous and it proves (14). It also shows that K is an epic R -field.

Now to prove that $\text{Ker } f = \mathcal{P}$, let $P \notin \mathcal{P}$, then (e_i, P) is an admissible system, so we have an equation

$$e_i + Pu_i = 0, \quad (15)$$

and this shows (e_i, P) to be admissible for the last component of u_i . By permuting the columns of P we obtain admissible systems for the other components of u_i . Let U be the matrix with the u_i as columns; the equations

(15) show that

$$I + PU = 0,$$

hence P is invertible over K and so $P \notin \text{Ker } f$. Next assume $P \in \mathcal{P}$ and denote by $P^{(i)}$ the matrix obtained from P by omitting the i th row and first column. If $P^{(i)} \in \mathcal{P}$ for all i , then by induction on the order of P , $P^{(i)} \in \text{Ker } f$ and hence $P \in \text{Ker } f$, as we wished to show. So we may suppose that $P^{(i)} \notin \mathcal{P}$ for some i , say $i = 1$. Then (P, e_1) is admissible, and it is admissible for 0, because $P \in \mathcal{P}$. Thus e_1 occurs with zero coefficient in the equation corresponding to (14) and we have

$$Pv = 0, \quad \text{where } v_0 = 1.$$

This shows that P is a zero-divisor, hence $P \in \text{Ker } f$ and it follows that $\text{Ker } f = \mathcal{P}$ as claimed. Finally $R_{\mathcal{P}}$ is a local ring with residue class field K , by Th. 2.2. ■

This result tells us that for any ring R there is a natural bijection between the (isomorphism classes of) epic R -fields and prime matrix ideals; moreover, by Th. 2.4, specialization corresponds to inclusion. Thus the set $X = X(R)$ of isomorphism classes of epic R -fields, partially ordered by specialization, may also be thought of as the set of all prime matrix ideals of R , partially ordered by inclusion. This way of looking at X also assures us that it is really a set.

Each square matrix A over R determines a subset of X , called its *singularity support*:

$$D(A) = \{x \in X \mid A \notin \mathcal{P}_x\}, \quad (16)$$

where $\mathcal{P}_x$ is the prime matrix ideal corresponding to x . It is easily verified that

$$D(A \oplus B) = D(A) \cap D(B), \quad D(I) = X.$$

Hence the collection of sets $D(A)$ admits finite intersections, and so forms a base for the open sets of a topology on X . The topological space $X(R)$ so defined is called the *field-spectrum* of R . It is analogous to the usual prime spectrum of a commutative ring, to which it reduces when R is taken to be commutative. We shall not pursue this point of view here but merely note that X satisfies the conditions for a spectral space, which characterize the prime spectrum of a commutative ring (cf. Ex. 13, see Hochster [69], Cohn [72']).

Exercises 7.4

1. Let $\mathcal{P}$ be a matrix ideal and $\mathcal{X} \subseteq \mathfrak{M}(R)$. Show that the set $\mathcal{P}:\mathcal{X} = \{A \in \mathfrak{M}(R) \mid A \oplus X \in \mathcal{P} \text{ for all } X \in \mathcal{X}\}$ is a matrix ideal, the quotient of $\mathcal{P}$ by $\mathcal{X}$ (not to be confused with $\mathcal{P}/\mathcal{X}$), and that it may be characterized as the largest matrix ideal satisfying $\mathcal{X}(\mathcal{P}:\mathcal{X}) \subseteq \mathcal{P}$. If the matrix ideal generated by X is denoted by (X) , show that $\mathcal{P}:X = \mathcal{P}:(X)$.

2. Show that if $\mathcal{P}, \mathcal{P}_\lambda$ are matrix ideals and $\mathcal{X} \subseteq \mathfrak{M}(R)$, then $\mathcal{P}:\mathcal{X} = \{\mathcal{P}:\{X\} \mid X \in \mathcal{X}\}$, $(\mathcal{P}_1:\mathcal{P}_2):\mathcal{P}_3 = \mathcal{P}_1:\mathcal{P}_2\mathcal{P}_3$, $(\cap \mathcal{P}_\lambda):\mathcal{X} = \cap (\mathcal{P}_\lambda:\mathcal{X})$.

3. A matrix ideal $\mathcal{Q}$ is said to be *primary* if $A \oplus B \in \mathcal{Q}$, $A \notin \mathcal{Q}$ implies that $\bigoplus_r B \in \mathcal{Q}$ for some r . Show that for any primary matrix ideal $\mathcal{Q}$, $\sqrt{\mathcal{Q}}$ is prime (in this case $\mathcal{Q}$ is also called $\sqrt{\mathcal{Q}}$ -primary). Give an example of a matrix ideal $\mathcal{Q}$ such that $\sqrt{\mathcal{Q}}$ is prime but $\mathcal{Q}$ is not primary. Show that if $\mathcal{Q}$ is $\mathcal{P}$ -primary, then for any $\mathcal{X} \subseteq \mathfrak{M}(R)$, $\mathcal{Q}:\mathcal{X}$ is $\mathcal{P}$ -primary, and is proper if and only if $\mathcal{X} \not\subseteq \mathcal{Q}$.

4. Let $\mathcal{A}$ be a matrix ideal which has a primary decomposition: $\mathcal{A} = \mathcal{Q}_1 \cap \dots \cap \mathcal{Q}_r$, where $\mathcal{Q}_i$ is $\mathcal{P}_i$ -primary. Show that $\sqrt{\mathcal{A}} = \mathcal{P}_1 \cap \dots \cap \mathcal{P}_r$, and show that $\mathcal{A}$ also has a primary decomposition in which the $\mathcal{P}_i$ are all different and no $\mathcal{Q}_i$ contains the intersection of all the others (such a primary decomposition is called *irredundant*). Show that any two irredundant primary decompositions of $\mathcal{A}$ have the same number of terms, and these can be numbered so that corresponding primes are equal. If $\mathcal{P}_1, \dots, \mathcal{P}_i$ are the primes disjoint from a given set $\mathcal{X}$, show that the corresponding intersection $\mathcal{Q}_1 \cap \dots \cap \mathcal{Q}_i$ is independent of the particular decomposition, but not necessarily each $\mathcal{Q}_i$.

5. Show that the union and intersection of any chain of prime matrix ideals are again prime. Show that for prime matrix ideals $\mathcal{A} \subset \mathcal{B}$ there exist prime matrix ideals $\mathcal{P}_1, \mathcal{P}_2$ such that $\mathcal{A} \subseteq \mathcal{P}_1 \subset \mathcal{P}_2 \subseteq \mathcal{B}$, but there is no prime matrix ideal between $\mathcal{P}_1$ and $\mathcal{P}_2$.

6. Let $\mathcal{P}$ be a minimal prime matrix ideal in a ring R . Show that for each $A \in \mathcal{P}$ there exists a matrix $C \notin \mathcal{P}$ and a positive integer r such that $C \bigoplus_r A$ is a determinantal sum of non-full matrices.

7°. Give an example of a saturated set of matrices (cf. Ex. 7.2.8) whose complement is not a union of prime matrix ideals. (Hint: Consider the invertible matrices in a general ring.) If $\mathcal{A}$ is a matrix ideal in a ring R and $\mathcal{P}_1, \mathcal{P}_2$ are prime matrix ideals such that $\mathcal{A} \subseteq \mathcal{P}_1 \cup \mathcal{P}_2$, does it follow that $\mathcal{A} \subseteq \mathcal{P}_1$ or $\mathcal{A} \subseteq \mathcal{P}_2$?

8. Show that for any matrix ideal $\mathcal{A}$ in a ring R the following are equivalent: (a) $\mathcal{A}$ is semiprime, (b) for any matrix ideal $\mathcal{Q}$, $\mathcal{Q}^2 \subseteq \mathcal{A} \Rightarrow \mathcal{Q} \subseteq \mathcal{A}$ and (c) for any matrix ideal $\mathcal{Q}$, $\mathcal{Q} \supseteq \mathcal{A} \supseteq \mathcal{Q}^2 \Rightarrow \mathcal{Q} = \mathcal{A}$. Verify that for any matrix ideal $\mathcal{A}$, $\sqrt{\mathcal{A}}$ is semiprime and that $\sqrt{\sqrt{\mathcal{A}}} = \sqrt{\mathcal{A}}$. Deduce that $\sqrt{\mathcal{A}}$ is the least semiprime matrix ideal $\supseteq \mathcal{A}$.

9°. Investigate rings in which every matrix ideal is finitely generated (e.g. examine the commutative case). Investigate rings in which every finitely generated matrix ideal can be generated by a single matrix.

10. (Gerasimov [79]) Let R be a ring, $\Sigma \subseteq R$ and $\mathcal{Z}$ the least matrix ideal (generated by the empty set). Show that the universal Σ -inverting map $\lambda: R \rightarrow R_\Sigma$ is injective if and only if $\mathcal{Z}/\Sigma$ contains no 1×1 matrix other than 0, where Σ denotes the saturation of Σ .

11. A topological space is called *irreducible* if it is non-empty and cannot be written as a union of two proper closed subsets. Show that if $X = \mathbf{X}(R)$ is the field-spectrum of a ring R , then a closed subset of X is irreducible if and only if it is the closure of a single point.

12. Show that $D(A)$ is compact, for any matrix A . (Hint: Assume that $D(A) = \bigcup D(A_i)$ and denote by $\mathcal{C}$ the matrix ideal generated by the A_i . Show that $A \in \sqrt{\mathcal{C}}$ and deduce that $A \in \sqrt{\mathcal{C}_0}$, where $\mathcal{C}_0$ is the matrix ideal generated by a finite subfamily of the A_i .)

13*. Let R be a ring and $X = X(R)$ its field spectrum. Verify that X satisfies Hochster's axioms for a spectral space (Hochster [69]):

- (i) X is a T_0 -space, i.e. given $x, y \in X$, $x \neq y$, there is an open set containing precisely one of x, y ;
- (ii) X is compact;
- (iii) the compact open sets admit finite intersections and form a base for the open sets in X and
- (iv) any irreducible closed subset is the closure of a single point.

14. Show that a ring homomorphism $f: R \rightarrow S$ induces a continuous map $f^*: X(S) \rightarrow X(R)$.

15°. Describe the rings whose field spectrum is a T_1 -space (i.e. every prime matrix ideal is maximal).

16°. Which rings have a Noetherian field spectrum (i.e. ACC on prime matrix ideals)?

17*. Given $A \in \mathfrak{M}(R)$, define Σ_A as the complement of $\bigcup \{\mathcal{P} \mid \mathcal{P} \in D(A)\}$ and write $R(A)$ for the universal localization at Σ_A . Show that there is a canonical homomorphism $R_A \rightarrow R(A)$, which is an isomorphism when R is commutative, but not in general. Show also that the canonical homomorphism $\phi: R \rightarrow R(A)$ induces a homeomorphism $\phi^*: X(R(A)) \rightarrow D(A)$. Given $A, B \in \mathfrak{M}(R)$, verify that $D(A) \supseteq D(B) \Rightarrow \Sigma_A \subseteq \Sigma_B$ and hence obtain a homomorphism $\phi_{AB}: R(A) \rightarrow R(B)$. Show that $\{R(A), \phi_{AB}\}$ is a presheaf of rings over X (cf. e.g. Macdonald [68]). Verify that the corresponding sheaf $\tilde{R}$ has as stalk over $x \in X$ the universal localization at $\mathcal{P}_x$ (cf. Cohn [72', 79]).

18°. Let R be a fir which is not right principal and let $\mathfrak{a}$ be a 2-sided ideal whose rank as right ideal is a finite number $r > 1$. If $u = (u_1, \dots, u_r)$ is a basis of $\mathfrak{a}$, verify that the map $\rho: R \rightarrow R_r$ defined by $au = u \cdot \rho(a)$ is a homomorphism. Determine conditions under which this map is honest.

7.5 Fields of fractions

The results of the last section can be used to answer various questions about the embeddability of rings in fields, or more generally, the existence of R -fields.

Let R be a ring; clearly R -fields exist if and only if there exist epic R -fields, or equivalently, by Th. 4.8, prime matrix ideals. To find whether R has prime matrix ideals we go back to the method of generating matrix ideals described

in 7.3. Let us denote by $\mathcal{L}_0$ the set of all determinantal sums of non-full matrices. Thus $A \in \mathcal{L}_0$ precisely if

$$A = C_1 \nabla \dots \nabla C_r,$$

where each C_i is non-full and the right-hand side is suitably bracketed. From the description in 7.3 it is clear that $\mathcal{L}_0$ is the least matrix pre-ideal in R . Let $\mathcal{L} = \mathcal{L}_0/\mathcal{I}$ be the matrix ideal generated by $\mathcal{L}_0$ and put $\mathcal{N} = \sqrt{\mathcal{L}}$; then $\mathcal{N}$ is proper if and only if $\mathcal{L}$ is proper. By Th. 4.4, $\mathcal{N}$ is the intersection of all prime matrix ideals in R , so $\mathcal{N}$ is proper if and only if R has prime matrix ideals. By Th. 4.8 this means that R has epic R -fields if and only if $\mathcal{N}$ is proper. We therefore obtain the following criterion for the existence of R -fields:

THEOREM 5.1. *Let R be any ring. Then there exists a homomorphism of R into a field if and only if no unit matrix in R can be written as a determinantal sum of non-full matrices. ■*

This includes the well-known necessary condition: if a ring can be mapped into a field, then the unit matrix (of any size) is full, i.e. R has UGN.

Secondly, we obtain a criterion for the invertibility of a matrix, which results from the following more general formulation.

THEOREM 5.2. *Let R be a ring and P, Q any square matrices over R . Then there is a homomorphism from R to a field mapping P to an invertible matrix and Q to a singular matrix if and only if no diagonal sum $I \overset{r}{\oplus} P$ can be written as a determinantal sum of non-full matrices and matrices $Q \oplus B$, where $B \in \mathfrak{M}(R)$.*

Proof. The condition for a homomorphism of the required sort to exist is that there should be a prime matrix ideal containing Q but not P . Let (Q) denote the matrix ideal generated by Q ; there is a prime matrix ideal containing Q but not P if and only if $P \notin \sqrt{(Q)}$. So the required condition is that $I \overset{r}{\oplus} P \notin (Q)$, i.e. there is no equation

$$I \overset{r}{\oplus} P = C_1 \nabla \dots \nabla C_s, \quad (C_i \text{ non-full or of form } Q \oplus B_i). \quad \blacksquare \quad (1)$$

In particular, taking $P = A, Q = 0$, we obtain

COROLLARY 5.3. *For any square matrix A over a ring R there is an R -field inverting A if and only if no diagonal sum $I \overset{r}{\oplus} A$ can be written as a determinantal sum of non-full matrices. ■*

Secondly, take $P = I$, $Q = A$, then we find

COROLLARY 5.4. *For any square matrix A over a ring R there is an R -field over which A becomes singular if and only if no unit matrix I can be written as a determinantal sum of non-full matrices and matrices of the form $A \oplus B$, where B is any matrix. ■*

From these results it is easy to obtain a criterion for the embeddability of a ring in a field. We recall that an integral domain R is embeddable in a field if and only if for each $a \in R^\times$ there is an a -inverting homomorphism into a field (cf. Cohn [71] or Ex. 5). By Cor. 5.3, this holds if and only if there is no equation

$$I \oplus aI = C_1 \nabla \dots \nabla C_r \quad (C_i \text{ non-full}). \quad (2)$$

In particular, aI cannot be expressed in this way. Conversely, if there is an expression (2) with I of order r and aI of order s , we multiply both sides by $aI_r \oplus I_s$ and observe that the determinantal sum is distributive with respect to multiplication by diagonal matrices. Thus we obtain aI as a determinantal sum of non-full matrices. This proves

COROLLARY 5.5. *A ring R can be embedded in a field if and only if it is an integral domain and no non-zero scalar matrix can be written as a determinantal sum of non-full matrices. ■*

An alternative formulation is given in

THEOREM 5.6. *A ring R is embeddable in a field if and only if no diagonal matrix with non-zero elements on the main diagonal can be written as a determinantal sum of non-full matrices.*

For if $ab = 0$, then

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} a & 0 \\ 1 & b \end{pmatrix} \nabla \begin{pmatrix} 0 & 0 \\ -1 & b \end{pmatrix} = \begin{pmatrix} a \\ 1 \end{pmatrix} (1 \quad b) \nabla \begin{pmatrix} 0 \\ 1 \end{pmatrix} (-1 \quad b),$$

and here both matrices on the right are non-full. Thus the condition of Th. 5.6 is sufficient to exclude zero-divisors, and by Cor. 5.5 it is sufficient for embeddability in a field. The converse is clear. ■

It may be of interest to note that the condition of Cor. 5.5, apart from the absence of zero-divisors, is in the form of quasi-identities, as required by general theory (cf. e.g. Cohn [81], p. 235, and Ex. 6, below).

We have already found a condition for the existence of a universal R -field: it was that the radical of the least matrix ideal, the set $\mathcal{N}$ constructed before Th. 5.1, should be prime. Moreover, there will be a universal field of fractions if and only if, further, the set $\mathcal{N}$ contains no non-zero elements of R . But frequently we are interested in the special case when there is a universal field of fractions in which every full matrix becomes invertible. This is the maximal set that can be inverted, since no non-full matrix can ever be inverted over an R -field.

We recall that a ring homomorphism is called *honest* if it keeps all full matrices full. In particular, a homomorphism to a field K is honest if it inverts all full matrices; this is possible only when the singular kernel of K consists precisely of the non-full matrices. The set of all full matrices over a given ring R will be denoted by $\Phi = \Phi(R)$, and an R -ring or homomorphism is called *fully inverting* if it is Φ -inverting. Any fully inverting homomorphism $f: R \rightarrow S$ to a non-zero ring S must be injective, for every non-zero element of R is full, as 1×1 matrix, and so maps to an invertible element of S . We shall need a condition for a universal localization to be a field, rather than just a local ring (Th. 2.2). In fact we have two slightly different situations with essentially the same proof.

A set Σ of matrices is said to be *factor-closed* if for any square matrices A, B such that $AB \in \Sigma$, A is invertible over R_Σ .

PROPOSITION 5.7. *Let R be a ring.*

(i) *If the set Φ of full matrices is multiplicative and $R_\Phi \neq 0$, then R_Φ is a field;*

(ii) *if Σ is a factor-closed multiplicative set of matrices over R and $f: R \rightarrow S$ is an honest Σ -inverting homomorphism, so that f can be factored by $f': R_\Sigma \rightarrow S$, then f' is injective.*

Proof. In both cases we have to show that certain elements in a localization are zero; in case (i) they are the non-units in R_Φ and in case (ii) the elements of R_Σ in the kernel of f' .

Let p be an element of R_Φ, R_Σ , respectively and let

$$Au = 0 \quad (3)$$

be an admissible system for p . In case (i) p is a non-unit in R_Φ so its numerator (A_0, A_*) is not invertible over R_Φ and hence is not full over R . In case (ii), $pf' = 0$, hence the numerator (A_0, A_*) is not full over S and by honesty of f , not full over R .

Thus in both cases, $(A_0, A_*) = PQ$, where $P \in {}^n R^{n-1}$, $Q \in {}^{n-1} R^n$ and so

$$A = (A_0 \ A_* \ A_n) = (PQ \ A_n) = (P \ A_n) \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix}.$$

We claim that (P, A_n) is invertible over R_Φ, R_Σ , respectively. For (P, A_n) is a left factor of the denominator (A_*, A_n) and in case (i) the latter is full over R , hence so is (P, A_n) , and so it is invertible over R_Φ . In case (ii) the denominator is in Σ , and since Σ is factor-closed, (P, A_n) becomes invertible over R_Σ . We can therefore in both cases cancel the left factor (P, A_n) in (3) and conclude that $p = 0$. ■

Let us take $\Sigma = \Phi$ in (ii) and note that Φ is necessarily factor-closed; combining (i) and (ii) we obtain one part of

COROLLARY 5.8. *Let $f: R \rightarrow S$ be a fully inverting homomorphism. If either (i) $S \neq 0$ and $\Phi(R)$ is multiplicative, or (ii) S has UGN, then f is injective and R_Φ is a field.*

It only remains to prove (ii). Clearly the set Φ' of matrices inverted by f contains Φ . If $\Phi' \neq \Phi$, then a non-full matrix becomes invertible over S , so a unit-matrix is non-full and S cannot satisfy UGN. ■

We shall use this result to derive a criterion for the universal localization R_Φ to be a field:

THEOREM 5.9. *Let R be any ring and Φ the set of all its full matrices. Then there is an honest map to an epic R -field, which is necessarily the universal localization R_Φ if and only if*

- (i) *the diagonal sum of any full matrices is full and*
- (ii) *the determinantal sum of any non-full matrices (where defined) is non-full.*

For the conclusion requires that the set of non-full matrices should be the unique least prime matrix ideal. If this is to be the case, (i) and (ii) must hold. Conversely, when they are satisfied, then the non-full matrices form a matrix ideal by (ii), necessarily the least, and it is prime by (i) (it is proper because 1, being the diagonal sum of the empty set of full matrices, is full). Hence there is an honest map to an epic R -field K , the residue class field of R_Φ . Now Cor. 5.8(i) shows that $R_\Phi = K$. ■

We remark that (i) generalizes the condition 'no zero-divisors'; (ii) is more subtle and has no analogue in the commutative case (but note that if the

product of full matrices is full and R has UGN, then (ii) holds by Lemma 9.1, below).

For an example of a ring satisfying UGN but not (ii) we use the following construction (cf. Bergman [74]). Let R be a k -algebra with 27 generators, arranged as 3×3 matrices P, U, V with defining relations in matrix form

$$UV = VU = I, \quad P^2 = P, \quad UP = (I - P)U. \quad (4)$$

By Th. 2.11.2, R is an integral domain (in fact a 2-fir), and clearly R can be realized by 2×2 matrices over a field, hence R has UGN; in particular, I is full. But there can be no R -fields, for over an R -field the image of P could be transformed to diagonal form with 0's and 1's on the main diagonal. Since P is conjugate to $I - P$, there must be equal numbers of both, which is impossible, because the order is odd. So there are no R -fields, and by Th. 5.1, I is a determinantal sum of non-full matrices.

We now return to the situation of Th. 5.9 and show that the class of rings considered there can be characterized in different ways:

THEOREM 5.10. *For any ring R the following conditions are equivalent:*

- (a) *the set of all non-full matrices is a prime matrix ideal,*
- (b) *if Φ denotes the set of all full matrices, then R_Φ is a field,*
- (c) *R has an inner rank preserving homomorphism to a field,*
- (d) *R is a Sylvester domain.*

Proof. (a) $\Leftrightarrow$ (b) was proved in Th. 5.9.

(a, b) $\Rightarrow$ (c). When (a), (b) hold, then the homomorphism $\lambda: R \rightarrow R_\Phi$ is honest and R_Φ is a field, hence by Cor. 5.4.10, λ preserves the inner rank.

(c) $\Rightarrow$ (d) follows because any field clearly satisfies Sylvester's law of nullity.

To prove (d) $\Rightarrow$ (a), we use Lemma 5.5.3; part (i) of that lemma gives condition (i) of Th. 5.9. To prove (ii), let $A = (A', a)$, $B = (A', b)$, $C = (A', a + b)$, where $A' \in {}^n R^{n-1}$, $a, b \in {}^n R$ and A, B are not full. Suppose first that A' is right full, then $n - 1 = \rho(A') = \rho(A', a) = \rho(A', b)$, hence $\rho(C) = n - 1$ by Lemma 5.5.3(ii). If A' is not right full, then $\rho(A') < n - 1$, hence $\rho(C) < n$, so in either case C is not full, as we had to show. ■

This result tells us that Sylvester domains form the precise class of rings which have a universal field of fractions over which every full matrix can be inverted. In particular, since any semifir is a Sylvester domain, we have

COROLLARY 5.11. *Every semifir R has a universal field of fractions K , such that every full matrix over R can be inverted over K . ■*

We also note the following consequence of Th. 5.10.

COROLLARY 5.12. *Over any Sylvester domain the product of any two right full matrices (if defined) is right full.*

This follows because a matrix is right full over R if and only if it is right regular over its universal field of fractions. ■

For pseudo-Sylvester domains we have the following analogue. Let us call a homomorphism of rings *stably honest* if it keeps stably full matrices stably full. Then Prop. 5.7(ii) remains true if f is merely stably honest. To see this we note that in Cramer's rule,

$$(A_* - A_0) = (A_* \ A_\infty) \begin{pmatrix} I & u_* \\ 0 & p \end{pmatrix},$$

the left-hand side is now stably full, hence so are the factors on the right-hand side. We also note that the analogue of Th. 5.9 for stably honest maps, with the set Φ^* of stably full matrices instead of Φ holds, with the same proof as before.

THEOREM 5.13. *For any ring R the following are equivalent:*

- (a) *the set of all matrices that are not stably full is a prime matrix ideal,*
- (b) *if Φ^* is the set of all stably full matrices, then R_{Φ^*} is a field,*
- (c) *R has a stable rank preserving homomorphism to a field K ,*
- (d) *R is an S -ring.*

Moreover, the field K is a universal R -field; it is a field of fractions if and only if R is weakly finite.

Proof. (a) $\Leftrightarrow$ (b) follows from the analogue of Th. 5.9. (a, b) $\Rightarrow$ (c) follows by Lemma 5.6.5 instead of Th. 5.4.9. (c) $\Rightarrow$ (d) follows as before, and (d) $\Rightarrow$ (a) follows by Lemma 5.6.6 instead of Lemma 5.5.3.

Now K is a universal R -field because the matrices that are not stably full clearly form the least prime matrix ideal. If R is weakly finite, then $\rho^*a = 0$ implies $a = 0$, by Prop. 5.6.1, and conversely, any subring of a field is weakly finite. ■

Finally, we examine the behaviour of homomorphisms under localization.

THEOREM 5.14. *Let R, R' be rings with prime matrix ideals $\mathcal{P}, \mathcal{P}'$ and corresponding epic fields K, K' , respectively. Then*

- (i) a homomorphism $f: R \rightarrow R'$ extends to a specialization of R -fields $K \rightarrow K'$ if and only if f maps $\mathcal{P}$ into $\mathcal{P}'$ and
 (ii) a homomorphism $f: R \rightarrow R'$ extends to a field homomorphism $K \rightarrow K'$ if and only if f maps $\mathcal{P}$ to $\mathcal{P}'$ and $\mathfrak{M}(R) \setminus \mathcal{P}$ to $\mathfrak{M}(R') \setminus \mathcal{P}'$.

Proof. The set $\mathcal{P}'f^{-1} = \{A \in \mathfrak{M}(R) \mid Af \in \mathcal{P}'\}$ is a prime matrix ideal of R and by hypothesis $\mathcal{P} \subseteq \mathcal{P}'f^{-1}$, which is the criterion for a specialization, by Th. 2.4. Now denote the complements of $\mathcal{P}, \mathcal{P}'$ by Σ, Σ' ; if $\Sigma f \subseteq \Sigma'$, then the denominators of elements of K remain invertible over K' , so we have a homomorphism, and again the condition is clearly necessary. ■

Exercises 7.5

1. Let R be a ring with a fully inverting R -field K . Show that every honest endomorphism of R extends to a unique endomorphism of K . Show that every derivation on R extends to a unique derivation on K .

2. Show that a ring R has a universal R -field if and only if the radical of the least matrix ideal is prime.

3. Show that R is embeddable in a direct product of fields if and only if no non-zero scalar matrix can be written as a determinantal sum of non-full matrices.

4. A ring R is called *strongly regular* if for each $a \in R$ there exists $x \in R$ such that $a^2x = a$. Show that such a ring is *reduced*, i.e. has no non-zero nilpotent elements. Deduce that a strongly regular ring which is also prime is a field. (Hint: Observe that in a reduced ring, $xy = 0 \Rightarrow yx = 0$.) Let R be strongly regular and $0 \neq a \in R$; show that any ideal $\mathfrak{p}$ maximal subject to $a \notin \mathfrak{p}$ is prime, and deduce that R is a subdirect product of fields.

5. Show that a subring of a strongly regular ring is embeddable in a field if and only if it is an integral domain. (Hint: For the sufficiency take $R \subseteq P = \prod K_\lambda$ (K_λ a field) by Ex. 4, with projections $\varepsilon_\lambda: P \rightarrow K_\lambda$ and for each $x \in P$ define $\Gamma_x = \{\lambda \mid x\varepsilon_\lambda = 0\}$, $I_x = \{K_\lambda \mid \lambda \in \Gamma_x\}$. Verify that I_x is an ideal in P , and if I is the ideal generated by all I_x with $x \in R^\times$, show that $R \cap I = 0$. Deduce that $P \rightarrow P/I$ is $R^\times$ -inverting, hence any homomorphism of P/I into a field K provides an embedding of R in K .)

6. A condition of the form $A_1 \wedge \dots \wedge A_r \Rightarrow A$, where A_i, A are atomic formulae or A is F , = false, is called a *quasi-identity*. The class of algebras defined by a set of quasi-identities is called a *quasi-variety*; it may be characterized as a universal class admitting direct products (cf. Cohn [81], p. 235). Verify that the following are quasi-varieties: (i) the class $\mathcal{G}$ of monoids embeddable in groups and (ii) the class $\mathcal{T}$ of subrings of strongly regular rings. If $\mathcal{D}_0$ is the class of integral domains, show that $\mathcal{T} \cap \mathcal{D}_0$ is the class of rings embeddable in fields.

7. Show that any filtered ring with inverse weak algorithm has a universal field of

fractions. If K is the universal field of fractions of the free power series ring $k\langle\langle X \rangle\rangle$, show that the subfield of K generated by $k\langle X \rangle$ is the universal field of fractions of the latter. (Hint: Use the inertia theorem, Th. 2.9.15.)

8. Let $R = k\langle x, y, z, t \rangle$ and define an endomorphism α of R by the rules: $x \mapsto xz$, $y \mapsto xt$, $z \mapsto yz$, $t \mapsto yt$. Show that α is injective, but not honest.

9. Let $R = k\langle x, y, z \rangle$, denote its universal field of fractions by K and the subalgebra generated by x, y, z and $t = zx^{-1}y$ by S . Show that the matrix $\begin{pmatrix} x & y \\ z & t \end{pmatrix}$ is full over S but not invertible over K .

10. Let R be a ring with a 2×2 matrix C satisfying $C^n = 0$, for some $n \geq 1$. If c is any entry of C^2 , show that for some unit matrix I , cI is a determinantal sum of non-full matrices. Find an explicit expression for cI when R is generated over k by the entries of C and $C^3 = 0$.

11. Let F be the free group on a set X ; show that the natural mapping $k\langle X \rangle \rightarrow kF$ is honest.

12. Let R be a semifir with centre a field k and with universal field of fractions K . If E is a finite commutative extension of k such that $R \otimes_k E$ is again a semifir, show that $K \otimes E$ is a field and is the universal field of fractions of $R \otimes E$.

13. Let K be the universal field of fractions of the free algebra $k\langle x, y \rangle$. Show that there is a 2×2 idempotent matrix over K which is different from $0, I$ and whose trace is not 1. (Hint: Try $a = x$, $b = y$, $c = y^{-1}(x - x^2)$, $d = 1 - y^{-1}xy$.) If an involution is defined on K by $x^* = x$, $y^* = y^{-1}(x - x^2)$, find a self-adjoint idempotent 2×2 matrix $\neq 0, I$ with trace $\neq 1$.

14*. (A. H. Schofield) Let R be any ring with an epic R -field K . Verify that $K(t)$ is an epic $R[t]$ -field, and deduce that the embedding $f: R \rightarrow R[t]$ induces the surjection $f^*: X(R[t]) \rightarrow X(R)$, where $X(R)$ denotes the set of all epic R -fields or equivalently, prime matrix ideals of R . Show that the inverse image of K under f^* consists of the different $K[t]$ -fields, and deduce that if R has a universal R -field U , then $U(t)$ is a universal $R[t]$ -field.

15. (Bergman and Dicks [78]) If R is a ring with a universal field of fractions K and Σ is the set of all matrices inverted over K , then there is a surjective homomorphism $R_\Sigma \rightarrow K$. Find an example where this is not an isomorphism.

16. Given $n > 1$, find a criterion for a ring to be embeddable in an $n \times n$ matrix ring over a field.

17. Let A be a square matrix over a free algebra. If the matrix ideal generated by A is improper, is A necessarily invertible?

18. (Bedoya and Lewin [77]) Let R be a ring and K a field of fractions of R .

Show that if every full matrix over $R[x]$ is invertible over $K(x)$, then R is a Sylvester domain and K its universal field of fractions.

19. (Cohn [71]) Let D be a field with two subfields A, B that are isomorphic, via an isomorphism $f: A \rightarrow B$. Show, by taking the universal field of fractions of a suitable HNN-construction (cf. 2.6) that D can be embedded in a field in which f is realized as an inner automorphism. (This process allows two elements transcendental over k to be made conjugate in an extension field. It can be shown to apply more generally to square matrices, cf. Cohn [73, 77].)

20. Let G be a group whose group algebra kG is embeddable in a field D , and let F be the free group on X . By considering $D\langle X, X^{-1} \rangle$ show that the group algebra of the direct product $G \times F$ can be embedded in a field.

7.6 Numerators and denominators

Let R be a ring and K an epic R -field; our object will be to compare different admissible systems for the same element or matrix over K . For simplicity we shall only discuss elements of K ; the only other case needed, that of a square matrix, is then easily deduced. However, we shall allow the admissible matrix A to have index other than 1, thus A will be $r \times (m+1)$, say. Then the system $Au = 0$ is said to have *order* m and it is admissible for p if it has a unique normalized solution u (i.e. $u_0 = 1$) and $u_\infty = p$. If $\mathcal{P}$ is the singular kernel of the R -field K , we shall also call A $\mathcal{P}$ -admissible, or K -admissible, but the reference to $\mathcal{P}$ or K is omitted when there is no risk of confusion.

In order to compare different systems we shall describe some operations which can be carried out on a system without changing its solution. Let A be an admissible matrix for p (over K); if u satisfies $Au = 0$, then it also satisfies $PAu = 0$ for any matrix P for which the product is defined, but the solutions of $Au = 0$ and $AQu = 0$ are not generally the same. However, we are only concerned with the ratio of the first and last components of u , and this ratio is preserved when Q is suitably restricted. Let us define a *bordered* matrix as in 7.4, as a matrix Q^* (not necessarily square), which agrees with the unit matrix in the first and last row. We shall define the following two operations on an admissible matrix A for p .

λ . Replace A by PA such that PA is again admissible for p .

ρ . Replace A by AQ^* , where Q^* is bordered, such that AQ^* is again admissible for p .

The inverse operations, cancelling a left or a bordered right factor from A so as to obtain again an admissible matrix for p , are denoted by λ', ρ' ,

respectively, and an operation is called *trivial* or *full* if the matrix P or Q^* is invertible or full respectively. We shall say that A *admits* a certain operation if the result of performing it on A is again admissible.

Just as matrices may be thought of as defining certain finitely presented modules, so admissible matrices, essentially matrices with distinguished first and last column (and a rank condition) define left modules with a distinguished pair of elements. We shall need a notion of equivalence for these matrices corresponding to stable association but preserving these two columns.

Two matrices A, A' are said to be *biassociated* if there exist invertible matrices U, V such that

$$AU = VA' \quad (1)$$

and U is bordered. The matrices $A = (A_0, A_*, A_\infty)$, $A' = (A'_0, A'_*, A'_\infty)$ are *stably biassociated* if they are stably associated, with a bordered matrix in which the unit rows correspond to the first and last columns of A . Over a weakly finite ring this condition can again be restated in terms of comaximal relations.

PROPOSITION 6.1. *Let R be a weakly finite ring and let $A \in {}^rR^{m+1}$, $A' \in {}^sR^{n+1}$, where $r+n = s+m$. Then A, A' are stably biassociated if and only if there is a comaximal relation*

$$AB' = BA' \quad (2)$$

with a bordered matrix B' .

Proof. If A, A' are stably biassociated, then by Prop. 0.6.5, we have a relation

$$\begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} U = V \begin{pmatrix} I & 0 \\ 0 & A' \end{pmatrix}, \quad (3)$$

where U, V are invertible. Moreover, U is obtained from a bordered matrix by certain row and column permutations and (by the formula after Prop. 0.6.5) has the partitioned form

$$U = \begin{pmatrix} D' & B' \\ C' & A' \end{pmatrix},$$

where B' is bordered and D' has the first and last row zero. If we partition V correspondingly and call the $(1, 2)$ -block B , then a comparison of $(1, 2)$ -blocks in (3) gives the comaximal relation (2) with bordered matrix B' . Conversely, when (2) is given as a comaximal relation, then by Prop. 0.6.5, we obtain two mutually inverse matrices

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad M^{-1} = \begin{pmatrix} D' & B' \\ C' & A' \end{pmatrix}. \quad (4)$$

Since the first row of B' is $e_1 = (1, 0, \dots, 0)$, we can subtract multiples of the $(r+1)$ th column of M^{-1} from the first r columns so as to reduce the first row of D' to 0, and make the corresponding change in M . This will only affect C and D , but not A or B , in M . Similarly, since the last row of B' is e_{m+1} we can reduce the last row of D' to zero (at the expense of further changes to C and D). Since the relation $M^{-1}M = I$ has been preserved, we see that the first and last rows of D are zero, while C is bordered. Now the argument of Prop. 0.6.5 shows that A and A' are stably biassociated. ■

From the proof we see that when A, A' are stably biassociated, they occur in pairs of inverse matrices, as in (4), with B' bordered. If in the expression (4) for M^{-1} we consider the cofactor of the $(1, 1)$ -entry of B' , we again obtain a pair of mutually inverse matrices, but with the first column of A, A' omitted. This shows that the denominators of A, A' are stably associated. A similar argument (omitting the last row and column of B' in (4)) shows that the numerators of A, A' are stably associated, as are the cores, and of course, here weak finiteness is not needed. Thus we obtain

PROPOSITION 6.2. *Over any ring R , if A, A' are stably biassociated, then the denominators (omitting the first column), the numerators (omitting the last column) and the cores (omitting the first and last columns) are stably associated. ■*

Let us next record conditions under which the operations $\lambda, \lambda', \rho, \rho'$ preserve admissibility.

PROPOSITION 6.3. *Let R be a ring with epic R -field K .*

(i) *The matrix $A \in {}^rR^{m+1}$ is admissible over K if and only if (A_*, A_m) is right regular over K but A is not; in particular this entails that $r \geq m$.*

(ii) *Suppose that $B = PA$, where $A \in {}^rR^{m+1}$, $B \in {}^sR^{m+1}$, $P \in {}^sR^r$. If A is admissible for $p \in K$, then B is admissible for p if and only if $\hat{P}(A_*, A_m)$ is right regular over K ; this is so if for example P is right regular over K . If B is admissible for p , then A is admissible for p if and only if A is not right regular over K .*

(iii) *Suppose that $B = AQ^*$, where $A \in {}^rR^{m+1}$, $B \in {}^rR^{n+1}$ and $Q^* \in {}^{m+1}R^{n+1}$ is bordered. If A is admissible for p , then B is admissible for p if and only if Q^* is right regular over K but B is not; when this is so, $m \geq n$. If B is admissible for*

p , then A is admissible for p if and only if its denominator (A_*, A_m) is right regular over K .

(iv) For a matrix P over R which is right regular over K , A is admissible if and only if PA is, and for a bordered matrix Q^* over R and right regular over K , A is admissible if and only if AQ^* is.

Proof. We have already seen (i), and (ii) follows because A and PA are admissible for the same element p (by uniqueness) if they are admissible at all, while admissibility follows because the rank (over K) is non-increasing under multiplication. Similarly (iii) and (iv) follow because the product of right regular matrices is right regular over K . ■

We remark that when R is a Sylvester domain and K its universal field of fractions, regularity over K can be replaced by fullness over R in this proposition.

We next examine conditions under which an admissible system can be simplified by operations $\lambda, \lambda', \rho, \rho'$. For any matrix A over R and any R -field K , the rank of A over K will be called the K -rank of A , written $\rho_K(A)$.

LEMMA 6.4. Let R be a ring with epic R -field K , and let $A \in {}^rR^{m+1}$ be a K -admissible matrix. Then the index of A is at most 1, i.e. $m \leq r$ and the following are equivalent:

- (a) the index is exactly 1, i.e. $m = r$,
- (b) A is left regular over K ,
- (c) the number of rows of A cannot be decreased by an operation λ .

Further, any K -admissible matrix can be transformed by an operation λ to a K -admissible matrix which is left regular over K .

If R is a Sylvester domain and K its universal field of fractions, then (a)–(c) are equivalent to

- (d) the number of rows of A cannot be decreased by an operation λ' ,

and, further, any K -admissible matrix can be transformed by an operation λ' to a K -admissible matrix which is left full, and every operation λ or λ' admitted by a left full matrix which does not increase the number of rows is full.

Proof. Since $Au = 0$ has a unique normalized solution u over K , the K -rank of A is m , whence $r \geq m$, with equality if and only if $\rho_K(A) = r$, i.e. A is left regular over K ; thus (a) $\Leftrightarrow$ (b).

Next let us write

$$A'' = EA, \quad E \in {}^mR^r, \quad A'' \in {}^mR^{m+1}, \quad (5)$$

where E consists of m rows of the $r \times r$ unit matrix, chosen so that the denominator formed from A'' is regular over K . Then A'' is K -admissible; this shows that (a) $\Leftrightarrow$ (c), and that we can reach a matrix satisfying (a) by an operation λ .

For the final part all full matrices become invertible over K and ρ_K is the inner rank. Since $\rho A = m$, we can write

$$A = PA', \quad \text{where } P \in {}^rR^m, A' \in {}^mR^{m+1},$$

where P is right full and A' is left full. Clearly A' is not right full, so by Prop. 6.3(ii), A' is admissible; moreover P is full precisely when $r = m$. This shows that (a) $\Leftrightarrow$ (d) and it also proves the last assertion. ■

We next examine when an admissible matrix is left prime; here we need to assume that the homomorphism to the R -field is fully inverting. In fact, the next two lemmas hold for Sylvester domains (except for the last sentence in each), but they will only be needed for semifirs.

LEMMA 6.5. *Let R be a semifir with universal field of fractions K , and let $A \in {}^rR^{m+1}$ be an admissible matrix. Then the following conditions are equivalent:*

- (a) A is left prime (hence $r = m$),
- (b) every non-trivial operation λ' admitted by A increases the number of rows.

Further, if R has left ACC_m then any admissible matrix $A \in {}^rR^{m+1}$ can be transformed by an operation $\lambda': A = PA'$ to a left prime admissible matrix $A' \in {}^mR^{m+1}$.

Proof. Suppose that A is left prime; then A is left full and no operation λ' can decrease the number of rows. Moreover, in any factorization of A ,

$$A = PB, \quad \text{where } P \in R_r, B \in {}^rR^{m+1}, \quad (6)$$

P is a unit, so any non-trivial operation λ' increases the number of rows, i.e. (b) holds. Conversely, if (b) holds, then by Lemma 6.4, A is left full and in any equation (6), P is a unit and hence A is left prime, i.e. (a) holds.

If R has left ACC_m we can apply Prop. 5.10.1 to obtain an equation (6), where B is left prime. ■

To establish a corresponding result for the operations ρ, ρ' , let us call an admissible matrix A *right *-prime* if its core A_* is right prime.

LEMMA 6.6. *Let R be a semifir with universal field of fractions K , and let A, A' be admissible matrices of orders m, m' , respectively. If $A = A'Q^*$, where Q^* is*

bordered, then $m \leq m'$, with equality if and only if Q^* is full. For any admissible matrix A the following are equivalent:

- (a) A is right $*$ -prime,
- (b) every bordered square right factor of A is a unit,
- (c) all full operations ρ' admitted by A are trivial.

Further, if R has right ACC_{m-1} , then any admissible matrix $A \in {}^R R^{m+1}$ can be reduced by a full operation $\rho' : A = A'Q^*$ to one which is right $*$ -prime.

Proof. If A, A' are admissible, where for a bordered Q^* ,

$$A = A'Q^*, \quad A \in {}^R R^{m+1}, A' \in {}^R R^{m'+1}, Q^* \in {}^{m'+1} R^{m+1}, \quad (7)$$

then Q^* is right full, by Prop. 6.3(iii), hence $m \leq m'$, with equality if and only if Q^* is full.

Now assume that $m' = m$ in (7); comparing cores in (7), we find that $A_* = A'_*Q_*$, and clearly Q^* is a unit if and only if Q_* is. It follows that A_* is right prime if and only if Q^* is a unit in every factorization (7). Thus (a) $\Leftrightarrow$ (b), and (b) $\Leftrightarrow$ (c) is clear because any full operation ρ' preserves the order and so has a square matrix Q^* . Conversely, when Q^* is square, ρ' preserves the order and so is full.

Finally, when R has right ACC_{m-1} , we can by Prop. 5.10.1 write

$$A_* = A'_*Q_*, \quad \text{where } A'_* \in {}^R R^{m-1}, Q_* \in R_{m-1},$$

where A'_* is right prime and Q_* is full. Hence on writing $A = A'Q^*$, where

$$A' = (A_0 \quad A'_* \quad A_\infty), \quad Q^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & Q_* & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

we find that Q^* is full, whence A' is admissible and right $*$ -prime. ■

From now on we shall restrict attention to the case of a semifir R with universal field of fractions U . We begin with a basic result giving conditions under which two admissible systems determine the same element of U . Let us call an operation ρ *strict* if its matrix is of the form $Q^* = 1 \oplus Q_* \oplus 1$; such a matrix is called *strictly bordered*. We note that every operation ρ consists of a trivial operation followed by a strict operation.

LEMMA 6.7. *Let R be a semifir and U its universal field of fractions. Then two admissible matrices $A \in {}^R R^{m+1}$ and $B \in {}^S R^{n+1}$ determine the same element $p \in U$ if and only if there exist matrices P, P' , strictly bordered matrices Q^*, Q'^* and admissible matrices A', B' such that*

$$A = PA', \quad B = B'Q^*, \quad (8)$$

$$P'A' = B'Q'^*. \quad (9)$$

Moreover, we can always take P to be right full, Q^* to be full and A' to be left full.

Proof. Assume that (8), (9) hold; by (8), A and A' determine the same element p of U , and B, B' determine the same element q . If $A'u = 0$, we apply (9); Q'^*u has the form $(1, v_*, p)^T$, so if $p \neq q$, then $B'Q'^*u \neq 0$ and we have a contradiction. Thus $p = q$.

Conversely, assume that A and B determine the same element p . It follows that

$$\begin{pmatrix} A_0 & A_* & A_\infty & 0 & A_\infty \\ B_0 & 0 & B_\infty & B_* & 0 \end{pmatrix}$$

is an admissible system for $p - p = 0$, so its numerator is not full:

$$\begin{pmatrix} A_0 & A_* & A_\infty & 0 \\ B_0 & 0 & B_\infty & B_* \end{pmatrix}$$

has inner rank $k < m + n$. We apply the partition lemma to a minimal factorization:

$$\begin{pmatrix} A_0 & A_* & A_\infty & 0 \\ B_0 & 0 & B_\infty & B_* \end{pmatrix} = \begin{pmatrix} P & 0 \\ P' & B'_* \end{pmatrix} \begin{pmatrix} A' & 0 \\ -Q'_* & Q_* \end{pmatrix}. \quad (10)$$

On equating blocks we find

$$A = (A_0 \ A_* \ A_\infty) = PA', \quad (11)$$

$$(B_0 \ 0 \ B_\infty) = P'A' - B'_*Q'_*, \quad (12)$$

$$B_* = B'_*Q_*. \quad (13)$$

Here (11) is the first equation of (8), and (13) yields the second equation of (8) if we set

$$B' = (B_0 \ B'_* \ B_\infty), \quad Q^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & Q_* & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Finally, if we put

$$Q'^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & Q'_* & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

then (12) takes on the form (9).

It remains to prove that P is right full, A' is left full and Q^* is full. This will also show (by Prop. 6.3) that A', B' are admissible.

We denote by t the number of columns of P, P' or, equivalently, the number of rows of A' (cf. (10)), and by t' the number of columns in B'_* , equivalently, the number of rows in Q'_*, Q_* . Since (10) was a minimal factorization, we have

$$t + t' = k < m + n. \quad (14)$$

But $t \geq \rho P \geq \rho A = m$, by (11) and $t' \geq \rho Q_* \geq \rho B_* = n - 1$, by (13). A comparison with (14) shows that $t = \rho P = m$ and $t' = \rho Q_* = n - 1$. Hence P is right full, A' has index 1 and so is left full, by Lemma 6.4, and Q_* , therefore also Q^* , is full, as claimed. ■

Note that although we proved P to be right full and Q^* full, we have not proved this for P' and Q'^* . Hence the common value of the two sides of (9) need not itself be an admissible matrix, and we have not obtained a chain of operations $\lambda, \lambda', \rho, \rho'$ connecting A and B .

To study the case where $Q'^* = 1 \oplus Q'_* \oplus 1$ may be non-full, take a minimal factorization $Q'_* = R_* S_*$ and write $R^* = 1 \oplus R_* \oplus 1, S^* = 1 \oplus S_* \oplus 1$; then

$$Q'^* = R^* S^*, \quad (15)$$

where R^* is strictly bordered and right full and S^* is strictly bordered and left full. By the first part of the above proof, $A'u = 0$; hence on combining this with (9) and (15) we get

$$0 = P'A'u = B'Q'^*u = (B'R^*)(S^*u).$$

Since S^* is bordered, the vector S^*u has first entry 1, therefore $B'R^*$ is not right full. But by Prop. 6.3(iii), $B'R^*$ is admissible, because R^* is right full. Thus $B' \mapsto B'R^*$ is an operation ρ taking B' to an admissible matrix, and this does not lead to a 'path' from A to B unless S^* is full. However, we can ensure that S^* is full by imposing a further condition on A and B .

Given admissible matrices A and B , B is said to be *minimal over A* if B cannot be transformed by any full operation ρ' , followed by any strict operation ρ , to an admissible matrix of smaller order than A . If B is minimal over itself, it will be called *minimal*.

Let $B'R^*$ be of order h ; if B is minimal over A , then since $B' \mapsto B'$ by a full operation ρ' and $B' \mapsto B'R^*$ by a strict operation ρ , it follows that $h = \text{order } B'R^* \geq \text{order } A = m$. Now S^* is $(h+1) \times (m+1)$ and is left full, so $h \leq m$, and hence $h = m$. It follows that $Q'^* = R^* S^*$ is right full, and so $P'A' = B'Q'^*$ is admissible, and by Prop. 6.3(iii), P' is right full. This proves

COROLLARY 6.8. *In the situation of Lemma 6.7, if B is minimal over A , then the common value of the two sides in (9) is an admissible matrix. Hence we can pass*

from A to B by a chain of four operations $\lambda', \lambda, \rho', \rho$, in that order. Here ρ is full, and if A is left full, then λ' is full. ■

In Lemma 6.7, if A is left prime, P must be a unit, and if B is right $*$ -prime, Q^* must be a unit. Let us call an admissible matrix A *reduced* if it is left prime and right $*$ -prime. Then we can state the result just proved as

COROLLARY 6.9. *In the situation of Lemma 6.7, if A is left prime and B right $*$ -prime, then there exist a matrix P and a bordered matrix Q^* such that $PA = BQ^*$. If, moreover, A, B are reduced, so that they both have index 1, then they are stably biassociated.*

Here the last part follows because the relation $PA = BQ^*$ is necessarily coprime, and hence comaximal, by Cor. 3.2.5. ■

We remark that if A is right $*$ -prime, any full operation ρ' must be trivial; hence a reduced matrix A which is not minimal admits a strict operation $\rho: A \mapsto AQ^*$ with a matrix Q^* of negative index.

We now come to our main result, the comparison theorem for numerators and denominators.

THEOREM 6.10. *Let R be a semifir, U its universal field of fractions and p an element of U .*

- (i) *We can pass between any two admissible matrices A, B for p by some sequence of operations $\lambda, \lambda', \rho, \rho'$ (at most six in all).*
- (ii) *The minimal admissible matrices for p are just the admissible matrices of least order for p , and we can pass from any one minimal admissible matrix for p to any other by four full operations, viz. $\lambda', \lambda, \rho', \rho$, in that order.*
- (iii) *If A and B are left full admissible matrices for p , then we can pass by full operations λ', ρ' from A to an admissible matrix A'' and from B to an admissible matrix B'' such that A'' and B'' are stably biassociated.*
- (iv) *Any two reduced admissible matrices A, B for p are stably biassociated, and if there exists a reduced matrix for p , then there exists a reduced minimal matrix for p .*

Proof. We begin with (ii); by Lemma 6.6 and an induction on the order we see that every element has a minimal admissible matrix. If A, B are minimal for p , with order $A \leq \text{order } B$, say, then B is certainly minimal over A , so by Cor. 6.8, the two sides of (9) represent an admissible matrix and we have a path from A to B by our four operations. By Lemma 6.7, P is right full, and since A is left full, P must be full; by the lemma, Q^* is also full. Since B is

minimal, Q^* must be full and since A' and $P'A' = B'Q^*$ are left full, P' must be full, by Lemma 6.4. So the four operations used are full; in particular, this implies that A, B have the same order, and this proves the last part of (ii).

Since all minimal admissible matrices for p have the same order, and since from any admissible matrix A for p we can construct a minimal admissible matrix by repeated operations ρ', ρ, λ' that do not increase the order, it follows that the common order of all minimal admissible matrices for p is the least order of any admissible matrix for p ; this proves the rest of (ii).

To prove (i) let B be any admissible matrix and A a minimal admissible matrix for p ; then B will certainly be minimal over A and we can pass from B to A by four operations $\rho', \rho, \lambda', \lambda$. Hence, using any minimal matrix as connecting link, we can pass between any two admissible matrices by a sequence of at most eight operations: $\rho', \rho, \lambda', \lambda, \lambda', \lambda, \rho', \rho$. By allowing the choice of the 'connecting link' to depend on one of the given matrices, we can in fact eliminate the middle two operations and obtain a chain of six; this proves (i).

Next let A and B be left full admissible matrices for p , as in (iii). By Lemma 6.7, (8), we can transform A by a full operation λ' , and B by a full operation ρ' to matrices A', B' such that

$$P'A' = B'Q^*, \quad \text{where } Q^* \text{ is strictly bordered.} \quad (16)$$

Since B is left full and Q^* is full, B' is left full, hence the matrix (P', B') is also left full. To show that $(A', Q^*)^T$ is right full, suppose that $(A', Q^*)^T x = 0$ for some $x \in {}^m U$. Since A' is admissible for p , the vector x is a right multiple of a vector $u = (1, u_*, p)^T$, so if $x \neq 0$, its first entry must be non-zero. But $Q^*x = 0$, so by the bordered form of Q^* , the first entry of x is 0. This shows that $x = 0$, so $(A', Q^*)^T$ is right full. Further, A' and B' , being left full admissible matrices, both have index 1, by Lemma 6.4, so by Prop. 3.2.4, we can cancel full left and right factors from the two sides of (16) so as to obtain a comaximal relation. Cancelling a full square left factor from B' clearly corresponds to an admissible operation λ' . Now any full right factor of Q^* can be taken to be bordered. For by the partition lemma, we have

$$Q^* = \begin{pmatrix} r & 0 \\ R' & R'' \end{pmatrix} \begin{pmatrix} s & 0 \\ S' & S'' \end{pmatrix}, \quad (17)$$

where $rs = 1$, hence r is a unit and we can by a trivial modification reduce r, s to 1 in (17). A similar argument applies to the last row. Now we can cancel the bordered full right factor, and this corresponds to an operation ρ' . The resulting comaximal relation

$$P''A'' = B''Q''^*$$

shows that A'' and B'' are stably biassociated.

Now the first statement of (iv) follows from (iii) by the definition of 'reduced'. For the second assertion we apply the reduction in (iii) to a minimal matrix A and a reduced matrix B . Of the resulting matrices A'' and B'' , the first will again be minimal, because the operations applied to A were full, while the second can be taken equal to B , because B admits no non-trivial operations λ' or ρ' . Hence A'' , being stably biassociated to the reduced matrix B , will be a minimal reduced admissible matrix for p . ■

In 7.1 we defined the left depth for p as the least value of the order of any admissible matrix for p . From Th. 6.10(ii) we see that it can be obtained as the order of any minimal admissible matrix for p . We note that whereas a minimal admissible matrix for a given $p \in U$ always exists, there may be no reduced matrix (in the absence of ACC). Moreover, a reduced matrix for p need not be minimal; e.g. over $k\langle x, y \rangle$ the matrix $\begin{pmatrix} x & -1 & 0 \\ 0 & 1 & -y \end{pmatrix}$ is admissible for $y^{-1}x$ and is reduced but not minimal. However, when there is a reduced matrix, then there is also a minimal reduced matrix, by Th. 6.10(iv).

We remark that analogues of these theorems hold when p is a matrix over U . The simplest case is that of a square matrix, say $t \times t$. This is most easily handled by taking T to be a semifir with universal field of fractions V and writing $R = T_t$, $U = V_t$. An $m \times n$ matrix over R is now an $mt \times nt$ matrix over T ; if its inner rank over T is r , then the rank of A over R may be defined as r/t . So the rank is now no longer an integer but a rational number with denominator dividing t . By a *left full* matrix we understand as before an $m \times n$ matrix of rank m ; clearly this will also be left full as matrix over T . Right full and full matrices are defined correspondingly. The definitions of left and right prime, maximal and reduced, and of the basic operations $\lambda, \rho, \lambda', \rho'$ are all as before. An admissible system for $p \in U$ can again be defined as a system with right full denominator, of rank equal to the rank of the matrix. If this rank is m and the numerator has rank s , then the rank of p over U is $s - m + 1$, a rational number between 0 and 1.

With these conventions Th. 6.10 and the lemmas leading up to it hold as stated, but for matrix rings over semifirs. The proofs are similar to those given here and will be left to the reader.

Exercises 7.6

1. Show that the conditions of Lemmas 6.4–6 are invariant under stable biassociation; interpret these conditions as conditions on the corresponding module with a pair of distinguished elements.

2. Show that all full (but no non-full) operations $\lambda, \lambda', \rho, \rho'$ preserve the conditions of Lemma 6.4, all full operations ρ' but not all full operations ρ preserve the condi-

tions of Lemma 6.5, all full operations λ' but not all full operations λ preserve the conditions of Lemma 6.6.

3. A matrix is called *strictly bordered* if it is of the form $1 \oplus Q \oplus 1$. Verify that any full operation ρ can be accomplished by carrying out a trivial operation ρ , followed by an operation ρ with strictly bordered matrix.

4. Let A be an admissible matrix; if its core has an invertible $t \times t$ submatrix, show that the order of A can be decreased by t .

5. Let U be the universal field of fractions of the free algebra $k\langle x, y, z \rangle$. Show that $A = \begin{pmatrix} x & -z & 0 \\ 0 & z & -y \end{pmatrix}$ is an admissible matrix for $y^{-1}zz^{-1}x = y^{-1}x$. Verify that no operation ρ will decrease the order of A , but that $y^{-1}x$ has left depth 1.

6. State and prove the matrix analogues of the results in this section.

7.7 The depth

In the depth we have a numerical invariant (for the elements of an epic R -field) which has no analogue in the commutative case; there we can write every element as $ab^{-1} = b^{-1}a$ and this has depth 1. By contrast we shall see that in the universal field of fractions of a semifir the depth is generally unbounded. But first we shall examine the case of elements of minimum depth. Given a ring R with epic R -field K , if a matrix $P \in {}^m K^n$ is given by an admissible system with null core (i.e. a core with 0 columns), then the system has the form

$$(A \ B) \begin{pmatrix} I \\ P \end{pmatrix} = 0, \quad A \in {}^m R^n, B \in R_m, \quad (1)$$

where B is invertible over K , and hence

$$P = -B^{-1}A, \quad B \text{ invertible over } K. \quad (2)$$

Conversely, any matrix P of the form (2) is given by the admissible system (1). Similarly, P has the form

$$P = -AB^{-1}, \quad B \text{ invertible over } K, \quad (3)$$

precisely if it is given by a right admissible system $(I, P) \begin{pmatrix} A \\ B \end{pmatrix} = 0$, or equivalently, by Prop. 1.5, by

$$\begin{pmatrix} -I & B & 0 \\ 0 & A & I \end{pmatrix} \begin{pmatrix} I \\ B^{-1} \\ P \end{pmatrix} = 0. \quad (4)$$

We remark that the matrix in (4) has the right annihilator $(B, I, -A)^T$ in R which is right regular over K .

PROPOSITION 7.1. *Let R be a ring and K an epic R -field. Given $P \in {}^m K^n$,*

- (i) *P can be written in the form (2) if and only if it has a core with left annihilator in R which is regular over K , or, equivalently, a null core;*
- (ii) *P can be written in the form (3) if and only if it has an admissible matrix with a right annihilator in R which is right regular over K .*

Proof. (i) If P is given by (2), it has the admissible system (1) with null core, and the latter has left annihilator I . Conversely, if A is an admissible matrix for P , where $CA_* = 0$ for a matrix C over R which is regular over K , then CA is again admissible for P , by Prop. 6.3(iv), and since $CA_* = 0$, the matrix (CA_0, CA_∞) with null core is again admissible for P .

(ii) If P has the form (3), it has the admissible system (4) with right annihilator $(B, I, -A)^T$ which is right regular over K . Conversely, let P have the admissible matrix A with right annihilator C , right regular over K . By the uniqueness of the normalized solution U of (4) over K , we have $C = UC_0$, where $C = (C_0, C_*, C_\infty)^T$; since C is right regular over K , C_0 must be regular over K and so $P = U_\infty = C_\infty C_0^{-1}$. ■

Let us restate the most important special case, where K is a fully inverting field. For simplicity we limit ourselves to elements. In that case (2), (3) just mean that the left, right depth respectively of p is 1.

COROLLARY 7.2. *Let R be a ring with fully inverting field K . Given $p \in K$, $d_l(p) = 1$ if and only if the core of one (and hence every) admissible matrix for p is a zero-divisor in R , and $d_r(p) = 1$ if and only if one (and hence every) admissible matrix for p is a zero-divisor.*

Further, if R is a semifir, then it is left Bezout if and only if the left depth is bounded by 1, and a similar condition ensures that R is right Bezout.

Here the last part follows because a semifir is left Bezout if and only if it is left Ore, or equivalently, every left regular square matrix is invertible over the field of fractions. ■

Let R be a semifir and U its universal field of fractions; then for any matrix A over R its rank over U , $\rho_U A$, equals ρA , as we have seen in 7.5. Now let K be any epic R -field; then the rank of A over K , $\rho_K A$, say, is the maximum of the orders of square submatrices of A which are regular over K , i.e. which do not lie in the singular kernel. Since this kernel is least for the universal field

of fractions, we have

$$\rho_K A \leq \rho_U A. \quad (5)$$

We shall use this observation to show that for fully atomic semifirs the depth cannot increase on specialization. We begin by proving a theorem on 'universal denominators', which is of independent interest.

THEOREM 7.3. *Let R be a semifir and U its universal field of fractions. If an element p of U can be defined by a reduced admissible matrix A , and there is a matrix for p which remains admissible over an epic R -field K , then A is also K -admissible.*

If, moreover, R is a fully atomic semifir, then every element of U can be defined by a reduced admissible matrix.

Proof. Let $p \in U$ and let A be a reduced admissible matrix defining p . By hypothesis there is a U -admissible matrix B for p , of index 1, which is also K -admissible. By Th. 6.10(iii) we can pass by full operations λ', ρ' from A to A'' and from B to B'' such that A'' and B'' are stably biassociated. Since A is left prime, λ' is trivial and we may take $A'' = A$; thus A is stably biassociated to B'' . Now B is K -admissible, so the system $Bz = 0$ has a unique normalized solution and $B = B''Q^*$, therefore $B''y = 0$ has a normalized solution $y = Q^*z$ in K . It follows that the denominator of B'' is invertible over K . Thus B'' is K -admissible, and hence so is A .

If R is fully atomic, and $p \in U$ is defined by an admissible matrix A , then A can be transformed to a left prime matrix A' by an operation λ' , using Prop. 5.10.1, and A' can be transformed to a minimal reduced matrix by an operation ρ' , using the left-right dual of Prop. 5.10.1, and Th. 6.10(iv). ■

We can now prove the result on the behaviour of the depths over different R -fields announced earlier.

THEOREM 7.4. *Let R be a fully atomic semifir, U its universal field of fractions and K any epic R -field. Then for any element p of U in the domain of the specialization $\phi: U \rightarrow K$ we have $d(p) \geq d(p\phi)$, where d is any one of the depths.*

Proof. Consider the left depth, say. By Th. 7.3, p may be defined by a minimal reduced admissible $m \times (m+1)$ matrix A , which is still admissible for $p\phi$ over K . Now $d(p) = m$ by definition, and since A is K -admissible, $d(p\phi) \leq m$, as claimed. ■

For more general rings the depth may increase on specialization, cf. Ex. 2.

We now turn to compute some examples of depths. They will show in particular that in the free algebra and the free power series ring (on more than one free generator) the depth assumes every value.

Let R be a semifir and U its universal field of fractions. First consider an element of the form

$$p = b_n^{-1} a_n b_{n-1}^{-1} a_{n-1} \dots b_1^{-1} a_1, \quad (6)$$

where no a_i or b_i is a unit or 0. An admissible matrix for $b^{-1}a$ is $(a, -b)$, hence by (5) of 7.1,

$$A = \begin{pmatrix} a_1 & -b_1 & 0 & 0 & \dots \\ 0 & a_2 & -b_2 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots a_n & -b_n \end{pmatrix} \quad (7)$$

is an admissible matrix for p . We shall need a condition for A to be left prime. We recall that two elements a, b of a ring are *totally coprime* if no non-unit factor of a is similar to a factor of b .

LEMMA 7.5. *Let R be a semifir and $a_i, b_i \in R$ ($i = 1, \dots, n$) not zero or units. Then the matrix A in (7) is left prime provided that each a_i is totally coprime to each b_j .*

Proof. Any $n \times n$ left factor P of A must also be a left factor of the square matrices

$$A_1 = \begin{pmatrix} a_1 & -b_1 & & \\ & a_2 & -b_2 & \\ & & \dots & \\ & & & a_n \end{pmatrix}, \quad A_2 = \begin{pmatrix} -b_1 & & & \\ a_2 & -b_2 & & \\ & \dots & & \\ & & & a_n & -b_n \end{pmatrix},$$

obtained by omitting the last and first column, respectively, from A . Now A_1 has $a_1, \dots, a_n$ down the main diagonal and 0 below it, hence by unique factorization over R , P must be a product of factors similar to factors of $a_1 \dots a_n$ and therefore totally coprime to $b_1 \dots b_n$. But A_2 has (apart from sign) $b_1, \dots, b_n$ down the main diagonal and 0 above it, so P is also a product of factors similar to factors of $b_1 \dots b_n$. This is possible only when P is a unit, hence A is indeed left prime. ■

We can now give sufficient conditions for the element (6) to have left depth n :

THEOREM 7.6. *Let R be a semifir and U its universal field of fractions. Consider the element (6) in U , where the a_i, b_i are in R and are neither 0 nor units. If each a_i is totally coprime to each b_j and there exists a proper ideal c of R containing all the a_i, b_i such that R/c is weakly finite, then the left depth of p , as element of U , is n .*

Proof. We know that the matrix A in (7) is admissible for p ; if we can show it to be minimal reduced, then it will follow that $d_l(p) = n$. By Lemma 7.5, A is left prime; its core A_* is of the same form as A but transposed, and the same argument shows A_* to be right prime, thus A is right $*$ -prime, hence reduced. By Lemma 3.4.11, A is not stably associated to a matrix of smaller order, hence A is minimal and the result follows. ■

To give an example, in the free algebra $R = k\langle x, y \rangle$, x and y are dissimilar atoms, and if c is the ideal generated by x, y , then $R/c \cong k$, hence $(y^{-1}x)^n$ has depth n . The same holds in the free power series ring $k\langle\langle x, y \rangle\rangle$. As an example not covered by Th. 7.6, consider in $k\langle\langle x, y, z \rangle\rangle$ the element $y^{-1}zx^{-1}yz^{-1}x$. This is determined by the admissible matrix

$$\begin{pmatrix} x & -z & 0 & 0 \\ 0 & y & -x & 0 \\ 0 & 0 & z & -y \end{pmatrix}. \quad (8)$$

It can be shown by a direct argument that the depth is 3 (cf. Ex. 5).

As a second type of element consider

$$q = a_1^{-1} + a_2^{-1} + \dots + a_n^{-1}, \quad \text{where } 0 \neq a_i \in R. \quad (9)$$

An admissible matrix for q is given by

$$B = \begin{pmatrix} 1 & -a_1 & -a_1 & \dots & \dots & -a_1 \\ 1 & a_2 & 0 & \dots & \dots & 0 \\ 1 & 0 & a_3 & \dots & \dots & 0 \\ & \dots & & \dots & \dots & \\ 1 & 0 & 0 & \dots & \dots & a_n & 0 \end{pmatrix}. \quad (10)$$

As in the proof of Lemma 7.5, by deleting each column in turn, we can show that B is left prime provided that any two of the a 's are totally coprime. Under this condition it also follows that B is right $*$ -prime, and if B is stably associated to an $m \times (m+1)$ matrix, then $m \geq n$, provided that there exists a proper ideal c containing all the a_i such that R/c is weakly finite. Hence we obtain

PROPOSITION 7.7. *Let R be a semifir with universal field of fractions U . Given $a_1, \dots, a_n \in R$, each a_i not 0 or a unit and such that any two are totally coprime and if, further, there exists a proper ideal c containing all the a_i such that R/c is weakly finite, then the element (9) has depth n . ■*

Exercises 7.7

1. In the universal field of fractions of $k\langle x, y, z \rangle$ compute the four depths of xy^{-1} , $y^{-1}z$ and $xy^{-1}z$.

2. (Bergman) Let R be the k -algebra generated by x, y, z, w with defining relation $wx = yz$ and denote by U, V the universal fields of fractions of the free algebras $k\langle x, y, z \rangle, k\langle x, z \rangle$, respectively. Verify that there is a homomorphism $f: R \rightarrow U$ with $x, y, z, w \mapsto x, y, z, yzx^{-1}$ and a homomorphism $R \rightarrow V$ with $x, y, z, w \mapsto x, 0, z, 0$. Show that U, V are epic R -fields with a specialization $U \rightarrow V$ defined by $x, y, z \mapsto x, 0, z$. Show also that xz^{-1} has left depth 1 in U but that its left depth in V is 2.

3. Let R be a semifir with universal field of fractions U . Given $p \in U$, let A, B be two admissible matrices for p . By applying Th. 6.10(iii), find an admissible matrix C for p which is K -admissible (for any epic R -field K) whenever either A or B is.

4. Obtain analogues of Lemma 7.5 and Th. 7.6 when R is a full matrix ring over a semifir.

5. Prove that the matrix (8) is reduced, by showing that it has no left factor similar to x, y or z . (Hint: If the matrix is A , consider a comaximal relation $uA = xA'$, where $u \in R^3$ and use the fact that the eigenring of x is k .) Apply Lemma 3.4.11 to show that the matrix is not stably biassociated to a matrix of lower order and hence show that $d_1(y^{-1}zx^{-1}yz^{-1}x) = 3$.

6. Use the method of Ex. 5 to show that the left depth of $xy^{-1}x + x^2y^{-2}x^2 + \dots + x^n y^{-n} x^n$ in $k\langle\langle x, y \rangle\rangle$ is $n+1$ (cf. Cohn [85]).

7. Find the depth of $xy^{-1}x + x^2y^{-2}x + \dots + x^n y^{-n} x$ and of $xy^{-1}x + x^2y^{-1}x^2 + \dots + x^n y^{-1} x^n$ in $k\langle x, y \rangle$.

8°. Find a non-Ore domain in which the lower depth is bounded by 1.

7.8 Centralizers in universal fields of fractions of firs

It is obvious that the centre of the free algebra $k\langle X \rangle$, where $|X| > 1$, is just k ; that the centre of its universal field of fractions is also k is less evident. It will follow from the results to be proved in this section, in which we determine more generally the centre of the universal field of fractions of a fir. We begin with a technical result on relations between elements of a semifir and of its universal field of fractions.

LEMMA 8.1. Let R be a semifir with a universal field of fractions U . Given $p, p' \in U$ and $c, c' \in R$ such that

$$cp' = pc', \quad (1)$$

suppose further that there are reduced admissible matrices $A \in {}^m R^{m+1}$ for $p, A' \in {}^n R^{n+1}$ for p' (this is so, for example, when R is fully atomic). Then there exist $P \in {}^m R^n$ and $Q^* \in {}^{m+1} R^{n+1}$ such that

$$PA' = AQ^*, \quad \text{where } Q^* = \begin{pmatrix} c' & 0 & 0 \\ Q_0 & Q_* & Q_\infty \\ 0 & 0 & c \end{pmatrix}. \quad (2)$$

Moreover, given any one solution P', Q'^* of (2), the general solution has the form

$$P = P' + A_* N, Q^* = Q'^* + N^* A', \quad \text{where } N^* = (0 \ N \ 0)^T, N \in {}^{m-1} R^n. \quad (3)$$

Proof. Since A is a reduced admissible matrix for p, AC , where $C = c' \oplus I_{m-1} \oplus c$, is an admissible matrix for p' . Now A is reduced, so AC is certainly right $*$ -prime and we may write $AC = P_1 A_1$, where $P_1 \in R_m$ is full and A_1 is left prime and hence a reduced admissible matrix for p' . By Th. 6.10(iv), A_1 and A' are stably biassociated, so there is a comaximal relation $P_0 A' = A_1 Q_1^*$, where Q_1^* is bordered. Hence $P_1 P_0 A' = AC Q_1^*$; writing $P = P_1 P_0$ and $Q^* = C Q_1^*$, we obtain (2).

Suppose now that P', Q'^* and P'', Q''^* both satisfy (2); then on writing $P = P' - P'', Q^* = Q'^* - Q''^*$, we have

$$PA' = AQ^*, \quad \text{where } Q^* = (0 \ Q \ 0)^T \text{ and } Q \in {}^{m-1} R^{n+1}. \quad (4)$$

From the form of Q^* the right-hand side does not involve the first or last column of A , so (4) may be written

$$PA' = A_* Q, \quad (5)$$

or also $(P, A_*)(-A', Q)^T = 0$. By the law of nullity it follows that

$$\rho(P \ A_*) + \rho \begin{pmatrix} A' \\ Q \end{pmatrix} \leq m + n - 1.$$

Since $\rho(P, A_*) \geq \rho A_* = m - 1$, $\rho((A', Q)^T) \geq \rho A' = n$, we have equality in both places and so

$$(P \ A_*) = B(N \ C), \quad \text{where } B \in {}^m R^{m-1}, N \in {}^{m-1} R^m, C \in R_{m-1}.$$

Since A_* is right prime, C is a unit, so we may replace B by BC and C by I and obtain

$$(P \ A_*) = A_*(N \ I).$$

This shows that $P = A_*N$; inserting this value in (5), we find that $A_*NA' = A_*Q$ and since A_* is right full, it is right regular, so we may cancel it and obtain $Q = NA'$. Now $Q^* = N^*A'$ follows, where Q^*, N^* are obtained by topping and tailing Q, N by rows of zeros. This shows that all solutions of (2) can be expressed in terms of a given one as in (3), and it is clear that all the expressions (3) are solutions of (2). ■

We remark that the same result holds, with the same proof, when p, p', c, c' are full matrices.

We shall use Lemma 8.1 with Cor. 7.2 to show that under suitable conditions on c, c' , any p occurring in an equation $cp = pc'$ has depth at most 1.

PROPOSITION 8.2. *Let R be a fully atomic semifir with universal field of fractions U . Given $p \in U, c, c' \in R$ such that*

$$cp = pc', \quad (6)$$

assume further that either (i) c, c' are not zero or units or (ii) R is a persistent semifir over a field k and c, c' are not algebraic over k . Then p has left and right depth at most 1, say

$$p = a_1 b_1^{-1} = b_2^{-1} a_2, \quad \text{where } a_2 b_1 = b_2 a_1 \text{ is coprime,} \quad (7)$$

and for any such representation (7) of p , there exists $c'' \in R$ such that

$$ca_1 = a_1 c'', \quad c' b_1 = b_1 c''. \quad (8)$$

Moreover, in case (ii), there exist $a \in R$ and a non-zero polynomial f over k such that

$$p = f(c)^{-1} a = a f(c')^{-1}, \quad ca = ac'. \quad (9)$$

Proof. Since R is fully atomic, there is a reduced admissible matrix A for p , and by Lemma 8.1 there exist P, Q^* satisfying

$$PA = AQ^*, \quad \text{where } Q^* \text{ is as in (2).} \quad (10)$$

We first take case (i), then c, c' are non-units. If A is $m \times (m+1)$, then from (10) we have

$$P(A_0 \ A_*) = (A_0 \ A_*) \begin{pmatrix} c' & 0 \\ Q_0 & Q_* \end{pmatrix}, \quad P(A_* \ A_\infty) = (A_* \ A_\infty) \begin{pmatrix} Q_*^* & Q_\infty^* \\ 0 & c \end{pmatrix}. \quad (11)$$

Let us assume that p has left depth greater than 1; then A_* is regular over R , by Cor. 7.2. Moreover, given one solution P', Q'^* of (2), we can write the

general solution in the form

$$P = P' + A_* N, \quad Q^* = Q'^* + N^* A', \quad \text{where } N^* = (0 \quad N \quad 0)^T.$$

Thus we have

$$Q_* = Q'_* + N A_*, \quad \text{where } N \in {}^{m-1}R^m. \quad (12)$$

Now A_* has rank $m-1$, so to each row of Q'_* we can add a linear combination of the rows of A_* so as to ensure that Q'_* has rank $m-1$ over U . Hence for a suitable choice of N in (12), Q_* has inner rank $m-1$ over R and so is full. Since $c, c' \neq 0$, it follows that Q^* is full, therefore the two sides in (11) have rank m , hence P is full. By unique factorization, P and $\begin{pmatrix} c' & 0 \\ Q_0 & Q_* \end{pmatrix}$ are products of similar atomic factors, in particular they have the same length. Now A is left prime, hence (10) is left coprime, therefore P is similar to a left factor P_0 , say, of Q^* , thus

$$Q^* = \begin{pmatrix} I & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} c' & 0 & 0 \\ Q_0 & Q_* & Q_\infty \\ 0 & 0 & 1 \end{pmatrix} = P_0 P_1, \quad (13)$$

where P_0 is similar to P and P_1 is a highest common right factor of A and Q^* (cf. (10)). It follows from this and (11) that P_0 and $\begin{pmatrix} c' & 0 \\ Q_0 & Q_* \end{pmatrix}$ are products of similar factors; by unique factorization applied to (13), P_1 and c are products of similar factors; in particular, P_1 and c have the same length. Moreover, as we saw, P_1 is a right factor of A ; thus A has a full right factor of length $l(c)$. Now $cp = pc'$ implies $c^n p = pc'^n$ for all $n \geq 1$. If we repeat the argument with c, c' replaced by c^n, c'^n , we find that A has a full right factor of length $l(c^n) = nl(c)$. But the length of full right factors of A is bounded, by Prop. 5.10.1, so we have a contradiction. This shows the left depth to be 1, and the same holds for the right depth, by symmetry.

We now have $p = a_1 b_1^{-1} = b_2^{-1} a_2$, hence

$$b_2 a_1 = a_2 b_1, \quad (14)$$

and we may take this relation to be coprime, without loss of generality, hence comaximal. Since $cp = pc'$, we also have

$$b_2 c a_1 = a_2 c' b_1.$$

Because (14) was comaximal, there exists $c'' \in R$ such that (8) holds, and this completes the proof of case (i).

We now turn to case (ii). If p had right depth greater than 1, A would be

regular, hence its eigenring would be algebraic over k , by Th. 4.6.4. Let f be a non-zero polynomial over k satisfied by Q^* (mod $R_m A$), which exists by (10), then the first row of $f(Q^*)$ takes the form

$$(f(c') \quad 0 \quad \dots \quad 0) = sA$$

for some $s \in R^m$. Thus $s(A_*, A_\infty) = 0$, but (A_*, A_∞) is full and so regular, hence $s = 0$ and it follows that $f(c') = 0$. This contradicts the assumption that c' is not algebraic over k , and it shows that the right depth of p is at most 1. By symmetry the same is true of the left depth, and we can prove (8) as before.

Now in (8), $b_1 \neq 0$, so by another application of Th. 4.6.4, we can find $f \in k[t]^\times$ such that

$$f(c') = b_1 d,$$

and $d \neq 0$, because c' is not algebraic over k . If we put $a = a_1 d$, then $p = a_1 b_1^{-1} = a f(c')^{-1}$, but by (6), $f(c)p = p f(c') = a$, hence

$$p = f(c)^{-1} a = a f(c')^{-1}.$$

Moreover, $ca = c f(c)p = f(c)cp = f(c)pc' = ac'$, so (9) holds. ■

In Prop. 8.2(i), if p has depth greater than 1, then c, c' are 0 or units, hence the elements of R , whose conjugates under p again lie in R , form a subfield of R :

COROLLARY 8.3. *Let R be a fully atomic semifir with universal field of fractions U . If $p \in U$ has depth greater than 1, then $pRp^{-1} \cap R$ is a subfield of R . ■*

As a further application of Prop. 8.2 we determine the centre of U :

THEOREM 8.4. *Let R be a fully atomic semifir with universal field of fractions U . Then either the centre of U coincides with the centre of R , or R is a principal ideal domain and the centre of U consists of all elements $ab^{-1} = b^{-1}a$ ($a, b \in R$) for which an automorphism α of R exists such that $ra = ar^\alpha$, $rb = br^\alpha$ for all $r \in R$.*

Proof. If $U = R$, there is nothing to prove; otherwise take a non-zero non-unit c in R and p in the centre of U , then $cp = pc$, so p has depth 1, by Prop. 8.2, say $p = ab^{-1}$. Now $pb = bp$, hence $ab = ba$ and so $p = b^{-1}a$. If we choose a representation $p = ab^{-1} = b^{-1}a$ with ab of minimal length, then a and b are left and right coprime, and by Prop. 8.2 we have, for each $r \in R$ an r' such that $ra = ar'$, $rb = br'$. It is clear that the map $\alpha: r \mapsto r'$ is an endomorphism; by symmetry it has an inverse and so is an automorphism of R . If

$r' = r$ for all $r \in R$ and a, b are units, then U and R have the same centre. Otherwise R is a two-sided Ore domain, by Th. 6.3.3, and hence a principal ideal domain. Now the form of the centre of R is given by Prop. 6.3.5. ■

The exception does, in fact, occur since, as we have seen in 6.1, every Krull domain can occur as the centre of a principal ideal domain. To illustrate these results, let D be a field which is a k -algebra and X a set with more than one element. Then the free D -ring over k , $D_k\langle X \rangle$ is a fir; its centre is k and it is not principal, therefore its universal field of fractions has centre k , by Th. 8.4. We state this as

COROLLARY 8.5. *Let D be a field and k a subfield of its centre, and let X be any set. If either $|X| > 1$ or $X \neq \emptyset$ and $D \neq k$, then the universal field of fractions of $D_k\langle X \rangle$ has centre k . ■*

Next we look at centralizers in the universal field of fractions. To put the result in context we recall that if D is any field with centre k and $a \in D$ is algebraic of degree n over k , then the centralizer C_a of a in D is a subfield such that $[D : C_a] = n$ (cf. e.g. Cohn [77], p. 366). This may be expressed loosely by saying that a has a 'large' centralizer; by contrast, in the universal field of fractions of a semifir R , the elements of R have 'small' centralizers, as the next result shows. We need to assume persistence; this holds e.g. for free algebras.

THEOREM 8.6. *Let R be a persistent semifir over k , and suppose that R is fully atomic, with universal field of fractions U . Given $c \in R$, not algebraic over k ,*

(i) *if C is the centralizer of c in R , then its centralizer in U is the localization $C_{k[c]^\times}$; moreover, C is an Ore domain and*

(ii) *if R contains the field of fractions $k(c)$ and c is conjugate in U to an element c' of R , say $cp = pc'$, then p is a unit in R and R also contains $k(c')$.*

Proof. Let $p \in C$. By Prop. 8.2 there exist $a \in R$ and $f \in k[c]^\times$ such that $p = f(c)^{-1}a = af(c)^{-1}$ and $ac = ca$, therefore $a \in C$ and the form of the centralizer follows. Now let $p, q \in C$, $q \neq 0$; then $q^{-1}p = af(c)^{-1}$ for some $a \in C$, by what has been proved, hence $pf(c) = qa$ and this shows C to be a right Ore domain; by symmetry it is also left Ore.

(ii) Again by Prop. 8.2 we have $p = f(c)^{-1}a$, but now $f(c)^{-1} \in R$, so $p \in R$. Applying the same proposition to p^{-1} we find that $p^{-1} = bg(c)^{-1}$, hence $p^{-1} \in R$ and this shows p to be a unit in R ; further R contains $p^{-1}k(c)p = k(c')$. ■

The condition of persistence cannot be omitted from Th. 8.6, as is shown by the example $R = F\langle X \rangle$, where F is a commutative transcendental field extension of k and $|X| > 1$. This is a fir, but the centralizer of any $x \in X$ contains $F(x)$ and so is larger than $F[x]_{k[x]}^*$.

If K is a field with a subring R , then the (right) *normal closure* of R in K may be defined as RN , where N is the right *normalizer* of R in K :

$$N = \{p \in K \mid Rp \subseteq pR\}. \quad (15)$$

Thus RN is the subring of K generated over R by the right R -invariant elements in K . The left normal closure is defined similarly and the normal closure is generated over R by the R -invariant elements in K . We shall examine some simple cases where R is equal to its normal closure.

PROPOSITION 8.7. *A free algebra $R = k\langle X \rangle$ on more than one free generator is equal to its right normal closure in its universal field of fractions.*

Proof. Let $p \in N$, where N is defined as in (15) in terms of R and U , then for any $c \in R$ there exists $c' \in R$ such that $cp = pc'$. By Prop. 8.2, there is a non-zero polynomial f over k and $a \in R$ such that

$$ca = ac', \quad p = f(c)^{-1}a = af(c')^{-1}. \quad (16)$$

Let x, y be distinct elements of X ; taking $c = x, y$ in turn, we find polynomials f, g and elements a, b such that $p = f(x)^{-1}a = g(y)^{-1}b$. Since we are in a semifir, there exist by Lemma 10.2, below, elements $a_1, b_1, c_1, d_1 \in R$ such that

$$a = a_1c_1, \quad b = b_1c_1, \quad f(x) = a_1d_1, \quad g(y) = b_1d_1. \quad (17)$$

Now $f(x), g(y)$ are totally coprime, because over an algebraic closure of k we can write $f(x) = \lambda \prod_i (x - \alpha_i)$, $g(y) = \mu \prod_j (y - \beta_j)$, and these two factorizations are clearly totally coprime. It follows that d_1 is a unit, and so may be taken to be 1. But then (17) reduces to $a = f(x)c_1$, $b = g(y)c_1$, hence $p = f(x)^{-1}a = c_1 \in R$, and the result follows. ■

The same conclusion can be proved in a second case, although it probably holds in a much wider class.

PROPOSITION 8.8. *Let R be a persistent semifir over a field k , which is fully atomic and a local ring. Then R is equal to its right normal closure in its universal field of fractions, unless R is a left discrete valuation ring.*

Proof. Let $p \in N$; for any non-unit $x \in R$ there exists $x' \in R$ such that $xp = px'$. By Prop. 8.2 there exist a non-zero polynomial f over k and $a \in R$ such that

$$xa = ax', \quad p = f(x)^{-1}a = af(x')^{-1}. \quad (18)$$

Now $f = \lambda_r x^r + \lambda_{r+1} x^{r+1} + \dots + \lambda_n x^n$, where $\lambda_i \in k$, $\lambda_r \neq 0$. Hence $f(x)$ is associated to x^r in R and so we may in (18) replace $f(x)$ by x^r and find that $p = x^{-r}a = ax'^{-r}$.

Let x, y be any atoms in R ; then this argument shows that

$$p = x^{-r}a = y^{-s}b;$$

since we are in a semifir, we can as before find $a_1, b_1, c, d \in R$ such that

$$a = a_1 c, \quad b = b_1 c, \quad x^r = a_1 d, \quad y^s = b_1 d. \quad (19)$$

If d is a unit, we can (as before) absorb it in a_1, b_1 and write $a = x^r c, b = y^s c$, hence $p = c \in R$. Thus if $p \notin R$, then d cannot be a unit. Then (19) shows that x is a right factor of d , hence x, y are left associated; thus any two atoms are left associated. Taking an atom x , say, we have for any unit u , $xu = u'x$ for some unit u' in R , hence every non-zero element of R is of the form ux^n , $n \geq 0$, u a unit. This means that R is a left principal valuation ring, by Prop. 2.9.11. ■

Exercises 7.8

1. Let R be a doubly persistent semifir over k , fully atomic, with universal field of fractions U . Show that if x is any element of R , not algebraic over k , then the centralizer of x in U is matrix algebraic over k .

2. Let D be a field with central subfield k such that $D \otimes k(s)$ is 1-inert in $D \otimes k(s, t)$ (for central indeterminates s, t) and D is a doubly persistent semifir over k . Show that D must be commutative.

3. (W. Dicks) Let $R = k\langle X \rangle$, U its universal field of fractions, and let $k(s)$ be a purely transcendental field extension. Show that $U \otimes k(s)$ is an integral domain and remains one under all commutative field extensions of $k(s)$. Deduce that $D \otimes k(s)$ is 1-inert in $D \otimes k(s, t)$ and hence that D is not doubly persistent.

4. Let $c \in k\langle X \rangle \setminus k$, so that the centralizer of c has the form $k[p]$, by Bergman's centralizer theorem (Th. 6.7.7). Show that the centralizer of c in the universal field of fractions of $k\langle X \rangle$ is $k(p)$.

5. Verify that the normal closure of R in K is $RN = NR$, where N is the intersection of the left and right normalizers of R .

6. Show that $R = k\langle x, y \rangle$ is a persistent semifir over k ; deduce that the normal closure of R in its universal field of fractions is R itself.

7⁰. For which semifirs R is it the case that R equals the normal closure in its universal field of fractions?

8*. Let $R = k\langle X \rangle$ and U its universal field of fractions. Suppose that $p \in U$ is quadratic over R : $p^2 - ap - b = 0$. If the companion matrix $Z = \begin{pmatrix} 0 & 1 \\ b & a \end{pmatrix}$ is algebraic over k , show that $p \in R$ (consider first the case that k is algebraically closed). When Z is not algebraic over k and p has left and right depth greater than 1, show that the admissible matrix for p has algebraic eigenring and obtain a contradiction. Deduce that again $p \in R$, thus R is 'quadratically closed' in U .

9⁰. Show that $k\langle X \rangle$ is 'relatively algebraically closed' in its universal field of fractions.

10. (Klein [72']) Show that for any commutative field k there is a skew field with an involution, and with k as centre.

11⁰. For a given commutative field k find a skew field K with an involution whose fixed point set is k .

7.9 Determinants and valuations

The determinant is an important scalar invariant associated with matrices over a commutative ring. For general rings there is no satisfactory analogue, although for skew fields one has the Dieudonné determinant, which in effect provides a universal homomorphism from $GL_n(K)$ to abelian groups. Our main object in this section is to calculate the Dieudonné determinant for the universal field of fractions of a fir, particularly for the free algebra $k\langle X \rangle$. However, we shall look at this problem in a more general setting, which will also allow us to describe valuations on these fields in terms of the rings.

Let R be any ring, Σ a multiplicative set of square matrices over R and G an abelian group, written additively. By a G -value on Σ we understand a mapping $v: \Sigma \rightarrow G$ such that

V.1. $v(E) = 0$ for any elementary matrix E ,

V.2. $v(A \oplus I) = v(A)$ for $A \in \Sigma$ and

V.3. $v(AB) = v(A) + v(B)$ for $A, B, AB \in \Sigma$.

If we write $v(-1) = \varepsilon$, then by V.1, 3, $2\varepsilon = 0$; thus if G is torsion-free, $\varepsilon = 0$, and we have

V.4. $v(-1) = 0$.

However, we shall not assume V.4 in what follows. We note the following consequences of V.1–3:

V.5. $v(I \oplus A) = v(A)$.

This follows from the equation

$$\begin{pmatrix} I & 0 \\ 0 & A \end{pmatrix} = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix},$$

by applying v and using V.1-3.

$$\text{V.6. } v \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = v(A) + v(B) = v \begin{pmatrix} A & 0 \\ D & B \end{pmatrix}.$$

The first equation follows by writing

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I & C \\ 0 & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix}.$$

The right-hand side is $v(I \oplus B) + v(A \oplus I)$, by V.1, 3 and this reduces to $v(A) + v(B)$, by V.2, 5. The second equation V.6 follows similarly.

By V.1-3, $v(A) = v(B)$ whenever A is stably E -associated to B . Further, two matrices 'differing' by a non-full matrix have the same value:

LEMMA 9.1. *Let R be a ring, Σ a multiplicative set of matrices, and v a G -value on Σ . In any equation*

$$C = A \nabla B,$$

where $A, C \in \Sigma$ and B is non-full, we have $v(A) = v(C)$.

Proof. Suppose that the determinantal sum is with respect to the first column, say $A = (A_1, A_2, \dots, A_n)$, $B = (B_1, A_2, \dots, A_n)$. Since B is non-full, $B = PQ$, where $P \in {}^n R^{n-1}$, $Q \in {}^{n-1} R^n$. Write $Q = (Q_1, Q')$, where $Q' \in R_{n-1}$; then

$$A = (A_1 \quad PQ') = (A_1 \quad P) \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix}, \quad (1)$$

$$B = PQ = (PQ_1 \quad PQ') = (PQ_1 \quad P) \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix}, \quad (2)$$

hence

$$C = (A_1 + PQ_1 \quad PQ') = (A_1 + PQ_1 \quad P) \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix} = (A_1 \quad P) \begin{pmatrix} 1 & 0 \\ Q_1 & I \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q' \end{pmatrix}. \quad (3)$$

Now a comparison of (1) and (3) (bearing in mind V.1) gives the result. ■

This result sheds further light on Th. 5.9; it shows that in any ring with UGN, if the product of two full matrices (when defined) is again full, then the determinantal sum of two non-full matrices (when defined) is non-full. For let $C = A \nabla B$, where A, B are non-full, then $A = C \nabla B'$, where B' is the matrix B

with the sign of a row or column changed, hence B' is again non-full. If C is full, then by (1), (3) A is full, which is a contradiction.

In the applications Σ will usually be Φ , the set of all full matrices; moreover Φ will need to admit diagonal sums and products (when defined). By the previous remark the set of non-full matrices is then a prime matrix ideal, hence by Th. 5.10, R is then a Sylvester domain. For these reasons we assume from now on that R is a Sylvester domain; then R has a universal field of fractions U , with a fully inverting homomorphism $\lambda: R \rightarrow U$.

Let v be a G -value on $\Phi(R)$; our object is to extend v to U . If $p \in U^\times$ has the admissible system $Au = 0$, our aim will be to define v by putting

$$v(p) = v(A_0 \ A_*) - v(-A_\infty \ A_*), \quad (4)$$

but to do so we must show that the right-hand side of (4) is independent of the choice of A . We shall in the first instance define v on admissible matrices. Let $A = (A_0, A_*, A_\infty)$ be an admissible matrix for a non-zero element of U , then (A_0, A_*) and $(-A_\infty, A_*)$ are both full, and we put

$$V(A) = v(A_0 \ A_*) - v(-A_\infty \ A_*). \quad (5)$$

If A is admissible for p and B is admissible for q , then the matrix

$$A \cdot B = \begin{pmatrix} B_0 & B_* & B_\infty & 0 & 0 \\ 0 & 0 & A_0 & A_* & A_\infty \end{pmatrix} \quad (6)$$

is admissible for pq , and we have

$$V(A \cdot B) = v \begin{pmatrix} B_0 & B_* & B_\infty & 0 \\ 0 & 0 & A_0 & A_* \end{pmatrix} - v \begin{pmatrix} 0 & B_* & B_\infty & 0 \\ -A_\infty & 0 & A_0 & A_* \end{pmatrix}.$$

Here the second term becomes, by a column interchange,

$$v \begin{pmatrix} -B_\infty & B_* & 0 & 0 \\ -A_0 & 0 & -A_\infty & A_* \end{pmatrix},$$

hence we obtain (using V.6),

$$\begin{aligned} V(A \cdot B) &= v(B_0 \ B_*) + v(A_0 \ A_*) - v(-B_\infty \ B_*) - v(-A_\infty \ A_*) \\ &= V(A) + V(B). \end{aligned}$$

Further, $(1, -1)$ is an admissible system for 1, and

$$V(1 \ -1) = v(1) - v(1) = 0.$$

If S is the semigroup of admissible systems with the multiplication (6) (clearly an associative operation), then $V: S \rightarrow G$ is a homomorphism, by what has been shown; we still have to show that V has the same value on all admissible matrices defining a given $p \in U$.

Let $A = (A_0, A_*, A_\infty)$ be admissible for p , then $\bar{A} = (-A_\infty, A_*, -A_0)$ is admissible for p^{-1} and hence $A \cdot \bar{A}$ is admissible for $pp^{-1} = 1$. Hence if A, B are any matrices admissible for the same p , then $A \cdot \bar{B}$ is admissible for $pp^{-1} = 1$, and clearly $V(\bar{B}) = -V(B)$, hence

$$V(A \cdot \bar{B}) = V(A) + V(\bar{B}) = V(A) - V(B).$$

So to prove that v is well-defined we need only show that $V(C) = 0$ for any admissible matrix C for 1.

Suppose then that C is admissible for 1; the condition for this is that $(C_0 + C_\infty, C_*)$ is non-full. Put $D = C_0 + C_\infty$; then

$$(C_0 \quad C_*) = (D - C_\infty \quad C_*) = (D \quad C_*) \nabla (-C_\infty \quad C_*),$$

and by Lemma 9.1, $V(C) = v(C_0, C_*) - v(-C_\infty, C_*) = 0$, as claimed.

Thus we have proved

THEOREM 9.2. *Let R be a Sylvester domain with universal field of fractions U and v any G -value on $\Phi(R)$. Then v can be extended to a homomorphism of $U^\times$ into G by (4), where A is any admissible matrix for $p \in U^\times$. ■*

Let us see how this applies to determinants. If K is any field and $n \geq 1$, then any matrix $A \in \text{GL}_n(K)$ can be reduced to diagonal form by elementary transformations; more precisely, we have

$$A = (I_{n-1} \oplus \alpha)U, \quad \text{where } \alpha \in K^\times, U \in E_n(K).$$

If $A = (I \oplus \beta)V$ is another such expression, then $\alpha\beta^{-1}$ belongs to $K^{\times'}$, the derived group of $K^\times$. Hence the coset of α in $K^{ab} = K^\times/K^{\times'}$ is an invariant of the matrix, called the *Dieudonné determinant* (cf. Dieudonné [43], Artin [57], p. 152). Let us denote this coset by $\text{Det } A$; it may be shown that $A \mapsto \text{Det } A$ is a homomorphism, more precisely, we have the diagram shown,

$$\begin{array}{ccc} \text{GL}_n(K) & \xrightarrow{\text{Det}} & K^{ab} \\ & \searrow & \downarrow \\ & & \Gamma \end{array}$$

where Det is universal for homomorphisms into abelian groups (cf. Bass [68]). For any field K and any $n \geq 1$ (except $K = \mathbb{F}_2$ and $n = 2$) there is an isomorphism

$$\text{GL}_n(K)^{ab} \cong K^{ab}. \quad (7)$$

We can define matrix multiplication generally for square matrices of different orders over a ring R by regarding A as $A \oplus I$, for I of sufficiently

large order. In this way Φ becomes a monoid (for a Sylvester domain) and we shall write Φ^{ab} for the universal abelian group of Φ . The same device allows us to embed $\mathbf{GL}_n(R)$ in $\mathbf{GL}_{n+1}(R)$ and define the *stable* general linear group $\mathbf{GL}(R) = \varinjlim \mathbf{GL}_n(R)$. Its abelianization is called the *Whitehead group*, written

$$\mathbf{K}_1(R) = \mathbf{GL}(R)^{ab}. \quad (8)$$

Now let R be a Sylvester domain and U its universal field of fractions. The natural map $\Phi \rightarrow \mathbf{GL}(U)^{ab}$ yields a homomorphism

$$f: \Phi^{ab} \rightarrow \mathbf{K}_1(U). \quad (9)$$

Now the natural map $\Phi \rightarrow \Phi^{ab}$ is a Φ^{ab} -value on Φ which by Th. 9.2 can be extended to $U^\times$; in this way we obtain a homomorphism from $\mathbf{K}_1(U) = \mathbf{GL}(U)^{ab}$ to Φ^{ab} which is clearly inverse to (9). Hence the latter is an isomorphism and we have proved

THEOREM 9.3. *Let R be a Sylvester domain, U its universal field of fractions and Φ^{ab} the universal abelian group of the monoid of full matrices over R . Then there is a natural homomorphism*

$$\mathbf{K}_1(U) \cong \Phi^{ab}. \blacksquare$$

To investigate the structure of this group more closely let us return to Th. 9.2 and consider the case $G = \mathbf{Z}$. By an $\mathbf{N}$ -value on R we understand a function v defined on $\Phi(R)$ with values in $\mathbf{N}$, satisfying V.1–3. Since $\mathbf{N} \subseteq \mathbf{Z}$, Th. 9.2 shows that v can be extended to a $\mathbf{Z}$ -value on $U^\times$; here $v(-1) = 0$, because $\mathbf{Z}$ is torsion-free.

PROPOSITION 9.4. *Let R be a Sylvester domain and v any $\mathbf{N}$ -value on R , then $v(P) = 0$ for $P \in \mathbf{GL}(R)$ and*

$$v(A) = v(A'), \quad (10)$$

whenever A and A' are stably associated.

For if $P \in \mathbf{GL}(R)$, then $v(P) \geq 0$, $v(P^{-1}) \geq 0$, but $v(P) + v(P^{-1}) = v(I) = 0$, hence $v(P) = 0$. Now (10) is an immediate consequence. $\blacksquare$

To give an example of $\mathbf{N}$ -values, let R be a fully atomic semifir and let us define a *prime* as a class of stably associated atomic matrices over R . For each prime p_i we define an $\mathbf{N}$ -value v_i as follows. Given $A \in \Phi$, we put $v_i(A) = r$ if the prime p_i occurs just r times in a complete factorization of A . By unique factorization, r is independent of the factorization chosen and V.1–3 are

easily checked. We shall call v_i the simple N-value associated with p_i . More generally, if for each prime p_i we pick an integer $n_i \geq 0$, then $w = \sum n_i v_i$ is an N-value, for it is defined on each full matrix A : $w(A) = \sum n_i v_i(A)$, where the sum is finite because $v_i(A) = 0$ for almost all i . In fact every N-value arises in this way, for if w is any N-value on R , take an atom P_i in the class p_i and put $n_i = w(P_i)$, then w and $\sum n_i v_i$ have the same value on each atom and hence on all of Φ . This proves

THEOREM 9.5. *Let R be a fully atomic semifir and let (v_i) be the simple N-values corresponding to the primes of R . Then for any family (n_i) of non-negative integers, $\sum n_i v_i$ is an N-value and conversely, every N-value is of this form. ■*

We remark that with every full matrix A there is associated an N-value w_A which is simple if and only if A is an atom, viz. $w_A = \sum n_i v_i$, where $n_i = v_i(A)$ and v_i runs over the simple N-values. We can also use N-values to characterize fully atomic semifirs.

PROPOSITION 9.6. *Let R be a semifir. Then R is fully atomic if and only if there is an N-value w on R such that $w(A) = 0$ precisely when A is a unit.*

Proof. If R is a fully atomic semifir and v_i are the simple N-values corresponding to the different primes of R , then $w = \sum v_i$ has the desired property. Conversely, when w exists, take any full matrix A and factorize it into non-units in any way:

$$A = P_1 \dots P_r. \quad (11)$$

Since $w(P_i) \geq 1$ by hypothesis, we have $w(A) = \sum w(P_i) \geq r$, and this provides a bound for the number of factors in (11). Hence we obtain a complete factorization of A by choosing r maximal. ■

This result can be extended to fully atomic Sylvester domains if we define a prime in this case as an equivalence class of atoms, where A, B are considered equivalent if $v(A) = v(B)$ for each N-value v on R . Such primes are of course unions of classes of stably associated atoms (by Prop. 9.4), but in general they may include atoms that are not stably associated. ✓

Consider a fully atomic semifir R and let p_i ($i \in I$) be the family of all primes. For each prime p_i we have a homomorphism $v_i: \Phi^{ab} \rightarrow \mathbb{Z}$, and combining all these maps, we have a homomorphism

$$\Phi(R)^{ab} \rightarrow \mathbb{Z}^I.$$

But each full matrix maps to 0 in almost all factors of $\mathbb{Z}^I$, hence the image lies in the weak direct power $\mathbb{Z}^{(I)}$. Let us write $D = D(R)$ for the free abelian

group on the p_i , written additively, and call D the *divisor group* of R . We have a homomorphism $\lambda: \Phi(R)^{ab} \rightarrow D$ and hence, by Th. 9.3, a homomorphism

$$\mu: K_1(U) \rightarrow D(R).$$

The map λ is surjective, by construction, hence so is (12). We claim that the kernel is $GL(R)/(GL(R) \cap E(U))$.

Any $A \in GL(R)$ satisfies $v_i(A) = 0$ for all i , by Prop. 9.4, hence $A \in E(U)$. Conversely, if $[A] - [B] = 0$, then $A^2 = B^2$, hence A, B have the same atomic factors, up to order and stable association. Let $A = P_1 \dots P_r$ be a complete factorization and let B be the product (in some order) of $Q_1, \dots, Q_r$ where Q_i is similar to P_i . Replacing A, B by $A \oplus I, B \oplus I$, for I of large enough order, we may assume Q_i to be associated to P_i , say $P_i = U_i Q_i V_i$, with $U_i, V_i \in GL(R)$. Then except for the order of the factors we have

$$A = Q_1 \dots Q_r U_1 \dots U_r V_1 \dots V_r = BF,$$

where $F \in GL(R)$, hence

$$A \equiv BF \pmod{GL(U)},$$

and so $[A] - [B] = [F] \in GL(R) \cdot GL(U)$. Here we may replace $GL(U) \cdot E(U)$ by Whitehead's lemma (cf. Bass [68], p. 226 or Ex. 7 below), and find

$$\ker \mu = GL(R) \cdot E(U)/E(U) \cong GL(R)/(GL(R) \cap E(U)).$$

Moreover, since D is free abelian, μ is split by D over its kernel and so we obtain

THEOREM 9.7. *Let R be a fully atomic semifir with universal field of fractions U and divisor group $D(R)$. Then*

$$K_1(U) \cong U^{ab} \cong D(R) \times [GL(R)/(GL(R) \cap E(U))]. \quad \blacksquare$$

The divisor group inherits a partial ordering from R , defined by writing $\pi > 0$ for $\pi \in D$ whenever π is positive on R . However, this ordering is not enough to define R within U , as is shown by the fact that the determinant of a matrix over R is usually a proper fraction, i.e. has no representative (cf. Ex. 3).

To illustrate Th. 9.7, consider the free algebra $R = k\langle X \rangle$. By Prop. 2.2.3 we have $GL(R) = E(R) \cdot k^\times$, hence

$$GL(R) \cdot E(U)/E(U) \cong E(U) \cdot k^\times / E(U) \cong k^\times / (k^\times \cap E(U)).$$

Now it can be shown (Cohn [82], Révész [83]) that $k^\times \cap E(U)$ is therefore we have for the universal field of fractions U of the free algebra $R = k\langle X \rangle$,

$$K_1(U) \cong D(R) \times k^\times.$$

Let us return to the case of a Sylvester domain and its universal field of fractions U , and examine the case of valuations on U . We recall that a valuation is a function v on U with values in $G \cup \{\infty\}$, where G is an ordered group, taken abelian for simplicity, such that

- v.1. $v(x) = \infty$ if and only if $x = 0$,
- v.2. $v(xy) = v(x) + v(y)$ and
- v.3. $v(x - y) \geq \min\{v(x), v(y)\}$.

Using v.2, we can weaken v.3 to

$$v(p - 1) \geq \min\{v(p), 0\} \quad \text{for any } p \in U; \quad (15)$$

moreover, it is clear that a G -value on R gives rise to a valuation on U , provided that (15) holds. To restate this condition in terms of R , let us take an admissible matrix $A = (A_0, A_*, A_\infty)$ for p , then an admissible matrix for $p - 1$ is $(A_0 + A_\infty, A_*, A_\infty)$, so the condition (15) becomes, after a slight rearrangement,

$$v(A_0 + A_\infty \quad A_*) \geq \min\{v(A_0 \quad A_*), v(A_\infty \quad A_*)\}. \quad (16)$$

Hence v gives rise to a valuation provided that

$$V.7 \quad v(A \nabla B) \geq \min\{v(A), v(B)\},$$

whenever the determinantal sum is defined.

In general this condition need not hold, e.g. in $R = k\langle x, y \rangle$ consider the simple N-value associated with x . We have $v(xy) = v(yx) = 1$, but $v(xy - yx) = 0$. Nevertheless, there is a valuation on the universal field of fractions U that is associated with x . To find it we write U as a skew function field $K(x; \alpha)$, where K is the universal field of fractions of $k\langle y_i | i \in \mathbb{Z} \rangle$ and α is the shift automorphism $y_i \mapsto y_{i+1}$; thus y_i is realized as $x^{-i}yx^i$. On $K(x; \alpha)$ the order of an element in x (i.e. the order of zeros, regarded as a function of x) is the required valuation. In terms of N-values this valuation may be obtained as the sum of certain simple N-values, but this does not seem a very efficient way of constructing the valuation.

Exercises 7.9

1. Verify that $v(I) = 0$ for any function v satisfying V.2, 3.

2. Let K be a field with centre k , such that for any $c \in K$ there exists $a \in K$ centralizing c and transcendental over $k(c)$, and any two elements transcendental over k are conjugate (cf. e.g. Cohn [71] or Ex. 5.19). Show that every element of K^* is a commutator and hence that $K_1(K) = 1$.

3. Let $R = k\langle x, y, z, t \rangle$ be the free algebra and U its universal field of fractions. Show that $\text{Det} \begin{pmatrix} x & y \\ z & t \end{pmatrix} = x(t - zx^{-1}y)$, but that this element of U^{ab} has no representative in R .

4°. In a fully atomic Sylvester domain R define the relation $A \sim B$ between atomic matrices to mean: $v(A) = v(B)$ for all N -values v . Find explicit conditions for $A \sim B$.

5. (Mahdavi-Hezavehi [82]) Let G be an ordered abelian group and define a matrix valuation v on R as a G -value satisfying V.7, defined on a multiplicative set Σ , with $v(A) = \infty$ for $A \notin \Sigma$. Show that any valuation v on a field K may be extended to a matrix valuation V by putting $V(A) = v(\text{Det } A)$. If $f: R \rightarrow S$ is a ring homomorphism and V a matrix valuation on S , show that V^f , defined for $A \in \mathfrak{M}(R)$ by $V^f(A) = V(A^f)$, is a matrix valuation on R .

6. (Mahdavi-Hezavehi [82]) If R is a ring with a matrix valuation V , show that $V^{-1}(\infty)$ is a prime matrix ideal, and if K is the corresponding epic R -field (with singular kernel $V^{-1}(\infty)$), then V induces a valuation v on K . Verify that all valuations on epic R -fields arise in this way. Show that this correspondence between matrix valuations on R and valuations on epic R -fields is bijective, but that different matrix valuations may well define the same valuation on R . (Hint: Consider the trivial valuation on R .)

7. (Whitehead lemma) Show that $A \oplus B \equiv B \oplus A \pmod{\mathbf{E}(R)}$. Deduce that $\text{GL}(R) = \mathbf{E}(R)$.

8*. Let $R = k\langle x, x^{-1}, y \rangle$ and denote its universal field of fractions by U . Writing $A = k\langle y_i | i \in \mathbb{Z} \rangle$, and denoting the shift automorphism $y_i \mapsto y_{i+1}$ by α , verify that $R \cong A[x; \alpha]_{(x)}$. Hence obtain a $\mathbb{Z}$ -valued valuation v on U such that $v(x) = 1$, $v(x^{-1}yx^i) = 0$ ($i \in \mathbb{Z}$). Deduce that in U , x cannot be expressed as a sum of (multiplicative) commutators.

9. Extend Th. 9.2 to pseudo-Sylvester domains.

10. Let R be a persistent semifir over k . What is the relation between the prime divisors of R and those of $R \otimes k(t)$?

7.10 Localization of semifirs and firs

Let R be a ring and Σ a set of square matrices over R . In this section we consider which properties pass from R to R_Σ ; in particular we shall consider the property of being a semifir or a fir.

We begin by examining what restrictions need to be placed on the set Σ . Let $R = k\langle x, y \rangle$ and take $\Sigma = \{xy\}$, or even $\{\overline{xy}\}$, its multiplicative closure. In R_Σ , xy will have an inverse u and we have $xyu = uxy = 1$, but $yux \neq 1$, hence the universal xy -inverting ring contains the idempotent $yux \neq 0, 1$ and so is not even an integral domain. We can exclude rings of this sort by

assuming Σ to be factor-closed, as defined in 7.5. Thus whenever $AB \in \Sigma$ for square matrices A, B , then A is invertible over R_Σ . We remark that B will also be invertible over R_Σ , for AB has an inverse C over R_Σ , hence B has inverse CA : Clearly $CA \cdot B = I$ and $A \cdot BC = I$. Since A^{-1} exists over R_Σ , we must have $A^{-1} = BC$ and so $B \cdot CA = A^{-1}A = I$.

If R_Σ is to be a semifir, there is a further similar condition obtained by taking A, B to be rectangular. Let $A \in {}^rR^n$, $B \in {}^nR^r$; if $AB \in \Sigma$, then AB is invertible over R_Σ , so if R_Σ is to be an Hermite ring, then $r \leq n$ and A must be completable in R_Σ , i.e. there exists $A' \in {}^{n-r}R_\Sigma^r$ such that $\begin{pmatrix} A \\ A' \end{pmatrix}$ is invertible.

A set of square matrices satisfying this condition is called *factor-complete*. This condition is again left-right symmetric, by Th. 0.4.1. Explicitly, if AB has the inverse C over R_Σ and $(A, A')^T$ has inverse (P, P') , then B can be completed to (B, P') , with inverse $(CA, A' - A'BCA)^T$, as is easily checked. From the definition it is clear that any factor-complete set is factor-closed, and the matrices in it are full.

We note that if Σ is factor-complete and $ABC \in \Sigma$, where A is square, then B is completable in R_Σ . For by factor-completeness, AB is completable in R_Σ , say $(AB, P)^T \in \text{GL}(R_\Sigma)$, and A is invertible, therefore so is $(B, P)^T = (A^{-1} \oplus I)(AB, P)^T$.

Our first observation is that these properties are preserved by passing to the multiplicative closure:

LEMMA 10.1. *Let R be a semifir, Σ a set of full matrices and $\bar{\Sigma}$ the set of all matrices of the form*

$$X = P \begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_r \end{pmatrix} Q, \quad P, Q \in \text{GL}(R), A_i \in \Sigma \cup \{1\}. \quad (1)$$

Then $\bar{\Sigma}$ is multiplicative, and if Σ is either factor-closed or factor-complete, then so is $\bar{\Sigma}$.

Proof. It is clear that $\bar{\Sigma}$ is multiplicative. Suppose now that Σ is factor-complete and assume that $AB = X \in \bar{\Sigma}$; we have to show that A or equivalently, B is completable in R_Σ . Thus we have to find B' such that $(B, B') \in \text{GL}(R_\Sigma)$. This comes to the same as showing that $(BQ^{-1}, B') \in \text{GL}(R_\Sigma)$, so we may assume that $P = I = Q$ in (1). We shall use induction on r ; when $r = 1$, then either $AB \in \Sigma$ and the result follows by definition of Σ , or $AB = I$ and then the result follows because R is an Hermite ring. Now let $r > 1$ and write

$$AB = \begin{pmatrix} X_1 & X_2 \\ 0 & X_4 \end{pmatrix}. \quad (2)$$

By the partition lemma there exists $T \in \mathbf{GL}(R)$ such that

$$AT = \begin{pmatrix} A_1 & A_2 \\ 0 & A_4 \end{pmatrix}, \quad T^{-1}B = \begin{pmatrix} B_1 & B_2 \\ 0 & B_4 \end{pmatrix},$$

where $A_1B_1 = X_1$, $A_4B_4 = X_4$. By induction on r there exist B'_1, B'_4 such that $(B_1, B'_1), (B_4, B'_4) \in \mathbf{GL}(R_\Sigma)$, hence we obtain B' such that

$$(B \ B') = T \begin{pmatrix} B_1 & B_2 & B'_1 & 0 \\ 0 & B_4 & 0 & B'_4 \end{pmatrix} \in \mathbf{GL}(R_\Sigma).$$

This shows $\bar{\Sigma}$ to be factor-complete. When Σ is factor-closed, the proof is similar but simpler, since A, B in (2) are now restricted to be square. ■

We shall need two auxiliary results, one on relations in the universal field of fractions, and one on the preservation of linear independence in passing from R to R_Σ .

LEMMA 10.2. *Let R be a semifir and U its universal field of fractions. Given a relation*

$$A = BD^{-1}C, \quad (3)$$

over U , where $D \in R_n$ is full, $A \in {}^mR^p$, $B \in {}^mR^n$ and $C \in {}^nR^p$, then there exist $P, Q \in R_n$, $B' \in {}^mR^n$, $C' \in {}^nR^p$ such that

$$D = QP, \quad B = B'P, \quad C = QC', \quad A = B'C'. \quad (4)$$

Proof. We have the following series of elementary transformations, leaving the inner rank unchanged:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \rightarrow \begin{pmatrix} A & B \\ D^{-1}C & I \end{pmatrix} \rightarrow \begin{pmatrix} 0 & B \\ 0 & I \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix},$$

where we have used (3). The inner rank over U is n , by inspection of the last matrix. Since the embedding $R \rightarrow U$ is rank preserving, by Cor. 5.4.10, the rank is the same in U as in R , hence

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} B' \\ Q \end{pmatrix} (C' \ P),$$

where $P, Q \in R_n$, $B' \in {}^mR^n$, $C' \in {}^nR^p$. On multiplying out we obtain (4). ■

LEMMA 10.3. Let R be a semifir and Σ a factor-complete set of matrices over R . Given any matrix X over R , there exists $T \in \text{GL}(R_\Sigma)$ such that XT is a matrix over R whose non-zero columns are right R_Σ -independent.

Proof. Suppose that the columns of X are right R_Σ -dependent, say

$$Xv = 0, \quad (5)$$

where $0 \neq v \in {}^n R_\Sigma$. Each component v_i of v is a component of the solution of a matrix equation with matrix in Σ . We can combine these equations into one, $Cv^* = b^*$, where C is an $N \times N$ matrix in Σ for some $N \geq n$, $b^* \in {}^N R$, $v^* \in {}^N R_\Sigma$ and the components of v occur among those of v^* . By permuting the columns of C so that the columns relating to v come first, we can write this system as

$$C \begin{pmatrix} v \\ v' \end{pmatrix} = b. \quad (6)$$

The original equation (5) may now be written as

$$(X \ 0) \begin{pmatrix} v & 0 \\ v' & I \end{pmatrix} = 0, \quad (7)$$

where I is the unit matrix of order $N-n$. If in (6) we partition C as $C = (C_1, C_2)$, where $C_1 \in {}^N R^n$, we can write (7) over R_Σ as

$$(X \ 0)C^{-1}(b \ C_2) = 0.$$

This equation still holds over U , so by Lemma 10.2 there exist matrices P, Q, X', V over R , where P, Q are full, such that

$$C = QP, \quad (X \ 0) = X'P, \quad (b \ C_2) = QV, \quad X'V = 0.$$

On replacing P, Q, X', V by $T^{-1}P, QT, X'T, T^{-1}V$, respectively, for suitably chosen $T \in \text{GL}(R)$, we may assume that the first r columns of X' are right R -independent and the rest are 0. Since $v \neq 0$, we have $b \neq 0$, so $V \neq 0$, therefore $r < n$. Writing $X' = (X_1, 0)$, we see that X_1 has fewer columns than X . Now $(X, 0) = (X_1, 0)P$ and the columns of X_1 are right R -independent, so we may write

$$P = \begin{pmatrix} P_1 & 0 \\ P_3 & P_4 \end{pmatrix}, \quad X_1 P_1 = X.$$

Hence

$$Q \begin{pmatrix} P_1 & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ P_3 & P_4 \end{pmatrix} = QP = C \in \Sigma.$$

Since $\bar{\Sigma}$ is factor-complete, $P_1 \otimes I$ is completable in R_{Σ} , hence so is P_1 , say $(P_1, F)^T \in \text{GL}(R_{\Sigma})$. Now $X = (X_1, 0)(P_1, F)^T$ and the result follows by induction on n . ■

We now come to the first main result of this section, giving conditions for the localization R_{Σ} to be a semifir.

THEOREM 10.4. *Let R be a semifir, Σ a set of matrices over R and R_{Σ} the universal localization. Then the following are equivalent:*

- (a) R_{Σ} is a semifir,
- (b) R_{Σ} is an Hermite ring,
- (c) Σ is factor-complete.

Proof. (a) $\Rightarrow$ (b) $\Rightarrow$ (c). This is clear by Th. 1.1.1 and the remarks before Lemma 10.1. To prove (c) $\Rightarrow$ (a), let $u \in R_{\Sigma}^n$ and consider the right ideal generated by the components of u . If the components are right linearly dependent, there exists $v \in {}^n R_{\Sigma}$, $v \neq 0$, such that $uv = 0$. As in the proof of Lemma 10.3 we can write

$$C \begin{pmatrix} v \\ v' \end{pmatrix} = b,$$

where $C \in \bar{\Sigma}$, $b \in {}^N R$. Now the equation $uv = 0$ may be written

$$(u \ 0) \begin{pmatrix} v & 0 \\ v' & I \end{pmatrix} = 0,$$

and we can again partition C as $C = (C_1, C_2)$, $C_1 \in {}^N R^n$, so that

$$(u \ 0)C^{-1}(b \ C_2) = 0.$$

By the left-right dual of Lemma 10.3 we can transform (b, C_2) by an invertible matrix over R_{Σ} to a matrix over R with r rows that are left R_{Σ} -independent, followed by 0's, say $T(b, C_2)$. Then $(u, 0)C^{-1}T^{-1}$ has at most $N - r$ non-zero columns and since $r \geq \rho(b, C_2) = N - n + 1$, we have at most $N - (N - n + 1) = n - 1$ non-zero columns. Thus we have a generating set of less than n elements for our right ideal. It follows that every finitely generated right ideal of R_{Σ} is free of unique rank, hence R_{Σ} is a semifir and (a) holds. ■

If the set Σ is not factor-complete, we cannot expect an Hermite ring, but as we shall see, the localization is a pseudo-Sylvester domain provided that Σ is factor-closed. We shall also need to assume Σ multiplicative, since Lemma 10.1 is no longer available now.

THEOREM 10.5. *Let R be an S -ring and let Σ be a multiplicative set of stably full matrices over R . Then the localization R_Σ is again an S -ring. Moreover, (i) if Σ is factor-closed, then R_Σ is weakly finite, hence a pseudo-Sylvester domain and (ii) if Σ is factor-complete, then R_Σ is a Sylvester domain.*

Proof. Let P be a stably full $n \times n$ matrix over R_Σ ; by Cramer's rule, applied to P , we have an equation

$$(A_* \quad -A_0) = (A_* \quad A_\infty) \begin{pmatrix} I & Q \\ 0 & P \end{pmatrix}, \quad (8)$$

where Q is over R_Σ , the A 's are over R and $(A_*, A_\infty) \in \Sigma$. Since P is stably full, so is $\begin{pmatrix} I & Q \\ 0 & P \end{pmatrix}$ and (A_*, A_∞) is invertible over R_Σ , so $(A_*, -A_0)$ is stably full over R_Σ , hence stably full over R and so invertible over U , the universal R -field. It follows that P is invertible over U , hence the map $R_\Sigma \rightarrow U$ is stably honest, therefore by Th. 5.13, R_Σ is an S -ring, as claimed.

Let us now return to (8) and suppose that P is an $n \times n$ matrix over R_Σ of stable rank r . Suppose that A_* has t columns, and so $t+n$ rows, since (A_*, A_∞) is square. By increasing t if necessary we may enlarge the unit factor on the right of (8) and thus we may assume that the rank of $\begin{pmatrix} I & Q \\ 0 & P \end{pmatrix}$ is stabilized at $t+r$. Hence $(A_*, -A_0)$ has the same inner rank and we have an equation

$$(A_* \quad -A_0) = B(C_* \quad C_0), \quad (9)$$

over R , where B has $t+r$ columns. Therefore

$$(A_* \quad A_\infty) = (BC_* \quad A_\infty) = (B \quad A_\infty) \begin{pmatrix} C_* & 0 \\ 0 & I \end{pmatrix}. \quad (10)$$

Further, $(A_*, -A_0) = (B, A_\infty) \begin{pmatrix} C_* & C_0 \\ 0 & 0 \end{pmatrix}$, so (8) now becomes

$$(B \quad A_\infty) \begin{pmatrix} C_* & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & Q \\ 0 & P \end{pmatrix} = (B \quad A_\infty) \begin{pmatrix} C_* & C_0 \\ 0 & 0 \end{pmatrix}.$$

We thus obtain

$$(B \quad A_\infty) \begin{pmatrix} C_* Q - C_0 \\ P \end{pmatrix} = 0, \quad (11)$$

where (B, A_∞) is $(t+n) \times (t+r+n)$.

We consider the two cases separately. (i) Σ is factor-closed. Let P be a square matrix of stable rank $r = 0$, then (B, A_∞) is a square left factor of

(A_*, A_∞) , by (10), and hence is invertible in R_Σ . Cancelling this factor from (11), we find that $P = 0$, so R_Σ is weakly finite, by Prop. 5.6.1. Thus R_Σ is a pseudo-Sylvester domain, as claimed.

(ii) Σ is factor-complete. Then the pair of factors in (10) is completable in R_Σ ; if

$$\begin{pmatrix} B & A_\infty \\ X & Y \end{pmatrix} \text{ and } \begin{pmatrix} S & T \\ U & V \end{pmatrix}$$

are mutually inverse, then by (11) we have $P = VZ$ for some matrix Z . Here the number of columns of V is the number of rows of X , which is the index of (B, A_∞) , i.e. $(t+r+n) - (t+n) = r$. Thus $\rho P \leq r = \rho^* P$, and this shows that the stable rank and inner rank of P over R_Σ are the same, therefore by Prop. 5.6.2 and 3, R_Σ is a Sylvester domain. ■

Finally, to obtain a localization theorem for firs, we shall use the following theorem of Bergman and Dicks:

THEOREM 10.6. *Let R be a left hereditary ring and Σ any set of square matrices. Then the universal localization R_Σ is again left hereditary.*

Proof. Let us write $S = R_\Sigma$; given $M \in {}_R\mathcal{M}$, $A \in R_n$, we have an abelian group endomorphism θ_A of ${}^nM$ given by

$$\theta_A : u \mapsto Au \quad (u \in {}^nM). \quad (12)$$

Let $\mathcal{M}^\Sigma$ be the full subcategory of ${}_R\mathcal{M}$ consisting of those modules M for which (12) is an automorphism for all $A \in \Sigma$. Then it is clear that $\mathcal{M}^\Sigma$ is equivalent to ${}_S\mathcal{M}$; by the 5-lemma it follows that $\mathcal{M}^\Sigma$ is closed under extensions in ${}_R\mathcal{M}$.

Now S is an epic R -ring, hence for $M \in \mathcal{M}^\Sigma$, $\text{Hom}_R(S, M) \cong M$ and it follows that $\text{Ext}_R^1(S, -)$ vanishes on $\mathcal{M}^\Sigma$. Given $C \in {}_S\mathcal{M}$, we have an injective resolution of C as R -module:

$$0 \rightarrow C \rightarrow I_0 \rightarrow I_1 \rightarrow 0.$$

This is of length at most 1 because R is left hereditary. Let us apply $\text{Hom}_R(S, -)$ and recall that if $\lambda: R \rightarrow S$ is the canonical map, then $\text{Hom}_R(S, A) = A^\lambda$ is the coinduced module; by Appendix, p. 552, this is S -injective whenever A is R -injective, while $C^\lambda = C$, as we have seen:

$$0 \rightarrow C \rightarrow I_0^\lambda \rightarrow I_1^\lambda \rightarrow 0.$$

Thus C has an injective resolution of length at most 1, and so $\text{gl. dim. } S \leq 1$ and S is left hereditary, as claimed. ■

As a consequence we have

THEOREM 10.7. *Let R be a fir and Σ a factor-complete set of matrices. Then R_Σ is a fir.*

For R_Σ is a semifir by Th. 10.4 and hereditary by Th. 10.6, therefore, by Cor. 1.4.2, R_Σ is a fir. ■

This shows $k_{\text{rat}}\langle X \rangle$ to be a fir.

Let us now examine how the divisor group of a fir or, more generally, a fully atomic semifir behaves under localization. We shall assume that R is a fully atomic semifir and Σ a factor-complete set of full matrices. An atom A over R and also the associated simple N-value is called Σ -irrelevant if A becomes a unit in R_Σ and Σ -relevant otherwise.

PROPOSITION 10.8. *Let R be a fully atomic semifir and Σ a factor-complete set of matrices over R . Then the localization R_Σ is again a fully atomic semifir and every atom over R either becomes a unit or remains an atom over R_Σ .*

Proof. We begin by proving the last part. Let A be an atom over R and suppose that over R_Σ we have $A = B_1 B_2$, where the B_i are square non-units. By Cramer's rule we have $U_i(B_i \oplus I)V_i = C_i$, $i = 1, 2$, where C_i is a matrix over R and $U_i, V_i \in \text{GL}(R_\Sigma)$. Hence

$$A \oplus I = U_1^{-1} C_1 V_1^{-1} U_2^{-1} C_2 V_2^{-1}. \quad (13)$$

Let v be the simple N-value defined by A , take complete factorizations of C_1, C_2 over R and let w_1, w_2 be the N-values corresponding to C_1, C_2 but counting only Σ -relevant atoms. Then by (13), $v = w_1 + w_2$, but $w_i(C_i) \geq 1$ and so

$$1 = v(A) = w_1(C_1) + w_2(C_2) \geq 2,$$

a contradiction; this shows A to be an atom or a unit over R_Σ .

Now let P be any full matrix over R_Σ and write

$$U(P \oplus I)V = A, \quad (14)$$

where $A \in \mathfrak{M}(R)$, $U, V \in \text{GL}(R_\Sigma)$. We can write A as a product of r atoms, say, over R ; each will be either an atom or a unit over R_Σ , hence P can be written as a product of at most r atoms over R_Σ and this shows R_Σ to be fully atomic; it is a semifir by Th. 10.4. ■

The fact that R_Σ is fully atomic may also be proved as follows. Denote by

w the sum of all Σ -relevant simple N -values on R ; then w is an N -value on R_Σ and by Cramer's rule, for any full matrix A over R_Σ , $w(A) = 0$ if and only if A is invertible over R_Σ . Hence we can apply the criterion of Prop. 9.6.

For any fully atomic semifir R , the divisor group may be defined as in 7.9, and if Σ is a factor-complete set of full matrices over R , then the divisor group can again be defined for R_Σ , by Prop. 10.8. The mapping between these divisor groups may be described as follows:

PROPOSITION 10.9. *Let R be a fully atomic semifir and Σ a factor-complete set of matrices over R , so that R_Σ is again fully atomic. Then (i) any two atoms over R that are not stably associated over R are not stably associated over R_Σ , unless both become units and (ii) every matrix P over R_Σ is stably associated to the image of a matrix P' over R , and if P is an atom, then so is P' .*

Proof. (i) Let A, A' be atoms over R , not stably associated over R , and suppose that A is Σ -relevant. Let v be the simple N -value corresponding to A , then v is an N -value on R_Σ and $v(A) = 1$, $v(A') = 0$, hence A, A' cannot be stably associated over R_Σ .

(ii) Let P be a matrix over R_Σ ; then we have an equation (14), hence P is stably associated to a matrix A over R . Now suppose that P is an atom and denote by w the sum of all Σ -relevant N -values on R ; then w is an N -value on R_Σ . Since P is an atom, we have $1 = w(P) = w(A)$; this means that in a complete factorization of A over R there is only one factor, P' say, which is Σ -relevant, and clearly P is stably associated over R_Σ to P' . ■

Let A be an atom over R and denote by $[A]_R$ the corresponding prime of R ; if A is Σ -relevant, it remains an atom over R_Σ and so defines a prime $[A]_\Sigma$ there. It is clear that stably associated atoms over R remain stably associated over R_Σ , hence the correspondence $[A]_R \mapsto [A]_\Sigma$ defines a homomorphism

$$\phi: \mathbf{D}(R) \rightarrow \mathbf{D}(R_\Sigma). \quad (15)$$

Let $\mathbf{D}_\Sigma(R)$ be the subgroup of $\mathbf{D}(R)$ generated by the Σ -relevant primes; we claim that $\mathbf{D}_\Sigma(R) \cong \mathbf{D}(R_\Sigma)$. For the restriction of ϕ to $\mathbf{D}_\Sigma(R)$ is injective, by Prop. 10.9(i) and surjective by (ii). Thus we have proved

THEOREM 10.10 *Let R be a fully atomic semifir, Σ a factor-complete set of matrices over R and denote by $\mathbf{D}_\Sigma(R)$ the subgroup of $\mathbf{D}(R)$ generated by the Σ -relevant primes of R . Then the embedding $R \rightarrow R_\Sigma$ induces an isomorphism*

$$\mathbf{D}_\Sigma(R) \cong \mathbf{D}(R_\Sigma).$$

Moreover, if (15) is the induced homomorphism, then

$$\mathbf{D}(R) = \mathbf{D}_{\Sigma}(R) \times \ker \phi, \quad (16)$$

where $\ker \phi$ is the subgroup of $\mathbf{D}(R)$ generated by the Σ -irrelevant primes.

Here (16) follows because $\mathbf{D}_{\Sigma}(R)$ is free and so provides a splitting; the rest is clear. ■

Exercises 7.10

1. Let R be a semifir with universal field of fractions U . If $a_i a_0^{-1} = b_i b_0^{-1}$ in U (for $a_i, b_i \in R$), find $a, b, c_i \in R$ such that $a_i = c_i a, b_i = c_i b$. Prove a corresponding result for matrices over R .

2. Verify that the group algebra of a free group is a fir.

3. (Schofield) Let R be a semifir and Σ the set of all full $n \times n$ matrices, for some $n \geq 1$. Show that Σ is factor-complete; deduce that R_{Σ} is a semifir. Show that the chain of rings obtained by iterating this process has as its union the universal field of fractions U of R . Verify that different values of n give cofinal chains.

4. Let R be a k -algebra with $2n(n+1)$ generators arranged as an $n \times (n+1)$ matrix A and an $(n+1) \times n$ matrix B , with defining relation $AB = I$. Show that R is an n -fir, whose universal localization at the set of all full $n \times n$ matrices is not an n -fir.

5. (Schofield) Let R be a k -algebra on 18 generators arranged as two 3×3 matrices A, B with defining relations $AB = 0$. Show that R is a 2-fir but $R_{R^{\times}}$ is not an integral domain. (Hint: Apply elementary operations to A, B to reduce some of the entries to 0.)

6. Let R be a ring and $\Sigma \subseteq \mathfrak{M}(R)$. Show that if R_{Σ} is a local ring, then Σ is factor-complete.

7. Let R be a semifir and $\Sigma \subseteq \mathfrak{M}(R)$. Show that R_{Σ} is a semifir if and only if R_{Σ} is weakly finite and whenever $AB \in \Sigma$, where $A \in {}^m R^n, B \in {}^n R^m, m \leq n$, then $I_n - B(AB)^{-1}A$ has inner rank $n - m$ over R_{Σ} .

8. Show that Th. 10.6 is still valid if Σ contains rectangular matrices. Deduce that the standard non-IBN ring $V_{m,n}$ is hereditary. Show also that $V_{n,n}$ is a fir.

9⁰. For a free algebra of infinite rank let Σ_0 be a finite set of matrices which is factor-closed. Find a finite set Σ_1 containing Σ_0 and factor-complete.

7.11 Universal localization: another route

Let R be a ring and $\Sigma \subseteq \mathfrak{M}(R)$; we have seen in 7.2 that the universal localization R_{Σ} is obtained by adjoining formal inverses of the matrices in Σ ,

and the elements of R_Σ are obtained as components of the solutions of matrix equations $Ax = b$. However, this mode of construction sheds no light on the kernel of the canonical map $\lambda: R \rightarrow R_\Sigma$. We shall now give a second method of constructing R_Σ , as the set of elements $bA^{-1}c$, or rather $a - bA^{-1}c$, for $A \in \Sigma$, but constructing all matrices at once, not merely the elements of R_Σ . This will enable us to determine $\ker \lambda$ and also to prove the theorem of Gerasimov and Malcolmson on the localization of n -firs.

We begin by defining the category $\mathcal{M}$ of matrix objects. An object in $\mathcal{M}$, also called *matrix object*, M , is a family of sets (${}^mM^n$) ($m, n \in \mathbb{N}$) with distinguished elements $0 \in {}^mM^n$, $1_r \in {}^rM^r$ and binary operations

$$\begin{aligned} {}^mM^n \times {}^mM^n &\rightarrow {}^mM^n, & (A, B) &\mapsto A \oplus B, \\ {}^mM^r \times {}^rM^n &\rightarrow {}^mM^n, & (A, B) &\mapsto A \odot B. \end{aligned}$$

We shall call ${}^mM^n$ the (m, n) -component of M ; it need not be empty even if m or n is zero. If each ${}^mM^n$ is an abelian group with respect to $\oplus$ with 0 as zero, $\odot$ is associative, with 1 as left and right neutrals, i.e. $1_m \odot A = A \odot 1_n = A$ for $A \in {}^mM^n$, and $\odot$ is distributive over $\oplus$, then M is said to be *ringlike*. In particular, ${}^mM^n$ is then a ring. For example, if R is any ring, then the set of all matrices over R is a matrix object $\text{Mat}(R)$, with (m, n) -component ${}^mR^n$, the set of all $m \times n$ matrices over R , with the usual addition and multiplication. It is clear that $\text{Mat}(R)$ is ringlike. We remark that a ringlike matrix object is essentially a preadditive category with $\mathbb{N}$ as object set.

Given two matrix objects M_1, M_2 , a morphism $f: M_1 \rightarrow M_2$ is a family of maps ${}^mf^n: {}^mM_1^n \rightarrow {}^mM_2^n$ which preserves the constants 0, 1 and the operations $\oplus, \odot$ whenever defined. For example, any ring homomorphism $f: R \rightarrow R'$ defines a morphism $\text{Mat}(f): \text{Mat}(R) \rightarrow \text{Mat}(R')$.

A *congruence* θ on a matrix object M is an equivalence on M which is compatible with $\oplus, \odot$ and such that elements from different components ${}^mM^n$ are inequivalent. Thus

- (i) if $A \theta A', A \in {}^mM^n$, then $A' \in {}^mM^n$ and
- (ii) if $A \theta A', B \theta B'$, then $A \oplus B \theta A' \oplus B'$ whenever either side is defined, and $A \odot B \theta A' \odot B'$, whenever either side is defined.

For example, the kernel of any morphism is a congruence. Conversely, given a congruence θ on M , we can define a matrix object structure in canonical fashion on the quotient M/θ . If the quotient is ringlike, we shall also say that the congruence θ is ringlike. Since a ringlike matrix object may be defined by identities, there is always a least ringlike congruence on any matrix object, leading to a universal ringlike object.

With each ring R we can associate a category $(R; \mathcal{M})$, usually called the comma category based on $\text{Mat}(R)$. An object of $(R; \mathcal{M})$, also called an R -

object, is a matrix object M together with a fixed morphism $\varepsilon: \text{Mat}(R) \rightarrow M$, and a morphism between R -objects (M_i, ε_i) ($i = 1, 2$) is a morphism $f: M_1 \rightarrow M_2$ such that the triangle shown commutes. In particular, if R' is any R -ring, then $\text{Mat}(R')$ becomes an R -object via the map $R \rightarrow R'$.

$$\begin{array}{ccc} & \text{Mat}(R) & \\ \varepsilon_1 \swarrow & & \searrow \varepsilon_2 \\ M_1 & \xrightarrow{\quad} & M_2 \end{array}$$

We observe that ringlike R -objects always arise from rings:

PROPOSITION 11.1. *Let R be a ring. If M is a ringlike R -object, then ${}^1M^1$ is an R -ring and $M \cong \text{Mat}({}^1M^1)$.*

Proof. It is clear that $N = {}^1M^1$ is a ring and ${}^1\varepsilon^1: R \rightarrow N$ is a ring homomorphism; thus we have morphisms from $\text{Mat}(R)$ to $\text{Mat}(N)$ and to M . Let ${}^mE_{ij}$ be the image of the matrix unit e_{ij} from ${}^mR^n$ in ${}^mM^n$. Given $A \in {}^mM^n$, write $a_{ij} = {}^1E_{1i} A {}^1E_{1j}$; then $a_{ij} \in {}^1M^1$, hence $(a_{ij}) \in {}^mN^n$. In the other direction, if $(a_{ij}) \in {}^mN^n$, we define $A = \sum {}^1E_{1i} a_{ij} {}^1E_{1j} \in {}^mM^n$; it is easily checked that these maps are mutually inverse homomorphisms. ■

This result shows that a ringlike R -object is completely determined by its $(1, 1)$ -component.

Let (M, ε) be an R -object and $\Sigma \subseteq \text{Mat}(R)$; a congruence θ on M is said to be (universal) Σ -inverting if θ is ringlike and the ring homomorphism ${}^1\varepsilon^1/\theta: R \rightarrow {}^1M^1/\theta$ is (universal) Σ -inverting. Our aim will be to construct an R -object $(M(R, \Sigma), \varepsilon)$ such that for any Σ -inverting homomorphism $f: R \rightarrow R'$, $\text{Mat}(f)$ can be factored uniquely by ε . It then follows that the universal ringlike quotient of $M(R, \Sigma)$ has the desired localization R_Σ as $(1, 1)$ -component.

Let $\mathfrak{B}(R)$ be the set of all rectangular matrices over R with a block decomposition of the form

$$A = \begin{pmatrix} A' & \tilde{A} \\ A^0 & A \end{pmatrix}; \quad (1)$$

thus we consider quadruples $(A', \tilde{A}, A^0, A)$ whose dimensions are such as to fit together so as to form a matrix as in (1). Two such matrices are considered equal only if they have the same block decomposition (as well as the same entries). Thus we have four mappings $A \mapsto A'$, etc., from $\mathfrak{B}(R)$ to $\text{Mat}(R)$. Any one of the four blocks may be absent, if certain dimensions are 0, and we indicate this by writing $(\underline{\tilde{A}})$, $(\overline{A^0} \mid A)$, etc.

With each ring R we associate an R -object $M(R)$ by putting

$${}^m M^n = \{A \in \mathfrak{B}(R) \mid \tilde{A} \in {}^m R^n\}, \quad (2)$$

with constants $(\underline{0})$, $(\underline{1})$, and operations defined as follows:

$$A \oplus B = \left(\begin{array}{cc|c} A' & B' & \tilde{A} + \tilde{B} \\ \hline A^0 & 0 & 'A \\ 0 & B^0 & 'B \end{array} \right), \quad (3)$$

$$A \odot B = \left(\begin{array}{ccc} A' & \tilde{A} & 0 \\ A^0 & 'A & 0 \\ 0 & 0 & 1 \end{array} \right) \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & B' & \tilde{B} \\ 0 & B^0 & 'B \end{array} \right) = \left(\begin{array}{cc|c} A' & \tilde{A}B' & \tilde{A}\tilde{B} \\ \hline A^0 & 'AB' & 'A\tilde{B} \\ 0 & B^0 & 'B \end{array} \right). \quad (4)$$

It is easily checked that $\oplus$ and $\odot$ are associative, with neutrals $(\underline{0})$, $(\underline{1})$ respectively. The map $\varepsilon: \text{Mat}(R) \rightarrow M(R)$ is defined as

$$A\varepsilon = (\underline{A}); \quad (5)$$

with this definition $M(R)$ becomes an R -object, as is clear from (3) and (4). We also remark that $A \mapsto \tilde{A}$ is a homomorphism from $M(R)$ to $\text{Mat}(R)$ and that $\varepsilon \cdot \sim = 1$. Further we shall put

$$\ominus A = (\underline{-1}) \odot A = \left(\begin{array}{c|c} -A' & -\tilde{A} \\ \hline A^0 & 'A \end{array} \right) \quad (6)$$

and $A \ominus B = A \oplus (\ominus B)$.

We shall need two further mappings. In the first place we shall define

$$A^* = \left(\begin{array}{c|c} -1 & 0 \\ \hline A & 1 \end{array} \right) \in {}^m M^n, \quad \text{if } A \in {}^m R^n; \quad (7)$$

this will be used later in the construction of the inverse. Secondly we define a map from $M(R)$ to $\text{Mat}(R)$ by writing

$$\hat{A} = \left(\begin{array}{ccc} 1 & A' & \tilde{A} \\ 0 & A^0 & 'A \\ 0 & 0 & 1 \end{array} \right).$$

In particular, for $P \in {}^m M^0$, $Q \in {}^0 M^n$, we have

$$\hat{P} = \begin{pmatrix} 1_m & P' \\ 0 & P^0 \end{pmatrix}, \quad \hat{Q} = \begin{pmatrix} Q^0 & 'Q \\ 0 & 1_n \end{pmatrix}.$$

We note that for any $A \in {}^m M^n$, if A^0 is invertible, then so is $\hat{A}$ and there is a unique $\omega(A) \in {}^m M^n$ such that

$$(\hat{A})^{-1} = \widehat{-\omega(A)}. \quad (8)$$

Explicitly, if $(A^0)^{-1} = \alpha$, then

$$\omega(A) = \left(\begin{array}{c|c} A'\alpha & \tilde{A} - A'\alpha'A \\ \hline -\alpha & \alpha'A \end{array} \right), \quad (9)$$

as is easily checked. From (8) it is clear that $\omega^2 = 1$.

Let Σ be an upper multiplicative set of matrices over R . We define a sub-object of $M(R)$ by putting

$$M(R, \Sigma) = \{A \in M(R) \mid A^0 \in \Sigma\}. \quad (10)$$

From (3) and (4) it is clear that $M(R, \Sigma)$ is closed under $\oplus, \odot$ and it contains the neutrals $0, 1$. Further, by (5), ε maps $\text{Mat}(R)$ into $M(R, \Sigma)$, so that $M(R, \Sigma)$ is an R -object, and $\wedge$ maps $M(R, \Sigma)$ into $\text{GL}(R_\Sigma)$, so that (8) can be expressed as a commutative diagram:

$$\begin{array}{ccc} M(R, \Sigma) & \xrightarrow{\wedge} & \text{GL}(R_\Sigma) \\ \downarrow -\omega & & \downarrow -1 \\ M(R, \Sigma) & \xrightarrow{\wedge} & \text{GL}(R_\Sigma) \end{array}$$

LEMMA 11.2. *Let R be a ring and Σ an upper multiplicative set of invertible matrices. Then the map ω defined by (8) is an R -morphism of $M(R, \Sigma)$.*

Proof. We have to verify that

$$\omega(A \oplus B) = \omega(A) \oplus \omega(B), \quad \omega(A \odot B) = \omega(A) \odot \omega(B),$$

whenever either side is defined, and $\omega(\lfloor \underline{A} \rfloor) = (\lfloor \underline{A} \rfloor)$. This is a routine calculation, which may be left to the reader. ■

As a consequence we see that the map $\tilde{\omega}: M(R, \Sigma) \rightarrow \text{Mat}(R)$, given by $A \mapsto \tilde{\omega}(A)$ is an R -morphism. This allows us to show that $M(R, \Sigma)$ is an R -object with the desired properties:

THEOREM 11.3. *Let R be any ring and Σ an upper multiplicative set of matrices over R . Then the R -object $M(R, \Sigma)$ defined in (10) has the following properties.*

(i) *For any Σ -inverting homomorphism $f: R \rightarrow R'$ there exists an R -morphism $h_f: M(R, \Sigma) \rightarrow \text{Mat}(R')$ of R -objects.*

(ii) *For any matrix $A \in \Sigma \cap {}^m R^n$ there exists $A^* \in {}^n M^m(R, \Sigma)$ such that $A^* h_f = (A f)^{-1}$ for any Σ -inverting homomorphism f .*

(iii) *$M(R, \Sigma)$ is generated as R -object by $\{A^* \mid A \in \Sigma\}$.*

Proof. We have the diagram shown, where $\Sigma' = \Sigma f$ and $h_f = M(f) \cdot \tilde{\omega}$. Here $\tilde{\omega}$ is given explicitly by

$$\tilde{\omega}(A) = \widetilde{\omega(A)} = \tilde{A} - A'(A^0)^{-1}A, \quad (11)$$

and this is defined whenever A^0 is invertible. In particular, for a Σ -inverting homomorphism f , $\tilde{\omega}(Af)$ is defined for any $A \in M(R, \Sigma)$. By the above remark $\tilde{\omega}$ is an R -morphism, hence so is h_f .

$$\begin{array}{ccc} \text{Mat}(R) & \xrightarrow{\varepsilon} & M(R, \Sigma) \\ & \searrow \text{Mat}(f) & \downarrow h_f \\ & & \text{Mat}(R') \end{array} \quad \begin{array}{c} \\ \\ \xleftarrow{M(f)} \end{array} \begin{array}{c} \\ \\ M(R', \Sigma') \end{array}$$

If $A \in \Sigma \cap {}^m R^n$, then $A^* \in {}^n M^m(R, \Sigma)$ is defined by (7) and satisfies

$$A^* M(f) = (Af)^*. \quad (12)$$

We therefore have

$$A^* \cdot h_f = \tilde{\omega}(A^* \cdot M(f)) = \tilde{\omega}((Af)^*) = \widetilde{\omega((Af)^*)} = (Af)^{-1}.$$

Finally, (iii) follows from the identity

$$\left(\begin{array}{c|c} A' & \tilde{A} \\ \hline A^0 & A \end{array} \right) = (\underline{\tilde{A}}) \oplus \{(\underline{-A'}) \odot (A^0)^* \odot (\underline{A})\}. \quad \blacksquare$$

For each Σ -inverting homomorphism $f: R \rightarrow R'$, $\ker h_f$ is a Σ -inverting congruence on $M(R, \Sigma)$ and every Σ -inverting congruence occurs in this way, by Prop. 11.1, therefore we have

COROLLARY 11.4. *With the notation of Th. 11.3, let (h_f) be the family of all Σ -inverting congruences. Then the universal Σ -inverting congruence θ_Σ on $M(R, \Sigma)$ is given by*

$$\theta_\Sigma = \bigcap \ker h_f. \quad \blacksquare$$

Next we shall examine θ_Σ more closely. Consider the universal Σ -inverting homomorphism $\lambda_\Sigma: R \rightarrow R_\Sigma$; if Σ contains every matrix A such that $A\lambda_\Sigma$ is invertible, Σ is said to be *saturated*. Any set Σ has a unique *saturation* $\bar{\Sigma} \supseteq \Sigma$, consisting of all matrices inverted over R_Σ . Thus $R_{\bar{\Sigma}} \cong R_\Sigma$ and the universal Σ - and $\bar{\Sigma}$ -inverting maps coincide.

A matrix $C \in M(R, \Sigma)$ is said to be Σ -*nonfull* if it can be written

$$\left(\begin{array}{c|c} C' & \tilde{C} \\ \hline C^0 & C \end{array} \right) = \left(\begin{array}{c|c} A' & \\ \hline A^0 & \end{array} \right) (\overline{B^0 | B}) = \left(\begin{array}{c|c} A'B^0 & A'B \\ \hline A^0 B^0 & A^0 B \end{array} \right), \quad A^0, B^0 \in \bar{\Sigma}, \quad (13)$$

where $\bar{\Sigma}$ is the saturation of Σ . We observe that if C is Σ -nonfull, then $\tilde{\omega}(C) = 0$. For if C is given by (13), then

$$\tilde{\omega}(C) = A''B - A'B^0(A^0B^0)^{-1}A^0B = 0.$$

Our next aim will be to describe θ_{Σ} by explicit relations; we begin by introducing a relation $\mathcal{R}$ between elements of $M(R, \Sigma)$. Given $A, B \in {}^mM^n(R, \Sigma)$, we shall write $A \rightarrow B$ or $A \mathcal{R} B$ and say A transforms to B if there exist $P \in {}^mM^0(R)$, $Q \in {}^0M^n(R)$ such that

$$\hat{P}A = B\hat{Q}. \quad (14)$$

In detail we have

$$\begin{pmatrix} 1 & P' \\ 0 & P^0 \end{pmatrix} \begin{pmatrix} A' & \tilde{A} \\ A^0 & 'A \end{pmatrix} = \begin{pmatrix} B' & \tilde{B} \\ B^0 & 'B \end{pmatrix} \begin{pmatrix} Q^0 & 'Q \\ 0 & 1 \end{pmatrix},$$

or written out in full:

$$\begin{aligned} A' + P'A^0 &= B'Q^0, & \tilde{A} + P' 'A &= B' 'Q + \tilde{B}, \\ P^0A^0 &= B^0Q^0, & P^0 'A &= B^0 'Q + 'B. \end{aligned}$$

These equations may be stated in a number of ways, which are listed in

LEMMA 11.5. *For any $A, B \in M(R, \Sigma)$ the following are equivalent:*

- (a) $\hat{P}A = B\hat{Q}$, thus $A \mathcal{R} B$,
- (b) $\begin{pmatrix} B' & P' \\ B^0 & P^0 \end{pmatrix} \begin{pmatrix} -Q^0 & -'Q \\ A^0 & 'A \end{pmatrix} = \begin{pmatrix} -A' & \tilde{B} - \tilde{A} \\ 0 & 'B \end{pmatrix}$,
- (c) $\begin{pmatrix} B' & \tilde{B} \\ B^0 & 'B \end{pmatrix} \ominus \begin{pmatrix} A' & \tilde{A} \\ A^0 & 'A \end{pmatrix} = \left(\begin{array}{cc|c} B' & P' & \\ B^0 & P^0 & \\ 0 & 1 & \end{array} \right) \left(\begin{array}{cc|c} 1 & -Q^0 & -'Q \\ 0 & A^0 & 'A \end{array} \right).$

The proof is straightforward and may be left to the reader. ■

From Lemma 11.5(c) we see that when $A \mathcal{R} B$, then $A \theta_{\Sigma} B$. We claim that θ_{Σ} is actually the equivalence generated by $\mathcal{R}$. In the first place it is clear that $\mathcal{R}$ is reflexive, since $1.A = A.1$; transitivity also follows, for if $\hat{P}A = B\hat{Q}$, $\hat{U}B = C\hat{V}$, then $\hat{U}\hat{P}A = \hat{U}B\hat{Q} = C\hat{V}\hat{Q}$ and $\hat{U}\hat{P} = \hat{Z}$ for suitable Z . Of course $\mathcal{R}$ is not symmetric, but we can show that it is compatible with $\oplus$ and $\odot$. Thus let $\hat{P}A = B\hat{Q}$, $\hat{U}C = D\hat{V}$ and suppose that $A \oplus C$ is defined; then

$$(\widehat{P \oplus U})(A \oplus C) = (B \oplus D)(\widehat{Q \oplus V}),$$

and if $A \odot C$ is defined, then

$$\left[\widehat{\begin{pmatrix} P' & \tilde{B} \\ P^0 & B \end{pmatrix}} \odot U \right] \cdot (A \odot C) = (B \odot D) \cdot \left[Q \odot \widehat{\begin{pmatrix} C' & \tilde{C} \\ V^0 & V \end{pmatrix}} \right].$$

For use in calculations we list some examples of transformations, with the appropriate factors P and Q ; for clarity we often put a dot instead of 0.

1. Omitting an inessential row:

$$\left(\begin{array}{ccc|c} C'_1 & \cdot & C'_3 & \tilde{C} \\ \hline C_{11} & \cdot & C_{13} & 'C_1 \\ C_{21} & C_{22} & C_{23} & 'C_2 \\ C_{31} & \cdot & C_{33} & 'C_3 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} C'_1 & C'_3 & \tilde{C} \\ \hline C_{11} & C_{13} & 'C_1 \\ C_{31} & C_{33} & 'C_3 \end{array} \right),$$

$$P = \left(\begin{array}{ccc|c} \cdot & \cdot & \cdot & \\ \hline 1 & \cdot & \cdot & \\ \cdot & \cdot & 1 & \end{array} \right), Q = \left(\begin{array}{ccc|c} 1 & \cdot & \cdot & \\ \hline \cdot & \cdot & 1 & \\ \cdot & \cdot & \cdot & \end{array} \right).$$

This operation is called *trivial* if the pivot C_{22} is in the north-west corner of C^0 , i.e. $(C_{11}, 0, C_{13}, 'C_1)$ and $(C'_1, C_{11}, C_{21}, C_{31})^T$ are absent, and $C_{22}, C_{33} \in \Sigma$.

2. Inserting an inessential column:

$$\left(\begin{array}{cc|c} C'_1 & C'_3 & \tilde{C} \\ \hline C_{11} & C_{13} & 'C_1 \\ C_{31} & C_{33} & 'C_3 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} C'_1 & C'_2 & C'_3 & \tilde{C} \\ \hline C_{11} & C_{12} & C_{13} & 'C_1 \\ \cdot & C_{22} & \cdot & \cdot \\ C_{31} & C_{32} & C_{33} & 'C_3 \end{array} \right),$$

$$P = \left(\begin{array}{ccc|c} \cdot & \cdot & \cdot & \\ \hline 1 & \cdot & \cdot & \\ \cdot & \cdot & 1 & \end{array} \right), Q = \left(\begin{array}{ccc|c} 1 & \cdot & \cdot & \\ \hline \cdot & \cdot & \cdot & \\ \cdot & 1 & \cdot & \end{array} \right).$$

This operation is called *trivial* if the pivot C_{22} is in the south-east corner of C^0 , i.e. $(C'_3, C_{13}, 0, C_{33})^T$ and $(C_{31}, C_{32}, C_{33}, 'C_3)$ are absent, and $C_{11}, C_{22} \in \Sigma$.

3. Multiplying by an inessential factor on the left:

$$C \rightarrow \hat{U}C, \quad \text{where } U \in {}^m M^0, P = U, Q = (\overline{1} \mid \bar{0}).$$

This operation is called *trivial* if $U^0 \in \Sigma$.

4. Cancelling an inessential factor on the right:

$$C\hat{V} \rightarrow C, \quad \text{where } V \in {}^0 M^n, P = \left(\begin{array}{c|c} 0 & \\ \hline 1 & \end{array} \right), Q = V.$$

This operation is called *trivial* if $V^0 \in \Sigma$.

We remark that the transformation $A \mathcal{R} B$ itself consists of first multiplying by an inessential factor on the left and then cancelling an inessential factor on the right. To give an illustration, we have

$$(\underline{0}) \odot A = \left(\begin{array}{c|c} 0 & 0 \\ \hline A^0 & 'A \end{array} \right),$$

as is easily checked, hence by omitting an inessential row we find that $(\underline{0}) \odot A \mathcal{R} (\underline{0})$. Similarly, by inserting an inessential column we find that $(\underline{0}) \mathcal{R} A \odot (\underline{0})$.

PROPOSITION 11.6. *The equivalence generated by $\mathcal{R}$ is the universal Σ -inverting congruence θ_Σ .*

Proof. Denote by $\sim$ the equivalence generated by $\mathcal{R}$. What has been said shows that $\sim$ is a congruence and the remark following Lemma 11.5 shows that this congruence is contained in θ_Σ , so it will be enough to show that $\sim$ is ringlike and Σ -inverting.

We have seen that $\underline{0}$ and $\underline{1}$ are neutrals for $\oplus, \odot$, respectively and the remark before Prop. 11.6 shows that $\underline{0}$ is a zero for multiplication; we have also seen that $\oplus, \odot$ are associative. To prove the distributive laws, we have

$$\hat{P}[(A \oplus B) \odot C] = [(A \odot C) \oplus (B \odot C)]\hat{Q},$$

where

$$P = \left(\begin{array}{ccc|c} \cdot & \cdot & \cdot & \\ \hline 1 & \cdot & \cdot & \\ \cdot & \cdot & 1 & \\ \cdot & 1 & \cdot & \\ \cdot & \cdot & 1 & \end{array} \right), \quad Q = \left(\begin{array}{ccc|c} \hline 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \end{array} \right),$$

and the other distributive law follows by symmetry; now the commutativity of addition also follows, as in Lemma 4.7. To show that $A \ominus A \sim 0$, we have

$$\left(\begin{array}{cc|c} A' & -A' & \cdot \\ \hline A^0 & \cdot & 'A \\ \cdot & A^0 & 'A \end{array} \right) \mathcal{R} \left(\begin{array}{cc|c} A' & \cdot & \cdot \\ \hline A^0 & A^0 & 'A \\ \cdot & A^0 & 'A \end{array} \right) \mathcal{R} \left(\begin{array}{cc|c} A' & \cdot & \cdot \\ \hline A^0 & \cdot & \cdot \\ \cdot & A^0 & 'A \end{array} \right) \\ \mathcal{R} \left(\begin{array}{c|c} \cdot & \cdot \\ \hline A^0 & 'A \end{array} \right) \mathcal{R}^{-1}(\underline{0}),$$

where, as usual, $\mathcal{R}^{-1}$ denotes the relation inverse to $\mathcal{R}$. It remains to show that $\sim$ is Σ -inverting; let $A \in \Sigma$:

$$\begin{aligned} A^* \odot \underline{A} &= \left(\begin{array}{c|c} -1 & 0 \\ A & A \end{array} \right) \mathcal{R} \left(\begin{array}{c|c} -1 & 1 \\ 1 & 0 \end{array} \right) \mathcal{R} \left(\begin{array}{c|c} 0 & 1 \\ 1 & 0 \end{array} \right) \mathcal{R} \underline{1}, \\ \underline{A} \odot A^* &= \left(\begin{array}{c|c} -A & 0 \\ A & 1 \end{array} \right) \mathcal{R} \left(\begin{array}{c|c} 0 & 1 \\ A & 1 \end{array} \right) \mathcal{R} \left(\begin{array}{c|c} 0 & 1 \\ 1 & 1 \end{array} \right) \mathcal{R}^{-1} \underline{1}. \quad \blacksquare \end{aligned}$$

To obtain a usable criterion for $\sim$ we need to analyse it further. Let us denote by $\mathcal{E}$ the equivalence $\sim$ as a subset of $M(R, \Sigma)^2$. For any subsets $\mathcal{A}, \mathcal{B}$ of $\mathcal{E}$ we define, as usual,

$$\begin{aligned} \mathcal{A} \cdot \mathcal{B} &= \{(x, z) \mid (x, y) \in \mathcal{A} \text{ and } (y, z) \in \mathcal{B} \text{ for some } y\}, \\ \mathcal{A}^{-1} &= \{(y, x) \mid (x, y) \in \mathcal{A}\}, \end{aligned}$$

and consider the following subsets of $\mathcal{E}$:

- $\mathcal{R}_r$: the set of pairs in $\mathcal{R}$ omitting a trivial row,
- $\mathcal{R}_c$: the set of pairs in $\mathcal{R}$ inserting a trivial column,
- $\mathcal{R}_i$: the set of pairs multiplying by a trivial factor on the left,
- $\mathcal{R}_j$: the set of pairs cancelling a trivial factor on the right.

It is clear that $\mathcal{R}_i^{-1} = \mathcal{R}_i$, $\mathcal{R}_j^{-1} = \mathcal{R}_j$, because trivial factors are invertible, and $\mathcal{R}_i \mathcal{R}_i \subseteq \mathcal{R}$, $\mathcal{R}_j \mathcal{R}_j \subseteq \mathcal{R}$. Further we note the following relations.

LEMMA 11.7. (i) $\mathcal{R}_c^{-1} \cdot \mathcal{R} \subseteq \mathcal{R} \cdot \mathcal{R}_c^{-1}$, (ii) $\mathcal{R} \cdot \mathcal{R}_r^{-1} \subseteq \mathcal{R}_r^{-1} \cdot \mathcal{R}$, (iii) $\mathcal{R} \subseteq \mathcal{R}_c \mathcal{R}_i \mathcal{R}_j \mathcal{R}_r$ and (iv) $\mathcal{E} = \mathcal{R}_r^{-1} \mathcal{R} \mathcal{R}_c^{-1}$.

Proof. (i) Suppose that $C \mathcal{R}_c^{-1} A \mathcal{R} B$, say $\hat{P}A = B\hat{Q}$ and

$$C = \left(\begin{array}{cc|c} A' & A_1 & \tilde{A} \\ A^0 & A_2 & 'A \\ \cdot & A_3 & \cdot \end{array} \right).$$

Then $\hat{P}_1 C = D\hat{Q}_1$, where

$$D = \left(\begin{array}{cc|c} B' & B_1 & \tilde{B} \\ B^0 & B_2 & 'B \\ \cdot & B_3 & \cdot \end{array} \right), \quad P_1 = \left(\begin{array}{c|c} P' & \cdot \\ P^0 & \cdot \\ \cdot & 1 \end{array} \right), \quad Q_1 = \left(\begin{array}{cc|c} Q^0 & \cdot & 'Q \\ \cdot & 1 & \cdot \end{array} \right)$$

and $(B_1, B_2, B_3)^T = (A_1 + P'A_2, P^0 A_2, A_3)^T$. Now (ii) follows by symmetry.

(iii) If $A \mathcal{R} B$, say $\hat{P}A = B\hat{Q}$, then $\hat{P}_1 C = D\hat{Q}_1$, where

$$C = \left(\begin{array}{cc|c} A' & B' & \tilde{A} \\ \hline A^0 & . & 'A \\ . & B^0 & . \end{array} \right), \quad D = \left(\begin{array}{cc|c} . & B' & \tilde{B} \\ \hline A^0 & . & 'A \\ . & B^0 & 'B \end{array} \right),$$

$$P_1 = \left(\begin{array}{cc|c} P' & . & \\ \hline 1 & . & \\ P^0 & 1 & \end{array} \right), \quad Q_1 = \left(\begin{array}{cc|c} 1 & . & \\ \hline Q^0 & 1 & 'Q \end{array} \right)$$

and now $A \mathcal{R}_c C$, $D \mathcal{R}_c B$ and $C \mathcal{R}_c \mathcal{R}_j D$.

To prove (iv), put $\mathcal{E}_1 = \mathcal{R}_r^{-1} \mathcal{R} \mathcal{R}_c^{-1}$; clearly $\mathcal{R} \subseteq \mathcal{E}_1 \subseteq \mathcal{E}$; so it will be enough to show that $\mathcal{E}_1$ is an equivalence, and this we can do using only (i)–(iii):

$$\begin{aligned} \mathcal{E}_1 \cdot \mathcal{E}_1 &= \mathcal{R}_r^{-1} \mathcal{R} \mathcal{R}_c^{-1} \mathcal{R}_r^{-1} \mathcal{R} \mathcal{R}_c^{-1} \subseteq \mathcal{R}_r^{-1} \mathcal{R} \mathcal{R}^{-1} \mathcal{R} \mathcal{R}_c^{-1} \\ &\subseteq \mathcal{R}_r^{-1} \mathcal{R} \mathcal{R}_r^{-1} \mathcal{R}_j^{-1} \mathcal{R}_i^{-1} \mathcal{R}_c^{-1} \mathcal{R} \mathcal{R}_c^{-1} \\ &\subseteq \mathcal{R}_r^{-1} \mathcal{R} \mathcal{R}_j \mathcal{R}_i \mathcal{R}_c^{-1} \subseteq \mathcal{R}_r^{-1} \mathcal{R} \mathcal{R}_c^{-1}. \end{aligned}$$

Secondly,

$$\begin{aligned} \mathcal{E}_1^{-1} &= \mathcal{R}_c \mathcal{R}^{-1} \mathcal{R}_r \subseteq \mathcal{R}_c \mathcal{R}_r^{-1} \mathcal{R}_j \mathcal{R}_i \mathcal{R}_c^{-1} \mathcal{R}_r \subseteq \mathcal{R} \mathcal{R}_r^{-1} \mathcal{R}_j \mathcal{R}_i \mathcal{R}_c^{-1} \mathcal{R} \\ &\subseteq \mathcal{R}_r^{-1} \mathcal{R} \mathcal{R}_c^{-1} = \mathcal{E}_1. \quad \blacksquare \end{aligned}$$

Using Lemma 11.7(iv) and Lemma 11.5(c), we obtain the following criterion for the equivalence of matrices:

Two matrices $A, B \in M(R, \Sigma)$ are equivalent (under $\mathcal{E}$) if and only if there is a relation of the form

$$\left(\begin{array}{cc|cc} B' & * & * & * \\ \hline B^0 & * & * & * \\ . & \Sigma & * & * \end{array} \right) \left(\begin{array}{cc|c} * & * & * \\ * & * & * \\ \hline \Sigma & * & * \\ . & A^0 & 'A \end{array} \right) = \left(\begin{array}{ccc} . & -A' & \tilde{B} - \tilde{A} \\ . & . & 'B \\ . & . & . \end{array} \right), \quad (15)$$

where Σ stands for a matrix in Σ and $*$ for an unspecified matrix of the appropriate size.

The results obtained so far may be summed up in a localization theorem.

THEOREM 11.8. *Let R be a ring and Σ an upper multiplicative set of matrices. Define the relation $\sim$ on the set $M(R, \Sigma)$ by writing $A \sim B$ whenever (15) holds. Then ' $\sim$ ' is the universal Σ -inverting congruence on $M(R, \Sigma)$. The class of*

$A \in {}^m M^n(R, \Sigma)$ is represented by a certain $m \times n$ matrix, written $[A]$, and we have

$$[A] + [B] = [A \oplus B], \quad [A] \cdot [B] = [A \odot B],$$

and for $A \in \Sigma$, $[\underline{A}]^{-1} = [A^*]$. ■

From (15) we obtain the following description of the kernel of λ_Σ or, rather, of $\text{Mat}(\lambda_\Sigma)$:

COROLLARY 11.9. *With the notation of Th. 11.8 the following are equivalent:*

- (a) *the matrix A represents 0 in R_Σ , i.e. $[A] = 0$,*
- (b) *there is a relation*

$$\begin{pmatrix} A' & * & * \\ A^0 & * & * \\ \cdot & \Sigma & * \end{pmatrix} \begin{pmatrix} * & * \\ * & * \\ \Sigma & * \end{pmatrix} = \begin{pmatrix} \cdot & \tilde{A} \\ \cdot & 'A \\ \cdot & \cdot \end{pmatrix},$$

- (c) *there is a relation*

$$\begin{pmatrix} * & * & * \\ \Sigma & * & * \end{pmatrix} \begin{pmatrix} * & * & * \\ \Sigma & * & * \\ \cdot & A^0 & 'A \end{pmatrix} = \begin{pmatrix} \cdot & A' & \tilde{A} \\ \cdot & \cdot & \cdot \end{pmatrix}.$$

In particular, the kernel of $\lambda: R \rightarrow R_\Sigma$ is the set of all $c \in R$ such that

$$\begin{pmatrix} \cdot & c \\ \cdot & \cdot \end{pmatrix} = \begin{pmatrix} * & * \\ \Sigma & * \end{pmatrix} \begin{pmatrix} * & * \\ \Sigma & * \end{pmatrix}. \quad \blacksquare \quad (16)$$

The following condition for certain entries of the matrix $[A]$ to vanish is sometimes useful:

COROLLARY 11.10. *Let $A \in M(R, \Sigma)$; if the first row of A is 0, then the first row of $[A]$ is 0; if the last column of A is 0, then the last column of $[A]$ is 0.*

Proof. Suppose that the first row of A is 0. If e_1 denotes a row vector of appropriate size, then

$$[\underline{e}_1] \cdot [A] = [\underline{e}_1 \odot A] = \left[\begin{pmatrix} e_1 A' & e_1 \tilde{A} \\ A^0 & 'A \end{pmatrix} \right] = \left[\begin{pmatrix} \cdot & \cdot \\ A^0 & 'A \end{pmatrix} \right] = 0.$$

The argument for the columns is similar. ■

Next we describe a lifting property for factorizations over $\text{Mat}(R_\Sigma)$, but first we shall need an auxiliary result, on the effect of applying trivial operations to products.

LEMMA 11.11. *Let $A, B, C \in M(R, \Sigma)$.*

(i) *If $C \mathcal{R}_r A \odot B$, then there exist $A_1, B_1 \in M(R, \Sigma)$ such that $A_1 \mathcal{R}_r A$, $B_1 \mathcal{R}_r B$ and $A_1 \odot B_1 \mathcal{R}_r C$.*

(ii) *If $A \odot B \mathcal{R}_c C$, then there exist $A_1, B_1 \in M(R, \Sigma)$ such that $A \mathcal{R}_c A_1$, $B \mathcal{R}_c B_1$ and $C \mathcal{R}_c A_1 \odot B_1$.*

Proof. We have

$$C = \left(\begin{array}{ccc|c} \cdot & A' & \tilde{A}B' & \tilde{A}\tilde{B} \\ C^0 & C_1 & C_2 & C_3 \\ \cdot & A^0 & 'AB' & 'A\tilde{B} \\ \cdot & \cdot & B^0 & 'B \end{array} \right), \quad C^0 \in \Sigma.$$

If we put

$$A_1 = \left(\begin{array}{cc|c} \cdot & A' & \tilde{A} \\ C^0 & C_1 & \cdot \\ \cdot & A^0 & 'A \end{array} \right), \quad B_1 = \left(\begin{array}{cc|c} \cdot & B' & \tilde{B} \\ C^0 & C_2 & C_3 \\ \cdot & B^0 & 'B \end{array} \right),$$

$$P = \left(\begin{array}{cccc|c} \cdot & \cdot & \cdot & \cdot & \\ 1 & \cdot & 1 & \cdot & \\ \cdot & 1 & \cdot & \cdot & \\ \cdot & \cdot & \cdot & 1 & \end{array} \right), \quad Q = \left(\begin{array}{cccc|c} 1 & \cdot & 1 & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot \end{array} \right)$$

then $\hat{P}(A_1 \odot B_1) = C\hat{Q}$ and (i) is proved. Now (ii) follows by symmetry. ■

PROPOSITION 11.12. *Let Σ be an upper multiplicative set of matrices over a ring R . For any relation*

$$UV = [\underline{F}], \quad U, V \in \text{Mat}(R_\Sigma), F \in \text{Mat}(R),$$

there exist $A, B \in M(R, \Sigma)$ such that $[A] = U$, $[B] = V$ and

$$\begin{pmatrix} A' & \tilde{A} & * \\ A^0 & 'A & * \end{pmatrix} \begin{pmatrix} * & * \\ B' & \tilde{B} \\ B^0 & 'B \end{pmatrix} = \begin{pmatrix} 0 & F \\ 0 & 0 \end{pmatrix}. \quad (17)$$

Proof. Let $U = [A_2]$, $V = [B_2]$, where $A_2, B_2 \in M(R, \Sigma)$. By Lemma 11.7(iv) there exist $C, D \in M(R, \Sigma)$ such that $C \mathcal{R}_r A_2 \odot B_2$, $C \mathcal{R}_r D$, $\underline{F} \mathcal{R}_c D$. By Lemma

11.11(i) there exist $A_1, B_1 \in M(R, \Sigma)$ such that $A_1 \mathcal{R} A_2, B_1 \mathcal{R} B_2, A_1 \odot B_1 \mathcal{R} C$, hence by transitivity, $A_1 \odot B_1 \mathcal{R} D$. Now D has the form

$$D = \begin{pmatrix} D' & F \\ D^0 & . \end{pmatrix},$$

so we have $\hat{P}(A_1 \odot B_1) = D\hat{Q}$; in detail, .

$$\begin{pmatrix} 1 & P_{11} & P_{12} \\ . & P_{21} & P_{22} \end{pmatrix} \begin{pmatrix} A'_1 & \tilde{A}_1 & . \\ A^0_1 & 'A_1 & . \\ . & . & 1 \end{pmatrix} \begin{pmatrix} 1 & . & . \\ . & B'_1 & \tilde{B}_1 \\ . & B^0_1 & 'B_1 \end{pmatrix} = \begin{pmatrix} D' & F \\ D^0 & . \end{pmatrix} \begin{pmatrix} Q_1 & Q_2 & Q_3 \\ . & . & 1 \end{pmatrix}.$$

By a calculation we find

$$\begin{pmatrix} D' & \tilde{A}_1 + P_{11}'A_1 & P_{12} \\ D^0 & P_{21}'A_1 & P_{22} \end{pmatrix} \begin{pmatrix} -Q_2 & -Q_3 \\ B'_1 & \tilde{B}_1 \\ B^0_1 & 'B_1 \end{pmatrix} = \begin{pmatrix} . & F \\ . & . \end{pmatrix};$$

now on writing

$$A = \begin{pmatrix} D' & \tilde{A}_1 + P_{11}'A_1 \\ D^0 & P_{21}'A_1 \end{pmatrix}, \quad B = B_1,$$

we see that this has the desired form (17). ■

In order to apply these results to n -firs we need to generalize the notion of a dependence relation. Let Σ be a set of matrices over a ring R . By a Σ -dependence relation $A \cdot B = 0$ we understand a relation of the form

$$\left(\begin{array}{c|c|c} A' & X_1 & U_1 \\ \hline A^0 & X_2 & U_2 \end{array} \right) \begin{pmatrix} V_1 & V_2 \\ Y_1 & Y_2 \\ B^0 & 'B \end{pmatrix} = 0, \quad A^0, B^0 \in \Sigma. \quad (18)$$

Here any of the blocks may be absent. The number of rows in (Y_1, Y_2) is called the *length* of (18).

The relation (18) is said to be *trivial* if it has the form

$$\left(\begin{array}{c|c|c|c} A' & X_{11} & . & . \\ \hline A^0 & X_{21} & . & . \end{array} \right) \begin{pmatrix} . & . \\ Y_{11} & Y_{12} \\ B^0 & 'B \end{pmatrix} = 0.$$

If there exists an invertible matrix T such that $AT^{-1} \cdot TB = 0$ is trivial, we shall say that the original relation (18) is *trivializable*, more precisely, it is trivialized by T^{-1}, T .

Let us write $\Sigma_0 = \Sigma \cup \{1_0\}$; we shall call Σ *n-independent* if any Σ_0 -dependence relation of length $\leq n$ can be trivialized. For example, the rings in which the empty set is *n-independent* are just the *n-firs*. Over an *n-fir* any set consisting of invertible matrices is easily seen to be *n-independent*. If Σ is *n-independent*, then for any $\Sigma' \subseteq \Sigma$, $n' \leq n$, Σ' is *n'-independent*.

A 0-independent set is also called *independent*. We note that an independent set consists of regular elements, because any relations of the form

$$(\overline{A^0} | \overline{\quad}) \left(\begin{array}{c|c} V \\ \hline \hline \hline \end{array} \right) = 0, \quad (\overline{\quad} | \overline{U}) \left(\begin{array}{c|c} \hline \hline \hline \\ \hline B^0 \end{array} \right) = 0$$

must be trivial, and so $V = 0 = U$.

We shall usually want to replace sets of matrices by their upper multiplicative closure.

LEMMA 11.13. *The upper multiplicative closure of an n-independent set of matrices is n-independent.*

Proof. Let Σ be *n-independent*; we may assume that $1_0 \in \Sigma$. Now consider a $\bar{\Sigma}$ -dependence of the form (18), where $\bar{\Sigma}$ is the upper multiplicative closure of Σ :

$$A \cdot B = 0. \quad (19)$$

We shall prove it to be trivializable by induction on the number of rows of A^0 and the number of columns of B^0 . If $A^0, B^0 \in \Sigma$, the result holds by hypothesis, so let $A^0 \notin \Sigma$. The relation (18) may be written in the form

$$\left(\begin{array}{cc|c|c} A_{11} & A_{12} & X_1 & U_1 \\ A_{21} & A_{22} & X_2 & U_2 \\ \cdot & A_{32} & X_3 & U_3 \end{array} \right) \left(\begin{array}{c|c} V_{11} & V_{12} \\ V_{21} & V_{22} \\ Y_1 & Y_2 \\ B^0 & 'B \end{array} \right) = 0, \quad A_{21}, A_{32} \in \bar{\Sigma}. \quad (20)$$

By induction we can trivialize the relation

$$(\overline{A_{32}} | \overline{X_3} | \overline{U_3}) \left(\begin{array}{c|c} V_{21} & V_{22} \\ Y_1 & Y_2 \\ B^0 & 'B \end{array} \right) = 0,$$

and hence we may take it to be trivial, so that (20) becomes

$$\left(\begin{array}{cc|cc|c} A_{11} & A_{12} & X_{11} & X_{12} & U_1 \\ A_{21} & A_{22} & X_{21} & X_{22} & U_2 \\ \cdot & A_{32} & X_{31} & \cdot & \cdot \end{array} \right) \begin{pmatrix} V_{11} & V_{12} \\ \cdot & \cdot \\ \cdot & \cdot \\ Y_{21} & Y_{22} \\ B^0 & 'B \end{pmatrix} = 0. \quad (21)$$

This yields the dependence relation

$$\left(\begin{array}{cc|c} A_{11} & X_{12} & U_1 \\ A_{21} & X_{22} & U_2 \end{array} \right) \begin{pmatrix} V_{11} & V_{12} \\ Y_{21} & Y_{22} \\ B^0 & 'B \end{pmatrix} = 0,$$

which can again be trivialized, by the induction hypothesis. Inserting this in (21), we find that (19) has now been trivialized. ■

There is a useful simplification of the conditions of Cor. 11.9 when Σ is independent:

PROPOSITION 11.14. *Let R be a ring, Σ an independent subset of $\text{Mat}(R)$ and $\bar{\Sigma}$ its saturation. Then the canonical map $\lambda: R \rightarrow R_\Sigma$ is injective and for any $A \in M(R, \Sigma)$ the following are equivalent:*

- (a) *the matrix A represents 0 in R_Σ ,*
- (b) *there is a relation*

$$\begin{pmatrix} A' & \tilde{A} \\ A^0 & 'A \end{pmatrix} = \begin{pmatrix} B' \\ B^0 \end{pmatrix} (C^0, 'C), \quad \text{where } B^0 \in \bar{\Sigma}, \quad (22)$$

- (c) *there is a relation (22), where $C^0 \in \bar{\Sigma}$.*

Proof. By Lemma 11.13 we can take Σ to be multiplicative. If $r \in R$ satisfies $r\lambda = 0$, then by Cor. 11.9,

$$\begin{pmatrix} 0 & r \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} A' & \tilde{A} \\ A^0 & 'A \end{pmatrix} \begin{pmatrix} B' & \tilde{B} \\ B^0 & 'B \end{pmatrix}.$$

By trivializing the relation $(\overline{A^0} | \overline{'A}) (\overline{B'} | \overline{B^0})^T = 0$, we may take $'A = 0$, $B' = 0$, and since A^0, B^0 are regular, it follows that $\tilde{A} = \tilde{B} = 0$, and so $r = 0$, and λ is injective as claimed.

Now (b) $\Rightarrow$ (a), (c) $\Rightarrow$ (a) follow because $\tilde{\omega}(A) = 0$ when A is Σ -nonfull, for we have $A^0 \in \Sigma$ and $A^0 = B^0 C^0$, so if one of B^0, C^0 is in Σ , then the other becomes invertible over R_Σ and so lies in the saturation $\bar{\Sigma}$. To prove

(a) $\Rightarrow$ (b), we have by Cor. 11.9 a relation

$$\begin{pmatrix} A' & P_1 & A_1 \\ A^0 & P_2 & A_2 \\ \cdot & P_3 & A_3 \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ Q_1 & Q_2 \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} \cdot & \tilde{A} \\ \cdot & 'A \\ \cdot & \cdot \end{pmatrix}, \quad P_3, B_{21} \in \Sigma. \quad (23)$$

By hypothesis, the Σ -dependence relation $(\overline{P_3} \parallel \overline{A_3})(\underline{Q_1} \parallel \underline{B_{21}})^T = 0$ can be trivialized, so we may take it to be trivial, i.e. $A_3 = 0$, $Q_1 = 0$. Since P_3 is regular, we have $Q_2 = 0$ and so (23) reduces to

$$\begin{pmatrix} A' & A_1 \\ A^0 & A_2 \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} \cdot & \tilde{A} \\ \cdot & 'A \end{pmatrix}.$$

We thus obtain the relation

$$\left(\begin{array}{c|c} A' & A_1 \\ \hline A^0 & A_2 \end{array} \right) \left(\begin{array}{c} B_{11} \\ \hline B_{21} \end{array} \right) = 0.$$

If this is trivialized by the pair S, T , we have equations of the form

$$\begin{pmatrix} A' & A_1 \\ A^0 & A_2 \end{pmatrix} \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} B' & \cdot \\ B^0 & \cdot \end{pmatrix}, \quad \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} \cdot & \tilde{D} \\ D^0 & 'D \end{pmatrix},$$

and now

$$\begin{pmatrix} A' & A_1 \\ A^0 & A_2 \end{pmatrix} = \begin{pmatrix} B' \\ B^0 \end{pmatrix} (T_{11} \quad T_{12}),$$

and this proves (b). Now (a) $\Rightarrow$ (c) follows similarly. ■

For independent sets of matrices we can also prove a condition somewhat like factor closure, although usually weaker than the latter.

PROPOSITION 11.15. *Let R be any ring and Σ an independent set of matrices. If $A = BC$, where $A, B \in \Sigma$, then there exists an invertible matrix U such that $UC \in \Sigma$.*

Proof. The relation $A = BC$ may be written in the form

$$(B \quad I) \begin{pmatrix} -C \\ A \end{pmatrix} = 0,$$

and this is trivializable, by S, T , say. We obtain

$$(B \quad I)S = (B^0 \quad 0), \quad B^0 \in \Sigma,$$

hence

$$\begin{pmatrix} B & I \\ I & 0 \end{pmatrix} S = \begin{pmatrix} B^0 & . \\ S_{11} & S_{12} \end{pmatrix},$$

and since the left-hand side is invertible, so is the right-hand side. Since B^0 is regular, it must be invertible, and so is S_{12} . Now we have

$$\begin{pmatrix} -C \\ A \end{pmatrix} = S \begin{pmatrix} 0 \\ A^0 \end{pmatrix} = \begin{pmatrix} S_{12} A^0 \\ S_{22} A^0 \end{pmatrix}, \quad A^0 \in \Sigma,$$

therefore $C = -S_{12}A^0$, and the desired conclusion follows. ■

We now come to a basic result establishing a connexion between n -independence and n -firs:

THEOREM 11.16. *Let R be a ring and Σ an n -independent set of matrices over R . Then R_Σ is an n -fir.*

Proof. By Lemma 11.13 we can take Σ to be multiplicative. Let $H \in R_\Sigma^n$, $K \in {}^n R_\Sigma$ be such that $HK = 0$; we have to find $S, T \in {}^n M^n(R, \Sigma)$ such that $[S], [T]$ are mutually inverse and the relation $H[S] \cdot [T]K = 0$ is trivial.

By Prop. 11.12 there exist $A \in {}^1 M^n(R, \Sigma)$, $B \in {}^n M^1(R, \Sigma)$ such that $[A] = H$, $[B] = K$ and we have a Σ -dependence of length n :

$$\begin{pmatrix} A' & \tilde{A} & P_1 \\ A^0 & 'A & P_2 \end{pmatrix} \begin{pmatrix} Q_1 & Q_2 \\ B' & \tilde{B} \\ B^0 & 'B \end{pmatrix} = 0.$$

Let U, V be a pair trivializing this relation. Then for suitable $C, D \in M(R, \Sigma)$ we have

$$\begin{pmatrix} A' & \tilde{A} & P_1 \\ A^0 & 'A & P_2 \end{pmatrix} U = \begin{pmatrix} C' & \tilde{C} & . \\ C^0 & 'C & . \end{pmatrix}, \quad V \begin{pmatrix} Q_1 & Q_2 \\ B' & \tilde{B} \\ B^0 & 'B \end{pmatrix} = \begin{pmatrix} . & . \\ D' & \tilde{D} \\ D^0 & 'D \end{pmatrix},$$

where

$$\begin{pmatrix} \tilde{C} \\ 'C \end{pmatrix} (D' \quad \tilde{D}) = 0 \quad (24)$$

is a trivial relation. We shall put

$$S = \left(\begin{array}{cc|c} B' & U_{21} & U_{22} \\ B^0 & U_{31} & U_{32} \\ . & C^0 & 'C \end{array} \right), \quad T = \left(\begin{array}{cc|c} D' & V_{21} & V_{22} \\ D^0 & V_{31} & V_{32} \\ . & A^0 & 'A \end{array} \right);$$

then

$$S \odot T = \left(\begin{array}{cccc|c} B' & U_{21} & U_{22}D' & U_{22}V_{21} & U_{22}V_{22} \\ B^0 & U_{31} & U_{32}D' & U_{32}V_{21} & U_{32}V_{22} \\ \cdot & C^0 & 'CD' & 'CV_{21} & 'CV_{22} \\ \cdot & \cdot & D^0 & V_{31} & V_{32} \\ \cdot & \cdot & \cdot & A^0 & 'A \end{array} \right).$$

We remark that $'CD' = 0$, by (24). If we multiply the last expression on the left by the trivial factor

$$\left(\begin{array}{ccccc} 1 & \cdot & \cdot & U_{23} & \cdot \\ \cdot & 1 & \cdot & U_{33} & \cdot \\ \cdot & \cdot & 1 & \cdot & -1 \\ \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \end{array} \right)$$

we obtain after some simplification,

$$\left(\begin{array}{cccc|c} B' & U_{21} & B' & -U_{21}V_{11} & 1 - U_{21}V_{12} \\ B^0 & U_{31} & B^0 & -U_{31}V_{11} & -U_{31}V_{12} \\ \cdot & C^0 & \cdot & -C^0V_{11} & -C^0V_{12} \\ \cdot & \cdot & D^0 & V_{31} & V_{32} \\ \cdot & \cdot & \cdot & A^0 & 'A \end{array} \right).$$

We now multiply by the trivial factor

$$\left(\begin{array}{ccccc} 1 & \cdot & -1 & \cdot & \cdot \\ \cdot & 1 & \cdot & V_{11} & V_{12} \\ \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \end{array} \right)$$

on the right and obtain

$$\left(\begin{array}{cccc|c} B' & U_{21} & \cdot & \cdot & 1 \\ B^0 & U_{31} & \cdot & \cdot & \cdot \\ \cdot & C^0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & D^0 & V_{31} & V_{32} \\ \cdot & \cdot & \cdot & A^0 & 'A \end{array} \right)$$

and this matrix becomes $\begin{bmatrix} 1 \end{bmatrix}$ by omitting inessential rows and columns. Thus $S \odot T \sim \begin{bmatrix} 1 \end{bmatrix}$, and $T \odot S \sim \begin{bmatrix} 1 \end{bmatrix}$ follows by symmetry. Next we have

$$A \odot S = \left(\begin{array}{ccc|c} A' & \tilde{A}B' & \tilde{A}U_{21} & \tilde{A}U_{22} \\ A^0 & 'AB' & 'AU_{21} & 'AU_{22} \\ \cdot & B^0 & U_{31} & U_{32} \\ \cdot & \cdot & C^0 & 'C \end{array} \right).$$

If we multiply on the left by the trivial factor

$$\begin{pmatrix} 1 & \cdot & P_1 & \cdot \\ \cdot & 1 & P_2 & -1 \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix}$$

we obtain

$$\left(\begin{array}{ccc|c} A' & -A'Q_1 & C' - A'U_{11} & \tilde{C} - A'U_{12} \\ A^0 & -A^0Q_1 & -A^0U_{11} & -A^0U_{12} \\ \cdot & B^0 & U_{31} & U_{32} \\ \cdot & \cdot & C^0 & 'C \end{array} \right).$$

Multiplying on the right by the trivial factor

$$\begin{pmatrix} 1 & Q_1 & U_{11} & U_{12} \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix}$$

we find

$$\left(\begin{array}{ccc|c} A' & \cdot & C' & \tilde{C} \\ A^0 & \cdot & \cdot & \cdot \\ \cdot & B^0 & U_{31} & U_{32} \\ \cdot & \cdot & C^0 & 'C \end{array} \right)$$

and by omitting inessential rows and columns this becomes $\begin{pmatrix} C' & \tilde{C} \\ C^0 & 'C \end{pmatrix}$.

Similarly we find that $T \odot B \sim \begin{pmatrix} D' & \tilde{D} \\ D^0 & 'D \end{pmatrix}$. Thus we have

$$H[S] \cdot [T]K = [A \odot S] \cdot [T \odot B] = [C] \cdot [D].$$

By the triviality of (24) we may assume that the last r columns of C and the first $n-r$ rows of D are 0; hence by Cor. 11.10, the same holds for $[C], [D]$, so the relation $[C] \cdot [D] = 0$ is trivial. ■

COROLLARY 11.17. *Let R be a ring and Σ an n -independent set of matrices over R . Then the canonical map $\lambda: R \rightarrow R_\Sigma$ preserves the inner rank up to $n+1$; more precisely, for any matrix F over R , if $\rho(F\lambda) \leq n$, then $\rho F = \rho(F\lambda)$.*

Proof. Let $F \in \text{Mat}(R)$ have inner rank $r \leq n$ over R_Σ and write $F\lambda = HK$, where H, K are matrices over R_Σ and K has r rows. By Prop. 11.12 there exist $A, B \in M(R, \Sigma)$ such that $[A] = H, [B] = K$ and

$$\begin{pmatrix} A' & \tilde{A} & P_1 \\ A^0 & 'A & P_2 \end{pmatrix} \begin{pmatrix} Q_1 & Q_2 \\ B' & \tilde{B} \\ B^0 & 'B \end{pmatrix} = \begin{pmatrix} \cdot & F \\ \cdot & \cdot \end{pmatrix}. \quad (25)$$

We have the relation

$$\left(\begin{array}{c|c|c} A' & \tilde{A} & P_1 \\ \hline A^0 & 'A & P_2 \end{array} \right) \left(\begin{array}{c} Q_1 \\ \hline B' \\ \hline B^0 \end{array} \right) = 0,$$

which we can trivialize, and so can take to be trivial; in particular, $P_1, P_2 = 0$. Now we have the relation

$$(A^0 \quad 'A) \begin{pmatrix} Q_1 & Q_2 \\ B' & \tilde{B} \end{pmatrix} = 0,$$

and we may again take this to be trivial, so $Q_1, Q_2 = 0$. Now (25) takes the form

$$\begin{pmatrix} A' & \tilde{A} & \cdot \\ A^0 & 'A & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ B' & \tilde{B} \\ B^0 & 'B \end{pmatrix} = \begin{pmatrix} \cdot & F \\ \cdot & \cdot \end{pmatrix}.$$

Hence $F = \tilde{A}\tilde{B}$, and this shows that $\rho F \leq r$, but clearly $\rho F \geq \rho(F\lambda) = r$, so equality holds. Now if $\rho(F\lambda) < \rho F \leq n+1$, then $\rho(F\lambda) \leq n$, a contradiction. ■

When R itself is known to be an n -fir, the hypothesis in Th. 11.16 can still be weakened a little:

PROPOSITION 11.18. *Let R be an n -fir and Σ a set of matrices over R including 1_0 . Then Σ is n -independent if and only if the relation*

$$(A \mid X \mid U) \begin{pmatrix} \underline{V} \\ \underline{Y} \\ B \end{pmatrix} = 0, \quad A, B \in \Sigma, \quad (26)$$

of length at most n is trivializable.

Proof. Clearly the condition is necessary, so we may assume that it is satisfied. Let us take an arbitrary relation (18) of length at most n ; we may assume that

$$(\overline{A^0 \mid X_2 \mid U_2})(\underline{V_1 \mid Y_1 \mid B^0})^T = 0$$

is trivial, and so obtain a relation of the form

$$\left(\begin{array}{c|cc|c} A' & X_{11} & X_{12} & U_1 \\ \hline A^0 & X_{21} & \cdot & \cdot \end{array} \right) \begin{pmatrix} \cdot & V_2 \\ \cdot & Y_{12} \\ Y_{21} & Y_{22} \\ B^0 & B \end{pmatrix} = 0.$$

Thus we have the relations

$$(X_{12} \quad U_1) \begin{pmatrix} Y_{21} \\ B^0 \end{pmatrix} = 0, \quad (A^0 \quad X_{21}) \begin{pmatrix} V_2 \\ Y_{12} \end{pmatrix} = 0, \quad (27)$$

and we may again take them to be trivial (using the fact that $1_0 \in \Sigma$). The original relation now takes the form

$$\left(\begin{array}{c|cccc|c} A' & X'_{11} & X'_{12} & X'_{13} & \cdot & \cdot \\ \hline A^0 & X'_{21} & \cdot & \cdot & \cdot & \cdot \end{array} \right) \begin{pmatrix} \cdot & \cdot & \cdot & \cdot & Y'_{41} & B^0 \\ \cdot & \cdot & Y'_{22} & Y'_{32} & Y'_{42} & B \end{pmatrix} = 0,$$

and now the remaining relation $(X'_{12}, X'_{13})(Y'_{22}, Y'_{32})^T = 0$ can be trivialized because R is an n -fir. Thus the given relation has been trivialized. ■

We remark that it is not necessary to assume that $1_0 \in \Sigma$, provided that we assume explicitly the trivializability of the cases where a factor from Σ has an empty set of rows or columns; these are just the two cases listed in (27).

If we apply the result for $n = 0$ and observe that when Σ consists of regular matrices, the relations (27) are then trivial, we obtain

COROLLARY 11.19. *A set Σ of matrices over a non-zero ring R is independent if and only if all the matrices in Σ are regular and all relations $(A \parallel U)(\underline{V} \parallel B)^T = 0$ ($A, B \in \Sigma$) are trivializable. ■*

This result leads to another sufficient condition for the canonical map λ_Σ to be injective.

PROPOSITION 11.20. *Let R be a non-zero ring and Σ a set of matrices over R which are regular and such that any relation*

$$(A|U)\begin{pmatrix} V \\ B \end{pmatrix} = 0, \quad A, B \in \Sigma,$$

can be transformed to the form

$$(A'|0)\begin{pmatrix} 0 \\ B' \end{pmatrix} = 0, \quad A', B' \in \Sigma.$$

Then the map $\lambda: R \rightarrow R_\Sigma$ is injective.

For by Cor. 11.19, Σ is independent, so the result follows from Prop. 11.14. ■

Putting all our results together, we obtain the following localization theorem for n -firs:

THEOREM 11.21. *Let R be an n -fir and Σ a set of matrices over R such that any relation*

$$(A, X, U)(V, Y, B)^T = 0, \quad A, B \in \Sigma_0 \quad (28)$$

of length at most n is Σ_0 -trivializable (where $\Sigma_0 = \Sigma \cup \{1_0\}$). Then R_Σ is again an n -fir and λ_Σ preserves the inner rank up to $n+1$.

For Σ is n -independent by Prop. 11.18, and by Lemma 11.13 we may take Σ to be upper multiplicative. Now the result follows from Th. 11.16 and Cor. 11.17. ■

Let $r \geq 1, n \geq 0$ and suppose that R is a $(2r+n)$ -fir. Then the set Φ_r of all full $r \times r$ matrices satisfies the conditions of Th. 11.21, for the relation (28) has length $\leq 2r+n$ and so is trivializable. Moreover, each factor has rank at least r , because both A and B are full; thus Φ_r is n -independent. It follows that R_Σ is an n -fir, and this proves the Gerasimov-Malcolmson localization theorem:

THEOREM 11.22. *Let R be a $(2r+n)$ -fir ($n \geq 0, r \geq 1$) and Σ the set of all full $r \times r$ matrices over R . Then R_Σ is an n -fir and the map $\lambda: R \rightarrow R_\Sigma$ preserves the inner rank up to $n+1$; in particular it is injective. ■*

Exercises 7.11

1. Prove Lemma 11.2.

2. Let A^0, B^0 be invertible matrices with inverses α, β respectively. If

$$F = (\tilde{A} - A'\alpha'A)(\tilde{B} - B'\beta'B),$$

show directly that an equation of the form (17) holds.

3. Malcolmson's criterion states that the kernel of the map $\lambda: R \rightarrow R_\Sigma$ consists of all $c \in R$ such that

$$\begin{pmatrix} A & 0 & 0 \\ 0 & B & q \\ p & 0 & c \end{pmatrix} = \begin{pmatrix} U \\ u \end{pmatrix} \begin{pmatrix} V & v \end{pmatrix}, \quad \text{for some } A, B, U, V \in \Sigma.$$

Show that (16) can be expressed in this form.

4. Verify that over an n -fir any set consisting of invertible matrices is n -independent.

5*. (Valitskas [82]) Let R be any k -algebra (not necessarily with unit element), denote by R^1 the algebra $R \oplus k$ obtained by adjoining a unit element and put $\Sigma_n = \{I + A \mid A \in R_n\}$. If Σ_n is independent, in the sense that every Σ_n -dependence can be trivialized by a member of Σ_{2n} , show that the universal localization $\lambda: R \rightarrow R_\Sigma$ is injective; show also that R_Σ , where $\Sigma = \cup \Sigma_n$, is a radical ring.

6*. (Valitskas [82]) Let R be a k -algebra and Σ_n as before. Show that Σ_n is independent provided that

A_n . All relations $(e_1 + x) \cdot y = 0, x \cdot (e_n + y) = 0$ ($x \in R^n, y \in {}^n R$) can be trivialized.

Such a ring is called an A_n -ring. Show that for an A_{2n} -ring R , the canonical homomorphism $\lambda: R \rightarrow R_\Sigma$ is injective. (Valitskas (l.c.) gives examples of rings satisfying A_{2n} for which $\lambda: R \rightarrow R_{\Sigma_{n+1}}$ is not injective, using Th. 2.11.2. This shows that the class of rings embeddable in radical rings is not finitely axiomatizable.)

7. Let R be a k -algebra generated by 16 elements arranged as a 4×4 matrix A with defining relation $A^5 = 0$. Verify that R is a 3-fir but not a 4-fir. Show that if $\Sigma = R^\times$, then R_Σ is an integral domain and the natural homomorphism $R \rightarrow R_\Sigma$ is injective, but R is not embeddable in a field. (Hint: Use Th. 2.11.2 and Prop. 1.5.5.)

8*. (After Malcolmson [84]) Let $m > r > 0$ and let R be a k -algebra with $2m(m+r)$ elements arranged as an $m \times (m+r)$ matrix A and an $(m+r) \times m$ matrix B with defining relations $AB = 0$, and denote by Σ the set of all full $r \times r$ matrices. Show that R_Σ is an n -fir, where $n = m - r - 1$, but not an $(n+1)$ -fir. (Hint: Regard $AB = 0$ as a dependence relation of length $n+1$; simplify it over R_Σ so as to obtain $PQ = 0$, where P, Q are full $(n+1) \times (n+1)$ matrices over R_Σ .)

9. (Malcolmson [84]) Let R be a ring, Σ an n -independent set of matrices and

$\lambda: R \rightarrow R_\Sigma$ the universal localization. Show that every $r \times s$ matrix over R , where $\min(r, s) \leq n$, is inert in R_Σ .

10°. If $f: R \rightarrow S$ is an injective homomorphism between non-zero rings and $\mathfrak{B}_n$ is the matrix reduction functor from 2.11, show that $\mathfrak{B}_n(f)$ preserves the inner rank up to n .

11°. Show that over a semifir every factor-closed set is independent. Generalize to n -independent sets.

12. (V. N. Gerasimov) Let R be a k -algebra without 1 and R^1 the algebra obtained by adjoining 1 to R . Show that if R^1 has a universal field of fractions, then R can be embedded in a radical algebra. (Hint: Show that for any square matrix A over R , $I + A$ cannot lie in the least prime matrix ideal of R^1 .)

13. (P. Malcolmson) Let R be a non-zero ring and Σ an independent set of full matrices over R . Then a matrix M over R has a right inverse over R_Σ if and only if there exist matrices $A \in \Sigma$ and B, C, D, N over R such that

$$\begin{pmatrix} N & M \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & I \end{pmatrix}.$$

(Hint: Use Prop. 11.14.)

Notes and comments on Chapter 7

It is well known that a commutative ring R can be embedded in a field if and only if R is an integral domain. This condition is clearly still necessary in the non-commutative case, but no longer sufficient. This was shown by Malcev [37] in answer to v. d. Waerden ([30], p. 49) who had written 'Die Möglichkeit der Einbettung nichtkommutativer Ringe ohne Nullteiler in einen sie umfassenden Körper bildet ein ungelöstes Problem, ausser in ganz speziellen Fällen'. Malcev's counter-example was in the form of a monoid ring QS , where S is a monoid with cancellation, but not embeddable in a group (cf. Ex. 2.11.9). The existence of such a monoid also provided a counter-example to the claim by Sushkevich [36], that every cancellation monoid is embeddable in a group. Malcev followed up his example by two papers (Malcev [39, 40]), which gave a set of necessary and sufficient conditions for a monoid to be embeddable in a group, in the form of quasi-identities, and showing that these conditions formed an infinite set which could not be replaced by a finite subset.

The following classification is taken from Bokut [81]. Let $\mathcal{D}_0$ be the class of integral domains, $\mathcal{D}_1$ the class of rings R such that $R^\times$ is embeddable in a group, $\mathcal{D}_2$ the class of rings R such that the universal $R^\times$ -inverting map is injective and $\mathcal{E}$ the class of rings embeddable in fields. Then it is clear that

$$\mathcal{D}_0 \supseteq \mathcal{D}_1 \supseteq \mathcal{D}_2 \supseteq \mathcal{E}.$$

Here all the inequalities are strict: in answering v. d. Waerden's question (whether $\mathcal{D}_0 = \mathcal{E}$), Malcev proved that $\mathcal{D}_0 \neq \mathcal{D}_1$ and he raised the question whether $\mathcal{D}_1 = \mathcal{E}$. This was answered 40 years later by three people, independently and almost simultaneously: Bowtell [67] and Klein [67] gave examples which showed that $\mathcal{D}_2 \neq \mathcal{E}$ (cf.

Ex. 11.7), while Bokut [67, 69] gave examples showing that $\mathcal{D}_1 \neq \mathcal{D}_2$. The examples of Bowtell and Klein can be obtained quite simply by the methods of 2.11, which provide n -firs that are not $(n+1)$ -firs (Ex. 11.7) but Bokut's proofs have not been simplified. The examples of n -firs that are not embeddable in $(n+1)$ -firs (Ex. 10.4) can also be used to show quite simply that the necessary and sufficient conditions of 7.5 for embeddability in a field cannot be replaced by a finite subset (they are not equivalent to a finite set of elementary sentences, cf. Cohn [74', 81]).

Ex. 5.6 shows that the conditions for an integral domain to be embeddable in a field take the form of quasi-identities, as expected from the general theory (Cohn [81]). Thus $\mathcal{E} = \mathcal{T} \cap \mathcal{D}_0$, where $\mathcal{T}$ is a quasi-variety, namely the class of all subrings of strongly regular rings. The result that a subdirect product of rings in $\mathcal{E}$ is in $\mathcal{E}$ if and only if it is in $\mathcal{D}_0$ was proved by Burmistrovich [63].

Malcev's solution of the embedding problem for monoids gave no hint for the corresponding problem of embedding rings in fields. Until 1970 the only purely algebraic methods of embedding rings in fields were based on Ore's method (Ore [31], cf. 0.8; for a topological method see Cohn [61']). Schwarz [49] defined a form of quotient ring for a free algebra $R = k\langle X \rangle$ (essentially a universal localization R_{R^*}), but he did not succeed in constructing a field of fractions for R (it may be of interest to note that Schwarz began in 1931 as a student of E. Noether and later became a student of H. Weyl).

In fact, the first field of fractions for $R = k\langle x, y \rangle$ had been obtained by Moufang [37] by embedding R in kG , where G is the free metabelian group on x, y , and constructing a field of power series for kG (Schwarz seems not to have been aware of her work). In the light of Moufang's construction it is relevant to observe that the free monoid X^* cannot be embedded in a nilpotent group on X , because such a group always includes non-trivial monoid identities (cf. Malcev [53]). Later, Malcev [48] and Neumann [49] independently showed that $k\langle X \rangle$ has a field of fractions, by embedding it in the group ring kF , where F is the free group on X , and constructing the field of formal power series of kF in terms of an ordering of F (cf. 8.7).

The basic idea underlying this chapter, to invert matrices rather than elements, was inspired by the rationality criteria of Schützenberger [62] and Nivat [69] for non-commutative power series rings. The observation that these criteria had quite general validity, coupled with the notion of a full matrix (derived from the 'inner rank' defined by Bergman [67]) was exploited in Cohn [71] to embed firs in fields. These criteria survive in Th. 1.2 (the forms given by Schützenberger and Nivat correspond to (b) and (c), respectively). The application to firs is based on the following theorem (Cohn [71]). Let R be a ring and M a subset of R . Then the universal localization $R \rightarrow R_M$ is injective provided that M consists of regular elements and for any $p, q \in M$, $\text{Hom}_R(R/pR, R/qR)$ is 0 or a field according as $p \neq q$ or $p = q$. This shows for example that any atomic 2-fir is in $\mathcal{D}_2$. Other interesting consequences depend on the fact that the coproduct (or free product) of fields is a fir and so is embeddable in a field (Cohn [77], Ch. 5).

The notion of a 'free' field was introduced by Amitsur [66] as a result of studying generalized rational identities (cf. also Bergman [70]). Lewin [74] showed that the subfield of the Malcev–Neumann power series field generated by kF is the universal field of fractions of $k\langle X \rangle$.

The development in 7.2 is based on Cohn [72]; matrix ideals, 7.3 and 4, were first described in Cohn [72]. In the first edition the construction of an epic R -field with prescribed singular kernel was based on an axiomatic description of fields as groups

with an extra element 0 and an operation $x \mapsto 1-x$ (cf. Cohn [61], Dicker [68], Leissner [71] and, for an amusing connexion with Mersenne primes, Hotje [83]); this has here been replaced by the direct construction of Th. 4.8. For a general survey of the subject, including recent work by Bergman, Dicks and Schofield, see Cohn [84].

The field spectrum $X(R)$ was introduced by Cohn [72]; it leads to an 'affine scheme' $(X; \bar{R})$ associated with a general ring (cf. also Cohn [79]). For any ring one has a sheaf of local rings over X , and a natural homomorphism

$$\gamma: R \rightarrow \Gamma(X, \bar{R}),$$

into the ring of global sections but, of course, one cannot expect an isomorphism (as in the commutative case, cf. e.g. Macdonald [68]). Thus, for example, if R has no homomorphisms into fields, then $X(R) = \emptyset$ and the scheme is trivial. Besides the sections arising from R (the 'integral' sections) one also has the following 'rational' sections: Let $A \in \mathfrak{M}(R)$; if A becomes invertible in every localization, then the entries of A^{-1} define *rational sections* and one may ask: (i) for which rings is every global section rational and (ii) for which rings is every rational section integral? For fully atomic semifirs (i) holds, by Th. 7.3. There is also a generalization of the field spectrum to take into account homomorphisms into simple Artinian rings (cf. Cohn [79], Ringel [79]). An even more general spectrum, the epi-spectrum, has been defined by Bergman [70]; for a commutative ring this agrees with the usual prime spectrum and, in general, it is never empty, but it is difficult to determine explicitly (for Noetherian rings it agrees with the previous case, cf. Cohn [79]).

Sections 7.6–8 are extracted with only small changes from the trilogy Cohn [82, 85, 85'], where further results on absolute rational identities, semifirs of bounded depth and centralizers in field coproducts are to be found. Section 7.9 is based on Cohn [82', 83], with some simplifications; for a different approach (and more precise results for free algebras) see Révész [83]. There is a corresponding development of valuations, briefly mentioned in Ex. 9.5 and 6 (cf. Mahdavi-Hezavehi [82]), and a description of orderings of epic R -fields in terms of 'matrix cones' on R , by Révész [83']. The localization theorem for semifirs, Th. 10.4, is taken from Cohn and Dicks [76], which corrects and complements results of Cohn [74']. The extension to Sylvester domains in Th. 10.5 was obtained by Cohn and Schofield [82]. Bergman and Dicks [78] first proved Th. 10.6, the localization theorem for hereditary rings; the very simple proof given here is essentially due to Dlab and Ringel [85], in a formulation by W. Dicks. The corresponding statement for flat epimorphisms is much easier, cf. Hudry [70]. The behaviour of atoms under localization (10.8–10) is described in Cohn [82'], where these results are used to prove the following theorem. Let $R = k\langle X \rangle$ be the free algebra on an infinite set X , let X_0 be a subset of X with an infinite complement, and denote by Σ the set of all full matrices totally coprime to the elements of X_0 . Then R_Σ is a simple principal ideal domain. This result is used in Cohn and Schofield [85] to construct a simple Bezout domain which is right but not left principal, thus answering Ex. 1.2.9^o of the first edition.

Section 7.11 is based on Gerasimov [82], with minor simplifications; in particular he proves Th. 11.22. This is proved independently by Malcolmson [84], who also gives examples of a $(2r+n)$ -fir whose localization at all full $r \times r$ matrices is not an $(n+1)$ -fir (cf. Ex. 11.8). The result has been applied by Valitskas [82], see Ex. 11.5 and 6.

8. Principal ideal domains

There is one important property of principal ideal domains which does not carry over to firs: the diagonal reduction of matrices. This is discussed in 8.1, and its consequences for modules are described in 8.2. In particular, these results apply to skew polynomial rings (8.3 and 4) and matrices over fields (8.5), and they allow us to analyse certain simple types of algebraic skew field extensions in 8.6. Many of the earlier results of the book can be used here, and this has been done wherever it simplifies matters. In 8.7 we describe the Malcev–Neumann construction of the ordered series field of an ordered group, and prove the Bergman conjugacy theorem (Th. 7.7). The concluding section 8.8 deals with Jategaonkar's iterated skew polynomial rings, which are here characterized as the rings with a unique remainder algorithm.

8.1 The diagonal reduction for matrices over a principal ideal domain

Let R be a fir and $A \in {}^mR^n$; we can always find $P \in \text{GL}_m(R)$ and $Q \in \text{GL}_n(R)$ such that $PAQ = A_1 \oplus 0$, where A_1 is a regular matrix, but in general no further reduction is possible. However, when R is a principal ideal domain, we can reduce A to diagonal form. In order to state the result, we need a definition. In an integral domain, an element a is said to be a *total divisor* of b , written $a \parallel b$, if there exists an invariant element c such that $a|c|b$. We observe that an element is not generally a total divisor of itself; in fact $a \parallel a$ if and only if a is invariant. If the ring R is simple, it has no non-unit invariant elements and $a \parallel b$ implies that either a is a unit or b is 0. Further, in a principal ideal domain R , every ideal is of the form $Rc = cR$, where c is 0 or an invariant element, by Prop. 6.1.4.

The notation $\text{diag}(a_1, \dots, a_r)$ for a matrix with $a_1, \dots, a_r$ on the main diagonal and 0's elsewhere will be used here even for matrices that are not square; the exact size will be indicated explicitly, unless it is clear from the context, as in (1), below.

THEOREM 1.1. *Let R be a principal ideal domain and $A \in {}^mR^n$. Then the row and column rank of A are the same; denoting the common value by r , we can find $P \in \text{GL}_m(R)$, $Q \in \text{GL}_n(R)$ such that*

$$P^{-1}AQ = \text{diag}(e_1, \dots, e_r, 0, \dots, 0), \quad e_i \parallel e_{i+1}, e_r \neq 0. \quad (1)$$

Proof. We have the following four types of operations on the columns of A , of which the first three are the well-known elementary operations:

- (i) *interchange two columns,*
- (ii) *multiply a column on the right by a unit factor,*
- (iii) *add a right multiple of one column to another,*
- (iv) *multiply two columns on the right by an invertible 2×2 matrix.*

As is well known, each of (i)–(iii) corresponds to right multiplication by an elementary matrix. The object of using (iv) is to replace the first two elements in the columns by their highest common left factor and 0, respectively. Thus if these elements are a, b , not both zero, we have $aR + bR = kR$, say $a = ka_1$, $b = kb_1$. This means that (a_1, b_1) has a right inverse, and so can be completed to an invertible 2×2 matrix C say. Hence $(k, 0)C = (a, b)$ and $(a, b)C^{-1} = (k, 0)$, as required. Corresponding operations can, of course, be carried out on the rows, acting on the left.

We can now proceed with the reduction. If $A = 0$, there is nothing to prove; otherwise we bring a non-zero element to the $(1, 1)$ -position in A , by permuting rows and permuting columns, using (i). Next we use (iv) to replace a_{11} successively by the HCLF of a_{11} and a_{12} , then by the HCLF of the new a_{11} and a_{13} , etc. After $n-1$ steps we have transformed A to a form where $a_{12} = \dots = a_{1n} = 0$. By symmetry the same process can be applied to the first column of A ; in the course of the reduction the first row of A may again become non-zero, but this can happen only if the length (i.e. the number of factors) of a_{11} is reduced; therefore by induction on the length of a_{11} we reach the form

$$A = a_{11} \oplus A_1.$$

By another induction, on $\max(m, n)$, we reach the form

$$A = \text{diag}(a_1, a_2, \dots, a_t, 0, \dots, 0).$$

Consider a_1 and a_2 ; for any $d \in R$ we have

$$\begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} = \begin{pmatrix} a_1 & da_2 \\ 0 & a_2 \end{pmatrix},$$

and now we can again diminish the length of a_1 unless a_1 is a left factor of da_2 for all $d \in R$, i.e. unless $a_1 R \supseteq Ra_2$. But in that case $a_1 R \supseteq Ra_2 R \supseteq Ra_2$; thus $a_1 | c a_2$, where c is the invariant generator of the ideal $Ra_2 R$. Hence $a_1 \parallel a_2$, and by repeating the argument we obtain the form

$$\text{diag}(e_1, e_2, \dots, e_t, 0, \dots, 0), \quad \text{where } e_i \parallel e_{i+1}.$$

If $e_t \neq 0$, this matrix has row and column rank t , but clearly A and $P^{-1}AQ$

have the same column rank. Similarly for the row rank and the assertion follows. ■

We shall return to this theorem in the next section, to see how far the e 's are unique. For the moment we note that if R is a Euclidean domain (hence a PID), we can instead of (iv) use the Euclidean algorithm, with an induction on the degree instead of the length, to accomplish the reduction in Th. 1.1. Most of the PID's we encounter will, in fact, be Euclidean.

We record two consequences of Th. 1.1.

COROLLARY 1.2. *Let R be a simple principal ideal domain; then any matrix over R is associated to $I \oplus a \oplus 0$, and hence stably associated to $a \oplus 0$, where $a \in R$.*

For if $a \parallel b$, then either $b = 0$ or a is a unit. Now any unit can be transformed to 1 by applying (ii) and, as we saw, there can be only one diagonal element not zero or a unit. ■

In the case of a field, every non-zero element is a unit, so in this case every matrix is associated to $I_r \oplus 0$, where r is the rank of A .

As a further application of Th. 1.1 we describe the rank of a matrix over $K(t)$, where K is a field.

LEMMA 1.3. *Let K be a field with infinite centre C , and consider the polynomial ring $K[t]$, with field of fractions $K(t)$. If $A = A(t)$ is a matrix over $K[t]$, then the rank of A over $K(t)$ is the supremum of the ranks of $A(\alpha)$, $\alpha \in C$. In fact, this supremum is assumed for all but a finite number of values of α .*

Proof. By Th. 1.1 we can find invertible matrices P, Q over $K[t]$ such that

$$P^{-1}AQ = \text{diag}(f_1, \dots, f_r, 0, \dots, 0), \quad \text{where } f_i \in K[t]. \quad (2)$$

The product of the non-zero diagonal terms on the right gives us a polynomial f whose zeros in C are the only points of C at which $A = A(t)$ falls short of its maximum rank, and the number of these values cannot exceed $\deg f$. ■

Let $R = K[t]$, where K is a field. If f is a polynomial of degree d , then $M = R/fR$ is an R -module which as K -space has dimension d . More generally, let M be the right R -module defined by the matrix A ; if A is regular, then M is finite-dimensional as K -space; in fact, its dimension is $\deg f_1 \dots f_r$, if A can be transformed to the diagonal form (2).

Exercises 8.1

1. Show that (iv) is not needed in the proof of Th. 1.1 if R is also an E_2 -ring. What simplifications are possible when R is commutative? Prove the uniqueness in this case (Th. 2.4).

2°. What kind of reduction theorem can be proved for R when R is (i) an atomic principal right ideal domain or (ii) a Bezout domain?

3. (Kaplansky [49]) By an *elementary divisor ring* is meant a ring over which every matrix admits a diagonal reduction as in Th. 1.1. Show that a ring over which every $m \times n$ matrix, with $m, n \leq 2$, admits a diagonal reduction is an elementary divisor ring.

4. (Kaplansky [49]) Show that an elementary divisor ring which is an integral domain is weakly finite.

5. Let R be an integral domain which is weakly finite. Show that if R has an infinite centre, then $R[t]$ is again weakly finite.

6. Show that an elementary divisor ring which is a fir is necessarily principal.

7°. Give an example of a commutative Bezout domain which is not an elementary divisor ring.

8°. (A. Hausknecht) Let R be a principal ideal domain; if the units together with 0 form a field k , is R necessarily a polynomial ring over k ?

9. A ring R is called *semi-Euclidean* (D. Goldschmidt) if there is a function $\phi: R^\times \rightarrow \mathbb{N}$ such that for any $a, b \in R^\times$ either $a = bq + r$, $\phi(r) < \phi(b)$, or $\phi(a) = \phi(b)$ and $b = aq + r$, $\phi(r) < \phi(a)$. Prove a triangular reduction for semi-Euclidean rings. Show that every valuation ring is semi-Euclidean.

10. Show that a Bezout domain with right ACC_1 is right principal.

11. Show that a principal ideal domain is primitive if and only if it has an unbounded atom.

12*. Show that Bezout domains form an elementary class (i.e. can be defined by elementary sentences; the class of PID's is not closed under ultraproducts and so cannot be elementary).

Show that a principal ideal domain may be characterized as a Bezout domain such that A : every non-unit has an atomic left factor, and B : left ACC_1 holds. Deduce that any principal right ideal domain which is elementarily equivalent to a PID is itself a PID. (Hint: A is elementary but not B , cf. Cohn [a].)

8.2 Finitely generated modules over principal ideal domains

Let R be a principal ideal domain; then R is in particular a right Ore domain, and so every right R -module M has a torsion submodule tM with

torsion free quotient M/tM (cf. 0.9). Suppose now that M is finitely generated torsion free over R ; then M can be embedded in a free R -module, by Cor. 0.9.5, and hence is free, by Th. 1.1.1. This proves

PROPOSITION 2.1. *Every torsion free R -module over a principal ideal domain R is free. ■*

Let us return to an arbitrary finitely generated R -module M . Then M/tM is finitely generated torsion free and hence free; therefore it can be lifted to a free submodule of M complementing tM :

COROLLARY 2.2. *Let R be a principal ideal domain. If M is a finitely generated R -module, there exists a free submodule F of M such that*

$$M = tM \oplus F.$$

Here tM is uniquely determined as the torsion submodule of M , while F is unique up to isomorphism. ■

This is, of course, just a special case of the decomposition over semifirs described in 5.1. The result shows that in classifying finitely generated modules over PID's we can restrict ourselves to torsion modules, but there would be no gain in making this simplification. However, there is another important (and well-known) fact which will be useful to us; it follows from Cor. 0.3.4, that we have

PROPOSITION 2.3. *Let R be a principal ideal domain and M an n -generator R -module. Then each submodule of M can be generated by n elements.*

This holds for free modules, by Cor. 0.3.4; since every module can be written as a homomorphic image of a free module, it holds generally. ■

In terms of the characteristic we can express the result by saying that over a principal ideal domain, every finitely generated module has non-negative characteristic.

Prop. 2.3 shows, in particular, that every finitely generated module is finitely presented. Let M be a finitely generated module with presentation

$${}^nR \xrightarrow{\alpha} {}^mR \longrightarrow M \longrightarrow 0, \quad (1)$$

and let A be the $m \times n$ matrix over R which represents the homomorphism $\alpha: {}^nR \rightarrow {}^mR$ relative to the standard bases in ${}^nR, {}^mR$. Then M is completely

specified by A , and if we change the bases in ${}^nR$ and ${}^mR$, this amounts to replacing A by $P^{-1}AQ$, where $P \in GL_m(R)$ and $Q \in GL_n(R)$. Using Th. 1.1 we now obtain the following generalization of the basis theorem for abelian groups:

THEOREM 2.4. *Let R be a principal ideal domain. Then any finitely generated R -module M is a direct sum of cyclic modules:*

$$M \cong R/e_1R \oplus \dots \oplus R/e_rR \oplus {}^{m-r}R, \quad (2)$$

where $e_i \parallel e_{i+1}$, and this condition determines the e_i up to similarity.

Proof. Let M be defined by a presentation (1) with matrix A . By Th. 1.1, A is associated to $\text{diag}(e_1, \dots, e_r, 0, \dots, 0)$ with $e_i \parallel e_{i+1}$, and since this change does not affect the module, we obtain (2). It only remains to prove the uniqueness.

We begin with the remark that finitely generated modules can be cancelled. Thus if

$$M \oplus N \cong M \oplus N', \quad (3)$$

where M, N, N' are finitely generated, then $N \cong N'$. For N and N' clearly have the same rank, and by applying Cor. 2.2 we may assume that M, N, N' are all torsion modules. Now $N \cong N'$ follows on taking complete decompositions on both sides of (3) and applying the Krull-Schmidt theorem, Th. 3.6.3. Let us write $R/a_1R \oplus \dots \oplus R/a_kR$ as $[a_1, \dots, a_k]$ for short and use $\sim$ to indicate isomorphism. Then $[a] \sim [b]$ if and only if a is similar to b , and what has been shown is that $[a, b_1, \dots, b_r] \sim [a, b'_1, \dots, b'_s]$ implies $[b_1, \dots, b_r] \sim [b'_1, \dots, b'_s]$.

Let us take two representations of M as direct sums of cyclic modules:

$$M \sim [d_1, \dots, d_r] \sim [e_1, \dots, e_r], \quad d_i \parallel d_{i+1}, e_i \parallel e_{i+1}. \quad (4)$$

There is no loss of generality in assuming the same number of summands on both sides, since we can always add zero summands, represented by unit factors: $R/R = 0$. Further we may suppose that the torsion free part of M has been split off, so that the d_i, e_i are all different from 0. If $r = 1$, the result is clear by what has been said, so let $r > 1$ and use induction on r . We shall write $l(a)$ for the length of a and assume that $l(d_1) \geq l(e_1)$; further, let $d_1 \mid c \mid d_2$, where c is invariant. If N is any module, Nc is a submodule; more specifically, if $N = R/aR$ and c is invariant, then $N/Nc \cong R/(aR + cR)$. Now consider M/Mc ; writing $e_iR + cR = e'_iR$, we have by (4),

$$M/Mc \sim [d_1, c, \dots, c] \sim [e'_1, e'_2, \dots, e'_r], \quad (5)$$

and $l(e'_i) \leq l(c)$ ($i = 1, \dots, r$), $l(e'_1) \leq l(e_1) \leq l(d_1)$. Comparing lengths in (5),

which must be equal, as the length of a composition series for M/Mc , we find that $l(d_1) + (r-1)l(c) = \sum l(e'_i)$, i.e.

$$l(d_1) - l(e'_1) + \sum_2^r (l(c) - l(e'_i)) = 0.$$

Since each term is non-negative, all are zero and $l(e'_1) = l(e_1) = l(d_1)$, $l(e'_i) = l(c)$. It follows that $e'_1 \sim e_1$, $e'_i \sim c$ ($i > 1$), and (5) now reads

$$[d_1, c, \dots, c] \sim [e_1, c, \dots, c].$$

By cancellation we find that e_1 is similar to d_1 ; so we may cancel the first term on both sides of (4) and obtain $[d_2, \dots, d_r] \sim [e_2, \dots, e_r]$. Now an induction on r gives the result. ■

For a torsion module the last term in (2) is absent. If, moreover, R is simple, then there are no invariant elements apart from units, and so at most one of the e_i in (2) can be a non-unit. But units can in any case be omitted. Thus (2) reduces to a single term and we see that over a simple PID every finitely generated torsion module is cyclic. In Prop. 2.8, below, we shall obtain this result in a somewhat more general context.

We also record a remark made during the proof:

COROLLARY 2.5. *If M, N, N' are finitely generated modules over a principal ideal domain such that $M \oplus N \cong M \oplus N'$, then $N \cong N'$. ■*

Th. 2.4 shows that the e_i in Th. 1.1 are determined up to similarity; we shall generally omit the units among them, since they do not contribute to M . The e_i are called the *invariant factors* of the matrix A or also of the module M . The condition imposed on the e_i (that each e_i be a total divisor of the next) ensures that (2) is a decomposition of M into cyclic modules with as few terms as possible. At the other extreme we have a decomposition into as many terms as possible, i.e. a complete direct decomposition into indecomposable modules. The indecomposable parts must then be cyclic, by Th. 2.4 and they are unique up to order, by the Krull-Schmidt theorem. The factors in this case are called the *elementary divisors* of the module (or the matrix). For example, over the integers, $M = \mathbb{Z}/3 \oplus \mathbb{Z}/15 \oplus \mathbb{Z}/750$ has the invariant factors 3, 15, 750 and the elementary divisors $3; 3, 5; 2, 3, 5^3$.

To obtain a complete decomposition we apply Th. 6.4.5, bearing in mind that all but the last non-zero invariant factor divide an invariant element and so are bounded:

PROPOSITION 2.6. *Let R be a principal ideal domain and M a finitely generated*

torsion module over R . Then

$$M \cong R/q_1R \oplus \dots \oplus R/q_kR \oplus R/uR,$$

where each q_i is a product of pairwise similar bounded atoms, while u is totally unbounded. The last term R/uR may be absent; this is so if and only if M is bounded, i.e. $\text{Ann}(M) \neq 0$. ■

It is of interest to extend the decomposition of Th. 2.4 beyond the principal case; here the following lemma is useful:

LEMMA 2.7. *Let R be a non-Artinian simple ring. Then any R -module of finite length is cyclic.*

Proof. Let M be a right R -module of finite length and suppose first that M has a simple submodule which is a direct summand:

$$M = M_1 \oplus S, \quad \text{where } S \text{ is simple.} \quad (6)$$

By induction on the length, M_1 is cyclic, say $M_1 = uR$. We denote by α the right annihilator of u in R ; it is a right ideal and since R is non-Artinian, $\alpha \neq 0$. If $S\alpha = 0$, then $S = SR = S\alpha R = 0$, a contradiction, so there exists $v \in S$ such that $\text{Ann}(v) \not\supseteq \alpha$. It is clear that $v \neq 0$, so $vR = S$ by simplicity; we claim that $u+v$ generates M . Consider $\text{Ann}(u+v)$; if $(u+v)x = 0$, then $ux = vx = 0$ by (6), so $x \in \alpha \cap \text{Ann}(v)$. Thus the map $x \mapsto (u+v)x$ gives an isomorphism

$$R/(\alpha \cap \text{Ann}(v)) \cong (u+v)R.$$

On the left we have a module of length $> l(M_1)$, hence of length $l(M)$, so $(u+v)R = M$ and M is cyclic, as claimed.

There remains the case when no simple submodule is a direct summand. Of course, M has a simple submodule (unless $M = 0$), S say. By induction hypothesis M/S is cyclic, generated by $u+S$, say. Hence $M = uR + S$ and by hypothesis, $uR \cap S \neq 0$, so $uR \supseteq S$ by the simplicity of S , therefore $uR = M$, and M is cyclic. ■

As a consequence we have

PROPOSITION 2.8. *Let R be an atomic simple principal right ideal domain, not a field. Then any finitely generated right torsion R -module is cyclic.*

For R is non-Artinian simple, so we need only verify that a finitely generated right torsion module is of finite length. By induction it is enough to check this for cyclic modules and here it is clear, by unique factorization and the fact that all right ideals are principal. ■

This result then shows that over an atomic simple principal right ideal domain R , every regular matrix A is stably associated to an element of R .

Exercises 8.2

1. Let R be a principal ideal domain. Show that if the elementary divisors of any module are powers of atoms, then R is invariant. (Hint: Use Th. 4.2.7.)

2. Obtain an analogue of Prop. 2.6 for modules, in which every finitely generated submodule is bounded.

3. Let M be a finitely generated torsion module over a principal ideal domain R . Given a decomposition (2) into cyclic modules which are as 'short' as possible, show directly that the e_i can be numbered so that each is a total divisor of the next.

4*. Let M be a finitely generated module over a principal ideal domain. Prove directly that the bounds of the elementary divisors of M are independent of the choice of decomposition of M .

5°. Let R be a 2-fir; if every finitely generated torsion free right R -module is free, is R necessarily left Bezout?

6°. Which of the results of this section go over to principal right ideal domains, or to Bezout domains?

8.3 Skew polynomial rings

An important example of principal ideal domains is given by the skew polynomial rings over a field, and we shall therefore examine these rings briefly. Let K be any field with an endomorphism α and an α -derivation δ , and consider the skew polynomial ring

$$R = K[t; \alpha, \delta]. \quad (1)$$

We recall that it has the defining relations

$$ct = tc^\alpha + c^\delta \quad (c \in K). \quad (2)$$

When R is given, t is determined up to a linear transformation; a change of variable changes the endomorphism α by an inner automorphism and when α is fixed, changes δ by an inner α -derivation, i.e. a derivation of the form

$$\Delta_c : a \mapsto ac - ca^\alpha.$$

We utilize this choice to simplify the expression (1) for R as follows. Let C be the centre of K and assume first that α does not leave C fixed, say $\gamma^\alpha \neq \gamma$ for

some $\gamma \in C$. On changing the variable to $t_1 = \gamma t - t\gamma$, we find that for any $a \in K$,

$$at_1 = a(\gamma t - t\gamma) = \gamma ta^\alpha + \gamma a^\delta - ta^\alpha \gamma - a^\delta \gamma = (\gamma t - t\gamma)a^\alpha,$$

i.e.

$$at_1 = t_1 a^\alpha,$$

so we see that δ is inner, and by a suitable choice of variable may be taken to be 0. Next we assume that α leaves C fixed, but there exists $\gamma \in C$ such that $\gamma^\delta \neq 0$. Then $\gamma^\alpha = \gamma$ and for any $a \in K$, $a\gamma = \gamma a$, hence $a^\delta \gamma + a\gamma^\delta = \gamma^\delta a^\alpha + \gamma a^\delta$, i.e. $a\gamma^\delta = \gamma^\delta a^\alpha$, and on putting $t_1 = t(\gamma^\delta)^{-1}$, we have

$$at_1 = ta^\alpha(\gamma^\delta)^{-1} + a^\delta(\gamma^\delta)^{-1} = t(\gamma^\delta)^{-1}a + a^\delta(\gamma^\delta)^{-1} = t_1 a + a^{\delta'},$$

where δ' is an ordinary derivation of K .

The only remaining possibility is that α leaves C fixed and δ maps it to 0. Thus we have

PROPOSITION 3.1. *Let $R = K[t; \alpha, \delta]$ be a skew polynomial ring over a field K with endomorphism α and α -derivation δ . Then either*

- (i) δ is inner, and by a suitable choice of t may be taken to be 0,
- (ii) α is inner, and by a suitable choice of t may be taken to be 1 or
- (iii) α leaves the centre C of K fixed and δ maps it to 0; in that case C is contained in the centre of R . ■

Next we examine the ideals of $R = K[t; \alpha, \delta]$. Every non-zero right ideal of R is of the form fR , where f is a polynomial in t which may without loss of generality be taken to be monic. The condition for f to be right invariant (and hence for fR to be two-sided) is that for each $g \in R$ there should exist $g_1 \in R$ such that $gf = fg_1$. In particular, we must have

$$\begin{aligned} cf &= fc_1 \quad \text{for all } c \in K, \\ tf &= f \cdot (tu + v). \end{aligned}$$

Here $c_1, u, v \in K$, by a comparison of degrees. If f has the form

$$f = t^n + t^{n-1}a_1 + \dots + a_n, \tag{3}$$

then $cf = t^n c^{\alpha^n} + \dots, fc_1 = t^n c_1 + \dots$, hence by equating highest terms we find $c_1 = c^{\alpha^n}$. Similarly we have $tf = t^{n+1} + t^n a_1 + \dots$,

$$\begin{aligned} f \cdot (tu + v) &= t^{n+1}u + t^n v + t^{n-1}a_1 tu + \dots \\ &= t^{n+1}u + t^n(v + a_1^{\alpha}u) + \dots \end{aligned}$$

It follows that $u = 1, v + a_1^{\alpha} = a_1$. Thus we can describe the right invariant elements of R as follows:

PROPOSITION 3.2. *In any skew polynomial ring $R = K[t; \alpha, \delta]$ over a field K with endomorphism α and α -derivation δ , the element f of R given by (3) is right invariant if and only if*

$$cf = fc^{\alpha} \text{ for all } c \in K \quad (4)$$

and

$$tf = f \cdot (t + a), \quad \text{where } a = a_1 - a_1^{\alpha}. \quad (5)$$

For we have seen that these conditions are necessary; conversely, when they hold, then by an easy induction on the degree, for any $g \in R$ there exists $g_1 \in R$ such that $gf = fg_1$ and so f is then right invariant. ■

If α is an automorphism, with inverse β , then (4), (5) may be restated as $fc = c^{\beta}f$, $ft = (t - a^{\beta}) \cdot f$, hence we have

COROLLARY 3.3. *If $R = K[t; \alpha, \delta]$, where α is an automorphism of K , then any right invariant element of R is invariant. ■*

Given an endomorphism α , an α -derivation δ of K and $n > 0$, we write $x^{\varepsilon} = x^{\delta\alpha^n} - x^{\alpha^n\delta}$. Then ε is clearly linear and

$$\begin{aligned} (xy)^{\varepsilon} &= x^{\alpha^n}y^{\delta\alpha^n} + x^{\delta\alpha^n}y^{\alpha^{n+1}} - x^{\alpha^n}y^{\alpha^n\delta} - x^{\alpha^n\delta}y^{\alpha^{n+1}} \\ &= x^{\alpha^n}y^{\varepsilon} + x^{\varepsilon}y^{\alpha^{n+1}}. \end{aligned}$$

Thus ε is an (α^n, α^{n+1}) -derivation. This derivation is necessarily inner when there is a right invariant element of degree n :

COROLLARY 3.4. *Let $R = K[t; \alpha, \delta]$ be as in Prop. 3.2 and assume that $f \in R$, of degree n as polynomial in t , given by (3), is right invariant. Then $\delta\alpha^n - \alpha^n\delta$ is inner, induced by $a_1 - a_1^{\alpha}$. In particular, if $\delta\alpha^n - \alpha^n\delta$ is outer for all $n > 0$, then R is simple.*

Proof. If $f \in R$ is right invariant, then $gf = fg_1$ for all $g \in R$, and here $g \mapsto g_1$ is an endomorphism of R . If we apply this endomorphism to (2) and write $a = a_1 - a_1^{\alpha}$, where a_1 is as in (3), then

$$c^{\alpha^n}(t+a) = (t+a)c^{\alpha^{n+1}} + c^{\delta\alpha^n}.$$

Using (2), we can simplify this to

$$c^{\alpha^n}a - ac^{\alpha^{n+1}} = c^{\delta\alpha^n} - c^{\alpha^n\delta},$$

and this shows that $\varepsilon = \delta\alpha^n - \alpha^n\delta$ is inner, induced by a . Therefore, if ε is outer, there can be no non-unit invariant element and R is simple. ■

With the help of Prop. 3.2 we can also describe the ideal structure of R , when α is surjective.

THEOREM 3.5. *Let $R = K[t; \alpha, \delta]$ be a skew polynomial ring over a field K with automorphism α and α -derivation δ . Assume further that no power of α is an inner automorphism of K . Then the monic right invariant elements of R form a monoid M , and either $M = \{1\}$ and R is simple, or M is generated by a single element $d \neq 1$, and the ideals of R are all of the form $d^v R$ ($v = 0, 1, \dots$).*

Proof. Let us call $f \in R$ right K -invariant if for each $c \in K$ there exists $c' \in K$ such that $cf = f'c'$. We have seen in Prop. 3.2 that each right invariant element f of R is necessarily K -invariant and that for a monic f of degree n we have $c' = c^{\alpha^n}$. We begin by showing that if there are non-unit monic right K -invariant elements, they are all powers of a single one. Since every ideal is generated, as right ideal, by a right invariant element, such elements always exist unless R is simple.

Let f be monic right K -invariant of degree n and choose u to be monic right K -invariant of least positive degree m . We have $f = uq + r$, where $\deg r < \deg u$. Now for any $c \in K$,

$$c(uq + r) = (uq + r)c^{\alpha^n} = uc^{\alpha^n}q + cr,$$

hence

$$u(qc^{\alpha^n} - c^{\alpha^n}q) = cr - rc^{\alpha^n}.$$

Here the right-hand side has lower degree than u , hence both sides are 0 and $cr = rc^{\alpha^n}$. By the minimality of the degree of u we have $r \in K$, but no power of α is inner, so $r = 0$ and $f = uq$. Now q is again right K -invariant and an induction on the degree of f shows that $f = u^v$ for some v .

In particular, this applies when f is monic right invariant, so any such element is a power of u , and of any two, one is a multiple of the other. Now if f and fg both satisfy (5) (possibly with different values of a), then so does g , as is easily checked. It follows that the monic right invariant elements are all powers of a single one, $d = u^u$, say, as claimed. ■

Let R be a skew polynomial ring and M the monoid of right invariant elements. Then we can always form the localization R_M and this will be a simple principal ideal domain, by Prop. 1.5.4. When $M \neq 1$, say $1 \neq d \in M$, then R_M is left as well as right principal. If α is an automorphism of K , this is clear; otherwise R_M contains $K^{[\alpha]} = \bigcup t^n K t^{-n}$ (cf. Ex. 0.10.4), because on R_M , α^m agrees with the inner automorphism induced by $d \in M$, where $m = \deg d$. Hence R_M is a localization of $K^{[\alpha]}[t; \alpha, \delta]$, and this is principal. Thus we obtain

COROLLARY 3.6. *Let $R = K[t; \alpha, \delta]$ be a skew polynomial ring, where no power of α is inner. Either (i) there is a non-constant polynomial which is right invariant; then the monoid M of monic right invariant elements is a right Ore set and R_M is a simple principal ideal domain, or (ii) there is no non-constant right invariant polynomial; then R is a simple principal right ideal domain, which is left principal if and only if α is an automorphism of K .*

Here the last part follows because for a non-surjective endomorphism α , R is not left Ore and hence not left principal (cf. Ex. 0.10.3). ■

The situation when a power of α is inner is described by

PROPOSITION 3.7. *Let $R = K[t; \alpha, \delta]$ be a skew polynomial ring over a field K , where α is an automorphism of which a positive power is inner on K . Then, unless R is simple, there exists an invariant element u of R inducing a power of α , and every invariant element either centralizes K (up to an inner automorphism) or is a power of u .*

Proof. Let $u \in R$ be a monic right K -invariant element of least positive degree v , say. Any right K -invariant element f , again taken monic for convenience, may be written

$$f = uq + r, \quad q, r \in R, \deg r < \deg u. \quad (6)$$

Write $cu = uc^\lambda$ (so that $\lambda = \alpha^v$, where $v = \deg u$, by Prop. 3.2), and $cf = fc^\mu$. Then $fc^\mu = cf = c(uq + r) = uc^\lambda q + cr$, hence

$$u(qc^\mu - c^\lambda q) = cr - rc^\mu, \quad (7)$$

and the right-hand side has degree less than u , so both sides must vanish; in particular, r is right K -invariant, and by the minimality of $\deg u$, we must have $\deg r = 0$, so $r \in K$. Now an induction on the degree shows

$$f = u^m + u^{m-1}a_1 + \dots + a_m, \quad a_i \in K. \quad (8)$$

Here the highest coefficient is 1 because both u, f are monic. For any $c \in K$, we have

$$\begin{aligned} fc^\mu &= u^m c^\mu + u^{m-1} a_1 c^\mu + \dots + a_m c^\mu, \\ cf &= u^m c^{\lambda^m} + u^{m-1} c^{\lambda^{m-1}} a_1 + \dots + c a_m. \end{aligned}$$

Equating powers of u we see that $\mu = \lambda^m$ and

$$ca_i = a_i c^{\lambda^i} \quad (i = 1, \dots, m). \quad (9)$$

Let α^d be the least power which is inner; then $\lambda^i = \alpha^{vi}$ is inner if and only if

$d \mid v_i$, so by (9), $a_i = 0$ unless $d \mid v_i$. Hence we either have $f = u^m$ or f is a polynomial in $u^{d'}$, where $d' = d/(d, v)$. By hypothesis, $cu^{d'} = u^{d'} \cdot ece^{-1}$ for some $e \in K^\times$, so $v = u^{d'}e$ centralizes K and f can be written as a polynomial in v , say $f = \sum v^j b_j$. We now have

$$0 = cf - fc^u = \sum v^j (cb_j - b_j c^u),$$

hence $c^u = b_j^{-1} c b_j$. So we have $f = f_1 \cdot b$, where f_1 centralizes K and b is an element of K inducing a power of α . ■

Exercises 8.3

1. Let $R = K[t; \alpha, \delta]$ be a skew polynomial ring and θ an automorphism of K commuting with α . Find conditions under which θ can be extended to an automorphism of R . Examine the case $\theta = \alpha$.

2. Let $R = K[t; \alpha, \delta]$ be a skew polynomial ring. Find conditions for an anti-automorphism of R to exist which maps t to $-t$ and fixes K . (Hint: Note that K has to be commutative.)

3. Let K be a field with an automorphism α , no power of which is inner. Show that for any α -derivation δ , the centralizer of K in $K(t; \alpha, \delta)$ is contained in the centre of K .

4. Let K be a commutative field with an automorphism α , whose fixed field is C . Determine the centre of $K(t; \alpha)$, (i) when α has infinite order and (ii) when α has finite order.

5. Show that in the skew polynomial ring $K[t; 1, \delta]$ the monic element f is right invariant if and only if it is central. Deduce that a monic invariant element in $K[t]$ has central coefficients.

6. Let $R = K[t; \alpha, \delta]$ be a skew polynomial ring in which the monoid of invariant elements is generated by d . Let M be the submonoid generated by d^m ; determine the ideal structure of R_M .

7. Let A be a square matrix over a skew polynomial ring R and let $\text{diag}(a_1, \dots, a_n)$ be its canonical diagonal form (cf. Th. 1.1). To what extent is the product $a_1 \dots a_n$ unique? Use this product to define a notion of determinant; examine the particular case when R is simple.

8. Let $R = K[t; \alpha, \delta]$ be a skew polynomial ring, where $\alpha\delta = \delta\alpha$ and $\text{char } K = 0$. If f is right invariant, given by (3), show that

$$nc^{\alpha^{n-1}\delta} = a_1 c^{\alpha^n} - c^{\alpha^{n-1}} a_1.$$

Deduce that δ is inner on $\text{im } \alpha^{n-1}$.

9. Let K be the universal field of fractions of the free algebra $k\langle X \rangle$, where k is a commutative field of characteristic 0 and $X = \{x_j \mid j = 1, 2, \dots\}$. Define α as endo-

morphism by $x_{ij}^\alpha = x_{i+1,j}$, and δ as α -derivation by $x_{ij}^\delta = x_{i,j+1}$. Verify that $\alpha\delta = \delta\alpha$, and show that if δ were inner on $\text{im } \alpha^{n-1}$, induced by a , then

$$nx_{ij} = ax_{i+1,j-1} - x_{i,j-1}a;$$

for all i, j . Show that no such a exists and deduce that R is a simple right but not left principal ideal domain.

10. Let K be the universal field of fractions of $k\langle x_1, x_2, \dots \rangle$ and define α, δ by $x_i^\alpha = x_{i+1}$, $x_i^\delta = x_{2i}$. Show that $\delta\alpha^n - \alpha^n\delta$ is outer for all $n > 0$, and deduce that $K[t; \alpha, \delta]$ is simple right but not left principal.

11. Let K be a field of finite dimension over its centre, and let α be an endomorphism of K . Show that every α -derivation of K is inner, unless α is an inner automorphism of K . (Hint: Use the proof of Prop. 3.1.)

12°. Find the ideal structure of $K[t; \alpha, \delta]$ when α is not surjective.

8.4 Theory of a single pseudo-linear transformation

Let K be a field with an automorphism α and an α -derivation δ . Given a right K -space V , we define a mapping $\theta: V \rightarrow V$ to be *pseudo-linear* (relative to (α, δ)), if

$$(x+y)\theta = x\theta + y\theta, \quad (xa)\theta = x\theta \cdot a^\alpha + x \cdot a^\delta \quad (x, y \in V, a \in K). \quad (1)$$

This includes the linear case when $\alpha = 1$, $\delta = 0$. Other important examples are the *semilinear* transformations (where $K = \mathbb{C}$, $\delta = 0$ and α is conjugation) and *differential* transformations ($K = F(t)$, $\alpha = 1$ and $\delta = d/dt$).

To obtain a matrix representation of θ , assuming V finite-dimensional, take a basis $v_1, \dots, v_n$ of V . If

$$v_j\theta = \sum v_i a_{ij}, \quad \text{where } a_{ij} \in K, \quad (2)$$

and we put $A = (a_{ij})$, then

$$(\sum v_j \xi_j)\theta = (\sum v_j \theta) \xi_j^\alpha + \sum v_j \xi_j^\delta = \sum v_i (\sum a_{ij} \xi_j^\alpha + \xi_i^\delta),$$

hence on writing $\xi = (\xi_1, \dots, \xi_n)^T$, we have

$$\theta: \xi \mapsto A\xi^\alpha + \xi^\delta. \quad (3)$$

Conversely, every $n \times n$ matrix A over K defines a pseudo-linear transformation θ on V by (3). Let us write the basis in V as a row vector $v = (v_1, \dots, v_n)$, so that the general vector $x \in V$ takes the form $x = v\xi$. If we change the basis in V , say $x = v\xi = v'\xi'$, where $v' = vP$, $\xi' = P^{-1}\xi$ and $P \in \text{GL}_n(K)$, then $v\theta = vA$, $v'\theta = v'B$, hence $vPB = vP\theta = v\theta P^\alpha + vP^\delta = v(AP^\alpha + P^\delta)$, so we obtain

$$PB = AP^\alpha + P^\delta. \quad (4)$$

Two matrices A, B related as in (4), where $P \in GL_n(K)$, are said to be (α, δ) -conjugate. Thus a given pseudo-linear transformation is represented in different coordinate systems by matrices that are mutually (α, δ) -conjugate, and our task will be to obtain a canonical form for a given matrix under (α, δ) -conjugacy. For $\alpha = 1, \delta = 0$ this includes the canonical form for a square matrix over a field under conjugacy: $B = P^{-1}AP$. Of course, the results in the general case are not as complete as in the classical case.

Writing $R = K[t; \alpha, \delta]$, we can define our space V as a right R -module by letting $\Sigma t^i a_i$ correspond to $\Sigma \theta^i a_i$; the equations (1) ensure that this is a homomorphism. Taking again a right K -basis $v = (v_1, \dots, v_n)$ of V , we find that V is generated as right R -module by $v_1, \dots, v_n$ and satisfies the relations $\Sigma v_i(a_{ij} - \delta_{ij}t) = 0$, or in matrix form

$$v(A - tI) = 0. \quad (5)$$

In fact, this is a complete set of defining relations for V , as is easily checked. Now by the diagonal reduction in R there exist $P, Q \in GL_n(R)$ such that

$$P(A - tI)Q = \text{diag}(\lambda_1, \dots, \lambda_n), \quad \lambda_{i-1} \parallel \lambda_i. \quad (6)$$

The polynomials λ_i in t are called the *invariant factors* of A . It follows that as right R -module, V is isomorphic to

$$R/\lambda_1 R \oplus \dots \oplus R/\lambda_n R. \quad (7)$$

We observe that $R/\lambda R = 0$ if and only if λ is a non-zero constant and $R/\lambda R \cong R$ if and only if $\lambda = 0$. Moreover, $\dim_K(R/\lambda R) = \deg \lambda$, hence $\Sigma \deg \lambda_i = n$. Denoting the centre of K by C , we see that $A - tI$ can become singular for at most n values of t in C . An element c of C such that $A - cI$ is singular is called a *central eigenvalue* of A . With this terminology we have, by Lemma 1.3 and its proof,

PROPOSITION 4.1. *An $n \times n$ matrix over a field has at most n central eigenvalues ■*

The expression (7) shows that V is a direct sum of cyclic R -modules, and as minimal decomposition into cyclic R -modules, this expression is unique up to isomorphism. We ask: when is V itself cyclic? By (7) this is the case when there is just one invariant factor that is not a unit. In this case the pseudo-linear transformation θ , or also the matrix A representing it, is called *cyclic*. More precisely it should be called (α, δ) -cyclic; for $\alpha = 1, \delta = 0$ this reduces to the usual notion of a cyclic (or non-derogatory) matrix. From (6) we obtain the following characterization:

PROPOSITION 4.2. *Let K be a field with endomorphism α and α -derivation δ .*

Then a square matrix A over K is (α, δ) -cyclic if and only if $tI - A$ is stably associated over $K[t; \alpha, \delta]$ to a 1×1 matrix. ■

For an (α, δ) -cyclic pseudo-linear transformation θ we have a normal form as in the commutative case. Let v be a generator of V as right R -module and put $v_i = v\theta^{i-1}$, then V is spanned over K by $v_1, v_2, \dots$ and if v_{r+1} is the first of these vectors linearly dependent on the preceding ones over K , let

$$v_{r+1} = v_1 a_1 + \dots + v_r a_r, \quad \text{where } a_i \in K. \quad (8)$$

By applying θ we obtain $v_{r+2} = (\sum v_i a_i) \theta = \sum v_{i+1} a_i^\alpha + \sum v_i a_i^\delta$, and this shows v_{r+2} to be linearly dependent on $v_1, \dots, v_r$. By induction the same is true for v_{r+v} ($v = 1, 2, \dots$), and since the v_i span V , we conclude that $r = n$ and $v_1, \dots, v_n$ is a basis for V over K . Relative to this basis θ has the matrix

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 & a_1 \\ 1 & 0 & 0 & \dots & 0 & a_2 \\ 0 & 1 & 0 & \dots & 0 & a_3 \\ & \dots & & \dots & & \\ 0 & 0 & 0 & \dots & 1 & a_n \end{pmatrix}. \quad (9)$$

This matrix is called the *companion matrix* of the polynomial

$$f = t^n - a_1 t^{n-1} - a_2 t^{n-2} - \dots - a_n. \quad (10)$$

We remark that $tI - A$ is stably associated to f , and note that (8) may be expressed in the form $vf(\theta) = 0$.

By (7), every square matrix over K is (α, δ) -conjugate to a diagonal sum of matrices of the form (9). We thus obtain an analogue of the rational canonical form; to obtain the full Jordan normal form one needs to make further assumptions about the field K , but we shall not pursue the matter here (cf. Cohn [73, 77]). We also note that the Cayley-Hamilton theorem no longer holds in general: if θ is a pseudo-linear transformation, and V the associated R -module, then the minimal polynomial of θ is the generator of the two-sided ideal of R annihilating V , and this may well be 0, even when V is finite-dimensional over K . In fact, as we see from (7), θ has a minimal polynomial if and only if its last invariant factor is bounded; in that case we also say that θ and the matrix representing it are *bounded*. The bound for the last invariant factor is then the minimal polynomial. For example, the ring of differential transformations (in characteristic 0) is simple, hence the only matrices that are bounded are the invertible matrices. We shall find that there is a relative notion that is more useful; this will be discussed in the next section.

Exercises 8.4

1. Find the normal form for aI , where a is a non-central element of K .
2. Let $A, B \in K_n$; show that A is (α, δ) -conjugate to B if and only if $tI - A$ is associated to $tI - B$ over $K[t; \alpha, \delta]$ (cf. Prop. 0.7.5).
3. Show that a system of linear differential equations in several unknown functions is equivalent to a single linear differential equation (use the simplicity of the ring of differential operators). What form does the system take if we apply the reduction to the Jordan normal form?
4. Show that a pseudo-linear transformation θ is bounded if and only if the kernel of the homomorphism from $R = K[t; \alpha, \delta]$ to $K[\theta]$ is a bounded ideal in R , and express this ideal in terms of the invariant factors of θ .
5. Show that $tI - A$, where A is given by (9) is stably associated over $K[t; \alpha, \delta]$ to f , where f is given by (10).
6. (Jacobson [37]) If θ is a bounded indecomposable pseudo-linear transformation, show that its minimal polynomial is a power of an I -atom.
7. (Jacobson [37]) Show that a bounded pseudo-linear transformation θ is completely reducible if and only if its minimal polynomial is a product of distinct I -atoms.

8.5 Eigenvalues and algebraic matrices

Let R be a k -algebra; a square matrix A over R is said to be *algebraic* over k if there is a polynomial $f \in k[t]^\times$ such that $f(A) = 0$. If for every non-zero polynomial f over k , $f(A)$ is regular, A is called *transcendental* over k ; this term will mainly be used when R is itself a field. In general a matrix is neither transcendental nor algebraic, but we always have a decomposition; to derive it we need a lemma.

LEMMA 5.1. *Let R, S be k -algebras and M an (R, S) -bimodule. Given $a \in R$, $b \in S$, assume that there is a polynomial f over k such that $f(a)$ is a unit, while $f(b) = 0$. Then for any $m \in M$, the equation*

$$ax - xb = m \quad (1)$$

has a unique solution $x \in M$.

Proof. In $\text{End}_k(M)$ write $\lambda_a: x \mapsto ax$, $\rho_b: x \mapsto xb$; then (1) may be written

$$x(\lambda_a - \rho_b) = m. \quad (2)$$

By hypothesis $f(\lambda_a)$ is a unit, $f(\rho_b) = 0$ and λ_a, ρ_b commute. Hence if we define $\phi(s, t)$ in commuting variables s, t by

$$\phi(s, t) = \frac{f(s) - f(t)}{s - t},$$

then $\phi(\lambda_a, \rho_b)(\lambda_a - \rho_b) = (\lambda_a - \rho_b)\phi(\lambda_a, \rho_b) = f(\lambda_a) - f(\rho_b) = f(\lambda_a)$, and this is a unit, hence (2) and with it (1) has a unique solution in M . ■

The result may be restated in matrix form. Consider the matrix ring

$\begin{pmatrix} R & M \\ 0 & S \end{pmatrix}$; given a matrix

$$\begin{pmatrix} a & u \\ 0 & b \end{pmatrix}, \quad (3)$$

where $f(a)$ is a unit and $f(b) = 0$, we can find a conjugate of (3) in diagonal

form. For if we transform (3) by $\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$, we obtain

$$\begin{pmatrix} a & u + ax - xb \\ 0 & b \end{pmatrix}$$

and by the lemma the north-east block is 0 for a suitable choice of x .

PROPOSITION 5.2. *Let K be a skew field which is a k -algebra. Then every square matrix A over K is conjugate to a diagonal sum of an algebraic and a transcendental matrix.*

Proof. Let $A \in K_n$ and consider $V = {}^nK$ as a $(K, k[t])$ -bimodule, in which the action of t for a given K -basis $u_1, \dots, u_n$ is given by

$$u_i t = \sum a_{ij} u_j, \quad A = (a_{ij}).$$

Since $K \otimes k[t] = K[t]$ is a principal ideal domain, V has a unique torsion submodule V_0 with torsion free quotient, which, being finitely generated, is free. Let V_1 be a complement of V_0 , so that

$$V = V_0 \oplus V_1. \quad (4)$$

Using a basis of V adapted to the decomposition (4), we find that A takes the form

$$\begin{pmatrix} A_0 & A' \\ 0 & A_1 \end{pmatrix},$$

where A_0 is algebraic and A_1 is transcendental. By Lemma 5.1 and the remark following it we can reduce A' to 0 and so obtain the desired conclusion. ■

We can get a criterion for algebraicity from the normal form obtained in 8.4. As we have seen there, if $A \in K_n$, then $tI - A$ is associated to $\text{diag}(\lambda_1, \dots, \lambda_n)$, where $\lambda_{i-1} \parallel \lambda_i$. It is clear from this that A is algebraic if and only if λ_n divides a polynomial with coefficients in k . When k is the precise centre of K , this leads to a criterion for A to be algebraic or transcendental.

THEOREM 5.3. *Let K be a field with centre k and let $A \in K_n$ have invariant factors $\lambda_1, \dots, \lambda_n$. Then (i) A is algebraic over k if and only if λ_n is bounded and (ii) A is transcendental over k if and only if λ_n is totally unbounded, and then $\lambda_1 = \dots = \lambda_{n-1} = 1$.*

Proof. We have seen that a polynomial f over K is invariant if and only if it is associated to a monic polynomial over the centre k . Thus A is algebraic if and only if λ_n divides an invariant polynomial, i.e. is bounded (cf. 6.4). To prove (ii), suppose that λ_n is not totally unbounded, say it has a bounded factor p , with bound p^* . Then the module V defined by A has a non-zero element annihilated by p and hence by p^* . Now p^* is invariant, hence with coefficients in k and $p^*(A)$ is singular, so A cannot be transcendental. Conversely, if A is not transcendental, then V has an element $\neq 0$ annihilated by an invariant polynomial, so some invariant factor λ_i has a non-unit factor which is bounded, hence λ_n has then a bounded factor. Thus A is transcendental if and only if λ_n is totally unbounded, and then no other λ_i can be a non-unit, because this would give rise to a non-unit invariant element dividing λ_n . ■

As a consequence we can now establish the claim made in 4.6.

COROLLARY 5.4. *A skew field K which is a k -algebra is matrix algebraic over k if and only if its centre is algebraic over k and every non-zero polynomial over K is bounded.*

Proof. Let C be the centre of K ; if K is matrix algebraic over k , then K is algebraic over k , and so is C . Further, any polynomial $f \neq 0$ is the sole invariant factor of its companion matrix B_f . By hypothesis B_f is algebraic, hence by Th. 5.3, f is bounded. Conversely, when the conditions hold, take any square matrix over K ; all its invariant factors are bounded, so A is algebraic over C and hence over k . ■

The minimal polynomial of a matrix over the centre of the field of coefficients has another more explicit description.

PROPOSITION 5.5. *Let K be a field with centre C and let $A \in K_n$ be algebraic over C , with minimal polynomial μ , which is irreducible over C . Then $\mu \cdot I$ is the least common right multiple of the $tI - P^{-1}AP$ for all $P \in \text{GL}_n(K)$,*

$$\mu(t) \cdot I = (tI - A_1) \dots (tI - A_r), \quad (5)$$

where $A_1 = A, A_2, \dots, A_r$ are certain conjugates of A , and the A_i may be permuted cyclically in (5).

Proof. By hypothesis, $\mu(A) = 0$, hence $\mu(P^{-1}AP) = 0$ for any $P \in \text{GL}_n(K)$, and reversing the argument we see that $P^{-1}AP$ has the same minimal polynomial μ as A . By the division algorithm for K_n we can write

$$\mu(t) \cdot I = (tI - A)B \quad \text{for some } B \in K[t]_n,$$

where the remainder is 0, by hypothesis. Similarly for $P^{-1}AP$, so $\mu \cdot I$ is a common right multiple for all the $tI - P^{-1}AP$. Let M be their least common right multiple; then

$$\mu \cdot I = MF, \quad \text{where } F \in K[t]_n.$$

Here the coefficient of the highest power of t in M is a unit and we can modify M so that this highest coefficient becomes 1. This determines M uniquely, and it is clearly unchanged by conjugation, i.e. it lies in the centre of the matrix ring $K[t]_n$, so we have $M = m(t) \cdot I$, where $m(t) \in C[t]$. By the minimality of μ we must have $m = \mu$, and the first assertion is proved.

To establish (5), let us write

$$\mu \cdot I = (tI - A_1) \dots (tI - A_r)M_2 = M_1M_2, \text{ say,} \quad (6)$$

where A_i is conjugate to $A_1 = A$ and $r \geq 1$ is chosen as large as possible. If M_2 is not a unit, there exists $A' = P^{-1}AP$ such that $tI - A'$ is not a left factor of M_1 , hence there is a comaximal relation

$$(tI - A')M'_1 = M_1(tI - A_{r+1}). \quad (7)$$

By Prop. 0.7.5, A_{r+1} is conjugate to A' , hence to A . Now $\mu \cdot I = (tI - A)B = M_1M_2$ and since (7) is comaximal, we have $M_2 = (tI - A'_{r+1})N$, but this contradicts the choice of r . Thus M_2 is a unit and on comparing leading terms in (6) we see that $M_2 = I$ and (5) follows.

Finally, if we split the product on the right of (5) in any way, say $\mu I = M_1M_2$, then $\mu M_1 = M_1M_2M_1$ and on cancelling the regular factor M_1 we have $\mu I = M_2M_1$. This shows that the factors in (5) can be permuted cyclically. ■

COROLLARY 5.6. *Two matrices over a field K which are algebraic over the centre C have the same minimal equation (assumed irreducible) over C if and only if they are conjugate.*

For if A, B have the same minimal polynomial μ , then

$$\mu \cdot I = (tI - A_1) \dots (tI - A_r) = (tI - B_1) \dots (tI - B_r),$$

where $A_1 = A, B_1 = B$, say. By unique factorization $tI - B$ is stably associated to some $tI - A_i$, hence B is conjugate to A_i and so is conjugate to A . The converse is clear. ■

We note that this corollary also follows from the Skolem-Noether theorem (cf. e.g. Cohn [77], p. 364). Two transcendental matrices need not, of course, be conjugate; however, it can be shown that they become conjugate when the field is suitably extended (cf. Cohn [77] or Ex. 7.5.19).

We now turn to the notion of eigenvalue, which will help to elucidate the connexion between matrices and equations.

Let K be a field and $A \in K_n$ (for some $n \geq 1$). An element α of K is called a *right eigenvalue* of A if there exists $u \in {}^nK, u \neq 0$, such that

$$Au = u\alpha.$$

Such a vector u is called an *eigenvector* for α ; the set of all right eigenvalues of A is called the *right spectrum* of A . Similarly a *left eigenvalue* of A is an element $\beta \in K$ such that $vA = \beta v$ for some $v \in K^n, v \neq 0$. The set of all such β is called the *left spectrum* of A , and by the *spectrum* of A , $\text{spec } A$, we understand the union of the left and right spectra. An element λ of K is called a *singular eigenvalue* of A if $A - \lambda I$ is singular; these λ are in general distinct from the left and right eigenvalues, and they do not share the properties described below. Of course, it is clear that a central eigenvalue, as defined in 8.4, is left, right and singular.

The importance of left and right eigenvalues is that they are invariant under similarity:

PROPOSITION 5.7. *The (left, right) spectrum of a matrix A over a field K consists of complete conjugacy classes of K and is a similarity invariant of A .*

Proof. Let α be a right eigenvalue of A , say $Au = u\alpha$, then for any $c \in K^\times$ we have $Auc = u\alpha c = uc \cdot c^{-1}\alpha c$, hence $c^{-1}\alpha c$ is again a right eigenvalue of A .

Secondly, let $P \in \text{GL}_n(K)$; then $PAP^{-1} \cdot Pu = PAu = Pu \cdot \alpha$, so α is also a right eigenvalue of PAP^{-1} . Similarly for left eigenvalues, and by combining these results, we find the spectrum of A to be a similarity invariant. ■

Since conjugacy classes are usually infinite, we see that the number of eigenvalues of a matrix is no longer finite; but as in the commutative case, there is a bound on the number of conjugacy classes of left or right eigenvalues.

PROPOSITION 5.8. *Let A be a matrix over a field K . Then the eigenvectors belonging to inconjugate right eigenvalues are right linearly independent over K . If α is a right and β a left eigenvalue of A and α, β are not conjugate, then the eigenvectors belonging to them are orthogonal, i.e. if $Au = u\alpha$, $vA = \beta v$, then $vu = 0$.*

Proof. Let $\alpha_1, \dots, \alpha_r$ be right eigenvalues and $u_1, \dots, u_r$ corresponding eigenvectors and assume that the u 's are linearly dependent. A minimal relation between them will have the form

$$u_1 = \sum_2^r u_i \lambda_i, \quad \text{where } \lambda_i \in K.$$

By definition $u_1 \neq 0$, hence by minimality $\lambda_i \neq 0$ and $r > 1$. Now $u_1 \alpha_1 = Au_1 = \sum_2^r Au_i \lambda_i = \sum_2^r u_i \alpha_i \lambda_i$; hence $\sum_2^r u_i (\lambda_i \alpha_1 - \alpha_i \lambda_i) = 0$. Since $u_2, \dots, u_r$ are linearly independent (by minimality), we have $\alpha_i = \lambda_i \alpha_1 \lambda_i^{-1}$ and so the α_i are all conjugate.

Next, if $Au = u\alpha$, $vA = \beta v$, then $vAu = vu \cdot \alpha = \beta \cdot vu$, and if $vu \neq 0$, this would mean that α, β are conjugate. ■

In the commutative case an $n \times n$ matrix with n distinct eigenvalues can be transformed to diagonal form. This result also has an analogue in the general case:

THEOREM 5.9. *Let K be a field and $A \in K_n$. Then $\text{spec } A$ cannot contain more than n conjugacy classes, and when it consists of exactly n classes, all except at most one algebraic over the centre of K , then A is similar to a diagonal matrix.*

Proof. We have seen that $\text{spec } A$ consists of conjugacy classes. Let r be the number of classes containing right eigenvalues and s the remaining classes in $\text{spec } A$; then the space spanned by the columns corresponding to the right eigenvalues is at least r -dimensional and the space of rows orthogonal to this space is at least s -dimensional, hence $r + s \leq n$, and $r + s$ is just the number of conjugacy classes.

Assume now that $r + s = n$; let $\alpha_1, \dots, \alpha_r$ be inconjugate right eigenvalues and $u_1, \dots, u_r$ the corresponding eigenvectors, while $\beta_1, \dots, \beta_s$ are the left

eigenvalues not conjugate among themselves or to the α 's, with corresponding eigenvectors $v_1, \dots, v_s$. By Prop. 5.8 the u 's are right linearly independent, the v 's are left linearly independent and $v_j u_i = 0$ for all i, j . Write U_1 for the $n \times r$ matrix consisting of the columns $u_1, \dots, u_r$ and V_2 for the $s \times n$ matrix consisting of the rows $v_1, \dots, v_s$. Since the columns of U_1 are linearly independent, we can find an $r \times n$ matrix V_1 over K such that $V_1 U_1 = I$, and similarly there is an $n \times s$ matrix U_2 such that $V_2 U_2 = I$. If we put $U = (U_1, U_2)$, $V = (V_1, V_2)^T$, then we have

$$VU = \begin{pmatrix} V_1 U_1 & V_1 U_2 \\ V_2 U_1 & V_2 U_2 \end{pmatrix} = \begin{pmatrix} I & W \\ 0 & I \end{pmatrix}.$$

The matrix on the right is invertible, hence

$$U(VU)^{-1} = (U_1, U_2 - U_1 W) = V^{-1},$$

and so

$$\begin{aligned} AV^{-1} &= A(U_1 \ U_2 - U_1 W) = (u_1 \alpha_1, \dots, u_r \alpha_r \ A(U_2 - U_1 W)) \\ VA &= (V_1 A \ \beta_1 v_1, \dots, \beta_s v_s)^T. \end{aligned}$$

It follows that $VA V^{-1} = \begin{pmatrix} \alpha & T \\ 0 & \beta \end{pmatrix}$, where $\alpha = \text{diag}(\alpha_1, \dots, \alpha_r)$, $\beta = \text{diag}(\beta_1, \dots, \beta_s)$ and $T \in {}^r K^s$. Now all the α 's and β 's are inconjugate and all but at most one are algebraic over the centre of K , hence their minimal equations are distinct, by Cor. 5.6. If only right or only left eigenvalues occur, we have diagonal form. Otherwise let $\beta_1, \dots, \beta_s$ be algebraic, say. Taking f to be the product of their minimal polynomials, we have $f(\beta) = 0$, while $f(\alpha)$ is a unit. By Lemma 5.1, we can find $X \in {}^r K^s$ such that $\alpha X - X\beta = T$, and transforming our matrix by $\begin{pmatrix} I & X \\ 0 & I \end{pmatrix}$, we reach diagonal form. ■

The restriction on the eigenvalues, that there is to be at most one transcendental conjugacy class, is not as severe as appears at first sight, but is to be expected because, as remarked earlier, K can be extended so that all transcendental elements are conjugate in the extension.

We have seen in 8.4 that every square matrix is conjugate to a direct sum of cyclic matrices (taking now $\alpha = 1$, $\delta = 0$), and that any cyclic matrix is conjugate to a matrix of the form (9) of 8.4. For such a matrix we have the following description of the eigenvalues:

THEOREM 5.10. *Let K be a field and let $A \in K_n$ be a cyclic matrix, of the form*

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 & a_1 \\ 1 & 0 & 0 & \dots & 0 & a_2 \\ 0 & 1 & 0 & \dots & 0 & a_3 \\ & \dots & \dots & & & \\ 0 & 0 & 0 & \dots & 1 & a_n \end{pmatrix}. \quad (8)$$

Then the right eigenvalues of A coincide with the conjugates of the roots of the equation

$$x^n - a_1 - a_2x - \dots - a_nx^{n-1} = 0, \quad (9)$$

and the left eigenvalues of A coincide with the conjugates of the roots of the equation

$$x^n - a_1 - xa_2 - \dots - x^{n-1}a_n = 0. \quad (10)$$

Proof. Let e_i ($i = 1, \dots, n$) be the row vectors of the unit matrix; then $e_1A = a_1e_n$, $e_2A = e_1 + a_2e_n$, ..., $e_nA = e_{n-1} + a_ne_n$. Now if β is a left eigenvalue of A , say $vA = \beta v$, where $v = \sum c_i e_i \neq 0$, then $(\sum c_i e_i)A = \beta \sum c_i e_i$, hence

$$\sum c_i e_{i-1} + \sum c_i a_i e_n = \sum \beta c_i e_i \quad (e_0 = 0),$$

and equating coefficients of e_i , we find

$$\beta c_i = c_{i+1} \quad (i = 1, \dots, n-1), \quad \beta c_n = \sum c_i a_i. \quad (11)$$

Suppose that $c_1 = 0$. Then (11) shows that $c_2 = \dots = c_n = 0$, so $v = 0$, a contradiction. Hence $c_1 \neq 0$, and on multiplying by an appropriate scalar we may assume that $c_1 = 1$. This changes β to a conjugate β_1 and now (11) shows that $c_i = \beta_1^{i-1}$ and

$$\beta_1^n - a_1 - \beta_1 a_2 - \dots - \beta_1^{n-1} a_n = 0,$$

so that β_1 satisfies (10).

Conversely, if β is a root of (10), put $v = \sum \beta^{i-1} e_i$, then $vA = (\sum \beta^{i-1} e_i)A = \sum \beta^{i-1} (e_{i-1} + a_i e_n) = \sum_{i=1}^{n-1} \beta^i e_i + \beta^n e_n = \beta v$, so β is a left eigenvalue of A .

Now the result for right eigenvalues follows by symmetry, because the matrix (8) is conjugate to its transpose, by the transforming matrix

$$\begin{pmatrix} -a_2 & -a_3 & -a_4 & \dots & -a_n & 1 \\ -a_3 & -a_4 & -a_5 & \dots & 1 & 0 \\ & \dots & \dots & \dots & & \\ -a_n & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}. \quad \blacksquare$$

Consider the roots of an equation

$$\sum x^i a_i = 0 \quad (12)$$

in K ; they are also called the *right zeros* of the polynomial $f = \sum x^i a_i$. We may take the polynomial monic, without loss of generality, and if A is its companion matrix, then by Th. 5.10, the roots of (12) are left eigenvalues of A . By Th. 5.9 we therefore obtain

COROLLARY 5.11. *The right zeros of a polynomial of degree n , in any field K , lie in at most n conjugacy classes of K . ■*

This will be proved again in the next section, in a more direct fashion.

Exercises 8.5

1. Let K be a field which is a k -algebra. Show that $K \otimes k(t) \subseteq K(t)$, with equality if and only if K is matrix algebraic over k .

2. Let A be a square matrix over a k -algebra R . If A is algebraic over k , show that A is conjugate over R to a diagonal sum of terms: $B_1 \oplus \dots \oplus B_r$, where each B_i has a power of an irreducible polynomial as minimal polynomial.

3. Let K be a field and p^* an I -atom of length r in a skew polynomial ring over K . Show that $r \mid \deg p^*$; further, show that $r^2 \mid \deg p^*$ if the eigenring of each atomic factor of p^* contains K .

4. Let K be a field, $R = K[t]$ and f a polynomial with no central factors, i.e. no factors (apart from 1) with all coefficients in the centre of R . Show that the ideal generated by f is the whole of R .

5. Verify that any left or right eigenvalue of a nilpotent matrix is 0. Show that a nilpotent matrix can have a non-zero singular eigenvalue.

6. Show directly that the right eigenvalues of the matrix (8) coincide with the conjugates of the roots of (9).

7. Let D be a finite-dimensional central simple algebra and ρ its regular representation. If F is a commutative subfield of D , show that $\rho(a)$ has an eigenvalue in F if and only if a is conjugate to an element of F . Deduce that every matrix over the real quaternions has left and right eigenvalues, and that every polynomial with quaternion coefficients has left and right zeros.

8°. (Cohn [77]) Is it true that every square matrix over a field K has a singular eigenvalue in some extension field of K ?

9°. (Cohn [77]) Let K be a field with centre C . Show that if every non-constant

equation over K has a solution in some extension, then C is relatively algebraically closed in K . (Hint: Examine the 'metro-equation' $ax - xa = 1$, where a is algebraic over C .)

10°. Let K be a field which is a k -algebra, where k is relatively algebraically closed in K . Is it true that every square matrix that is not conjugate over k to a triangular matrix has a non-zero singular eigenvalue in some extension field of K ?

11. Let D be a field of characteristic 0, matrix algebraic over its centre C . Show that the derivation $f \mapsto f' (= df/dx)$ of $D[x]$ is outer; more precisely, show that the derivation $f \mapsto f' - [f, g]$ (for fixed $g \in D[x]$) has kernel C .

8.6 Algebraic skew field extensions

As an application of the preceding ideas let us consider algebraic extensions of skew fields.

Let E be a field with a subfield K . An element $c \in E$ is said to be *right algebraic* over K if the right K -module generated by the powers of c is finitely generated over K . Clearly this is equivalent to the condition that there is a monic polynomial $f = \sum_1^n x^i a_i$ ($a_i \in K$, $a_n = 1$) such that $\sum c^i a_i = 0$, or also that f has the form

$$f = (x - c) \cdot g. \quad (1)$$

In other words, c is a right zero of f . The ring $E[x]$ is a principal ideal domain, as is its subring $K[x]$, and if we take a complete factorization of f in $K[x]$:

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

we know that $r \leq \deg f$. Moreover, any atomic factor of f is similar to one of the factors on the right of (2). Now two similar elements of $K[x]$ have the same degree, for the degree of $f \in R = K[x]$ may be characterized as the dimension of the right K -space R/fR . From Prop. 0.7.5 we see that $x - a$ and $x - b$ are similar in R if and only if a, b are conjugate in K . Applying the result to a polynomial of degree n , we again obtain

THEOREM 6.1. *The right zeros of a polynomial of degree n , in a field K , lie in at most n different conjugacy classes of K . ■*

To determine the number of zeros in a given conjugacy class we use Th. 4.2.4; this tells us that the number of left factors of f similar to $x - a$ is (up to associates) 0, 1 or at least α , where α is the cardinal of the eigenring of $x - a$ in $K[x]$. But the latter is just the centralizer of a in K , by Cor. 0.7.4. Thus we have

THEOREM 6.2. *Let K be a field, $f \in K[x]$, and let a be a right zero of f in some extension E of K . Denote by C_a the centralizer of a in E . Then the number of right zeros of f conjugate to a in E is either 1 or at least $|C_a|$. ■*

From this result it is easy to deduce that the number of conjugates of a , if greater than 1, is infinite. We need only show that in any non-commutative field every element has an infinite centralizer. This is clear for elements that are transcendental over the centre; in general it follows from

LEMMA 6.3. *Let K be a field with finite centre C . If $C \neq K$, then every element of K commutes with an element transcendental over C .*

Proof. It will be enough to prove the result for elements of K that are not in C , and clearly we need only consider elements algebraic over C . Thus let $a \in K$ be algebraic over C but not in C ; then $C(a)$ is again finite, with q^r elements, where $q = |C|$ and $r > 1$. The mapping $x \mapsto x^q$ is an automorphism of $C(a)$ over C , hence a and a^q satisfy the same irreducible equation over C , and so are conjugate, by Cor. 5.6. Hence there exists $b \in K$ such that

$$b^{-1}ab = a^q. \quad (3)$$

Since $a \notin C$, $a^q \neq a$ and therefore $ab \neq ba$. Consider the field $A = C(a, b^r)$; by (3), b^r induces the automorphism $x \mapsto x^{q^r}$ on $C(a)$. This is the identity, so A is commutative. If b^r is algebraic over C , then A as algebraic extension of C is finite and by (3), $A(b)$ is also finite, but non-commutative, as we have seen. This contradicts Wedderburn's theorem (that all finite fields are commutative), hence b^r is transcendental over C , and it commutes with a . ■

We note that as a consequence of this lemma, if a non-commutative field is algebraic over its centre, then the centre must be infinite.

To find the relation between the right zeros of a polynomial and its factors, let $f(x) = \sum x^i a_i$, $g(x) = \sum x^j b_j$; then $fg = \sum x^{i+j} a_i b_j$, hence on writing $f(c)$ or f_c for the result of replacing x by c in f , keeping the coefficients on the right, we obtain

$$(fg)_c = \sum c^{i+j} a_i b_j = \sum c^j f_c b_j = f_c \cdot \sum (f_c^{-1} c f_c)^j b_j,$$

and so

$$(fg)_c = f_c \cdot g(f_c^{-1} c f_c), \quad (4)$$

provided only that $f_c \neq 0$. We thus obtain

LEMMA 6.4. *For any $f, g \in K[x]$ and $c \in K$,*

$$(fg)_c = \begin{cases} 0 & \text{if } f_c = 0, \\ f_c \cdot g(f_c^{-1}cf_c) & \text{if } f_c \neq 0. \end{cases}$$

Hence the right zeros of fg are either right zeros of f or conjugates of right zeros of g . ■

With the help of this result it is easy to construct polynomials with prescribed zeros:

PROPOSITION 6.5. *Let $c_1, \dots, c_n \in K$ be pairwise inconjugate. Then there is a unique monic polynomial of degree n with $c_1, \dots, c_n$ as right zeros.*

Proof. If f, g are two monic polynomials of degree n with the c 's as right zeros, then $f-g$ is a polynomial of degree less than n with the c 's as right zeros, and so must vanish by Th. 6.1. Thus there can be at most one such polynomial. To find it we may, by induction on n , assume that g is monic with right zeros $c_2, \dots, c_n$. Then $g(c_1) \neq 0$, by Th. 6.1, hence by Lemma 6.4 $g \cdot (x - g(c_1)^{-1}c_1g(c_1))$ is a monic polynomial of degree n with $c_1, \dots, c_n$ as right zeros. ■

Of course, it is also possible to construct polynomials with conjugate zeros, but then we no longer have uniqueness. Thus let $c_1, \dots, c_n$ be any elements of R , possibly conjugate, and not necessarily distinct, and suppose that we already have a monic polynomial f of degree $n-1$ with $c_2, \dots, c_n$ as right zeros. We can write $f = (x - c_1)r$, where r is chosen so that $g(c_1) \neq 0$, and now $f \cdot (x - g(c_1)^{-1}c_1g(c_1))$ is a monic polynomial of degree n with $c_1, \dots, c_n$ as right zeros. We cannot expect to have uniqueness now, for if f has two right zeros that are conjugate, then by Th. 6.2 it has infinitely many right zeros in this conjugacy class, so we can find a monic polynomial of degree less than n with n zeros, and hence infinitely many monic polynomials of degree n with these n zeros. We remark that if m of the c_i are equal, to c_1 , say, then f constructed above will have $(x - c_1)^m$ as a left factor.

We shall now consider a particular class of finite-dimensional extensions of K , namely the homomorphic images of the skew polynomial ring $R = K[t; \alpha, \delta]$. Thus we have a field E which is generated (as a ring) by K and an element θ satisfying

$$a\theta = \theta a^\alpha + a^\delta \quad \text{for all } a \in K. \quad (7)$$

Such an extension is said to be *pseudo-linear*. Since E is a field, it must be a proper homomorphic image of R , and this is obtained by factoring out a non-zero ideal of R . Since R is right principal, this ideal has the form fR ,

where f is a right invariant polynomial, which may be taken to be monic. If f has degree n in t and θ denotes the residue class of t , then E is a right K -space with basis $1, \theta, \dots, \theta^{n-1}$. It follows that for each non-zero ideal fR , R/fR is either a field extension of K or it has zero-divisors, for each left multiplication is a linear transformation of a finite-dimensional K -space and so is either invertible or has a non-zero kernel. A necessary condition for the absence of zero-divisors is that f be irreducible over K ; when this is so, f is an I -atom of length 1, and hence R/fR is then a field by Th. 6.4.3. So to obtain all pseudo-linear extensions of K associated with α, δ we have to find all irreducible right invariant monic polynomials in $K[t; \alpha, \delta]$.

We shall restrict ourselves to the case $\delta = 0$. Since f is right invariant, we have by Prop. 3.2, $cf = fc^{\alpha^n}$ for all $c \in K$. A comparison of constant terms gives $ca_n = a_n c^{\alpha^n}$, and $a_n \neq 0$ by the irreducibility of f . This shows α^n to be inner, so that we are in the case of Prop. 3.7. Let r be the least positive integer such that α^r is inner on K , induced by e^{-1} , say, and write $x = t^r e$, then x centralizes K , and if $t \mapsto \theta$, $x \mapsto \xi$ in the homomorphism $R \rightarrow R/fR$, then we have

$$K \subset K(\xi) \subset K(\theta). \quad (8)$$

Let us call a pseudo-linear extension *central* if $\alpha = 1, \delta = 0$ and the generator satisfies a monic equation with coefficients in the centre C of K . Thus a central extension is a simple extension of the form $K \otimes_C F$, where F/C is a commutative simple field extension. For example, the above extension $K(\xi)/K$ is central. Secondly, we define a *binomial* or *pure* extension as a pseudo-linear extension (with zero derivation) whose generator satisfies a binomial equation

$$x^n - u = 0,$$

and whose associated endomorphism α is such that α^n is the inner automorphism induced by u . Then it is clear that the extension $K(\theta)/K(\xi)$ in (8) is of this form, since θ satisfies the equation $\theta^r = \xi e^{-1}$. We thus have

THEOREM 6.6. *Let K be a skew field with endomorphism α , and assume that the least power α^r (if any) which is inner on K is also the least power leaving the centre C of K elementwise fixed. Then every pseudo-linear extension of K with endomorphism α (and zero derivation) is obtained by taking a central extension, followed by a binomial extension. In particular, every pseudo-linear extension (with zero derivation) of prime degree is either central or binomial. ■*

We can therefore confine our attention to central and binomial extensions in what follows. Consider first central extensions. Given a field K with centre C , any monic polynomial with coefficients in C defines a ring extension of K .

As before, we see that R/fR is a field if and only if f is irreducible over K , so we have

PROPOSITION 6.7. *Let K be a field with centre C and let $F = C[\xi]$ be a finite extension with minimal polynomial f for ξ over C . Then the central extension $K \otimes_C F$ is simple if and only if f is irreducible over C , and it is a field if and only if f is irreducible over K . Moreover, all central extensions of K are obtained in this way. ■*

Next consider binomial extensions. Let again $R = K[t; \alpha]$ and assume that α^n is the inner automorphism induced by $u = u^\alpha$, then the element $f = t^n - u$ of R is invariant. To determine when R/fR is a field we need to know when $t^n - u$ is an atom in R . Now every binomial extension can be obtained as a succession of binomial extensions of prime degrees, for if $K(\theta)/K$ is binomial of degree $n = pm$, then $K(\theta)/K(\theta^p)$ and $K(\theta^p)/K$ are binomial extensions of degrees p, m respectively. Thus we may take $n = p$, a prime.

Clearly $t^p - u$ is an I -atom, because $K(\theta)/K(\theta^p)$ is a field extension. If $t^p - u$ is not itself an atom, it must be a product of similar atoms, by Th. 6.4.3. Now any two similar atoms have the same degree, so $t^p - u$ is a product of linear factors. Further, two linear factors $t - a, t - b$ are similar if and only if they are α -conjugate: if $(t - a)v = u(t - b)$, we may take u, v to be of degree 0, and a comparison of terms shows that

$$b = c^{-1}ac^\alpha \quad \text{for some } c \in K^\times. \quad (9)$$

Let $t - a$ be a left factor of $t^p - u$; then

$$t^p - u = (t - a)(t^{p-1} + t^{p-2}b_1 + \dots + b_{p-1}). \quad (10)$$

Equating coefficients, we find

$$b_1 = a^{\alpha^{p-1}}, b_2 = a^{\alpha^{p-2}}a^{\alpha^{p-1}}, \dots, b_{p-1} = a^\alpha a^{\alpha^2} \dots a^{\alpha^{p-1}}.$$

Therefore

$$u = aa^\alpha \dots a^{\alpha^{p-1}}. \quad (11)$$

Conversely, given any element u of K of the form (11), for some $a \in K$, we have a factorization (10). Thus $t^p - u$ is an atom if and only if u does not satisfy (11), for any $a \in K$. So we have proved

PROPOSITION 6.8. *Let K be a skew field and α an automorphism of K whose p th power (for some prime p) is inner, induced by $u \in K$. If u satisfies no equation (11), for any $a \in K$, then the equations*

$$t^p - u = 0, \quad at = ta^\alpha \quad (a \in K) \quad (12)$$

define a binomial extension of K and, conversely, all such extensions of K are obtained by this process, repeated if necessary. ■

In the special case where K is commutative this is, of course, well known, and it leads to the construction of cyclic crossed products of prime degree (cf. e.g. Cohn [77], Ch. 10).

In the commutative case we know that a binomial expression $t^p - u$ of prime degree either is irreducible or splits into linear factors, provided only that the ground field contains a primitive p th root of 1. The same result still holds in the general case:

PROPOSITION 6.9. *Let p be a prime number and K a field containing a primitive p th root of 1, ω , say, in its centre. Further, let α be an automorphism of K whose p th power is inner on K , induced by u , say. Then the polynomial*

$$t^p - u \quad (13)$$

is either an atom in $K[t; \alpha]$ or it splits into linear factors, a decomposition being

$$t^p - u = (t - a)(t - \omega a^\alpha) \dots (t - \omega^{p-1} a^{\alpha^{p-1}}). \quad (14)$$

where a is any left zero of (13).

Proof. We know that the polynomial (13) is an atom in $R = K[t; \alpha]$ if and only if there is no representation of the form (11) for u . More precisely, the proof of Prop. 6.8 shows that an expression (11) is possible whenever $t - a$ is a left factor of $t^p - u$, and conversely. Let $t - a$ be a left factor of (13), then since $at = ta^\alpha$, we have $t^{-1}a = a^\alpha t^{-1}$ in $K(t; \alpha)$. Hence

$$\begin{aligned} (at^{-1})^p &= at^{-1}at^{-1} \dots at^{-1} \\ &= aa^\alpha \dots a^{\alpha^{p-1}} t^{-p} \\ &= ut^{-p}. \end{aligned}$$

By hypothesis ut^{-p} is central, hence we have a decomposition $1 - ut^{-p} = 1 - (at^{-1})^p = (1 - at^{-1})(1 - \omega at^{-1}) \dots (1 - \omega^{p-1} at^{-1})$; now (14) follows by multiplying up by t^p on the right and taking the commutation formula into account. ■

Exercises 8.6

1. Let L/K be an extension generated as a K -ring by an element θ such that $K\theta \subseteq \theta K + K$. Show that the extension is pseudo-linear; deduce that every quadratic extension is pseudo-linear.

2. Let K be a field and $R = K[t]$. If f is an I -atom in R , show that f has a central zero in an extension of K if and only if f is an atom.

3. Let K be a field which is a k -algebra and let $E = k(\alpha)$ be a (commutative) simple extension of k generated by α , with minimal polynomial f over k . Show that $K \otimes E$ is Artinian and (i) is simple if and only if f is irreducible over the centre of K and (ii) it is a field if and only if f is irreducible over K .

4. Let K be a field and α an automorphism of order n . If A is a matrix satisfying $AA^\alpha \dots A^{\alpha^{n-1}} = I$, show that there exists a matrix B such that $A = B^{-1}B^\alpha$ (this is a form of Hilbert's theorem 90).

5. (K. C. Smith) Let $f \in K[x; \alpha]$, say $f = \sum x^i a_i$; show that $x - c$ is a left factor of f if and only if $\sum c c^\alpha \dots c^{\alpha^{i-1}} a_i = 0$.

6. Let K be a field with an automorphism α whose p th power is inner, induced by u , but α itself is not inner. Show that the equation (11) has no solution in K .

7. Let E/F be a commutative field extension. If f is a polynomial over E which is irreducible over F but has a linear factor over E , show that f splits into linear factors over the universal field of fractions U of $E_F\langle x \rangle$. Hence show that there are irreducible polynomials over E which split over U . (Hint: Take a polynomial with 2-fold transitive Galois group and adjoin one zero to the ground field.)

8*. Let K be a totally ordered field and α an isotone automorphism whose square is inner, induced by $c \in K$. If $c > 0$ and $c = xx^\alpha$ has no solution in K , show that the binomial extension of K by γ subject to $\alpha\gamma = \gamma\alpha^\alpha$, $\gamma^2 = c$, can be ordered so as to extend the ordering on K , by defining $a + \gamma b \geq 0$ if (i) $a \geq 0$, $b \geq 0$, (ii) $a > 0 > b$ and $c < ab^{-1}(ab^{-1})^\alpha$ or (iii) $a < 0 < b$ and $c > ab^{-1}(ab^{-1})^\alpha$ (cf. Cohn [b]).

9. Let E/K be a pseudo-linear extension and denote the left and right degrees by $[E:K]_L, [E:K]_R$. Show that if α is the endomorphism and $[E:K]_R = n$, then

$$[E:K]_L = 1 + [K:K^\alpha]_L + [K:K^\alpha]_L^2 + \dots + [K:K^\alpha]_L^{n-1}.$$

Deduce that $[E:K]_L = n$ if and only if α is an automorphism.

10. (Artin's problem) Let k be a commutative field, define E as the universal field of fractions of $k\langle x, y \rangle$ and K as the subfield of E generated over k by x, y^2 and $xy - yx^2$. Show that E, K admit the endomorphism $\alpha: x \mapsto x^2, y \mapsto -y$, and that K admits the inner α -derivation induced by y ($c \mapsto cy - yc^\alpha$). Show further that E/K is a pseudo-linear extension of right degree 2 and left degree greater than 2. (Hint: Use Ex. 9 and the fact that $x \notin K^\alpha$, cf. Cohn [84].)

8.7 Laurent series and the Malcev-Neumann construction

Let K be a field with an automorphism α and an α -derivation δ ; consider the skew function field $K(x; \alpha, \delta)$, the field of fractions of $K[x; \alpha, \delta]$. To study this field it is convenient to write its elements as Laurent series. Suppose first

that $\delta = 0$; we have the equation

$$ax = xa^{\alpha} \quad (a \in K),$$

and if $\beta = \alpha^{-1}$, we have the relation

$$ax^{-1} = x^{-1}a^{\beta}.$$

Using these two relations to bring all coefficients to the right, we can easily show that the set of all Laurent series

$$f = \sum_{n=-k}^{\infty} x^n a_n \quad (1)$$

forms a field; if in (1), $a_{-k} \neq 0$, then we can write $f = x^{-k}a_{-k}(1-g)$, where $o(g) > 0$, and so f has the inverse $f^{-1} = (\sum g^i)a_{-k}^{-1}x^k$. The resulting field is denoted by $K((x; \alpha))$ and may be obtained from the power series ring $K[[x; \alpha]]$ by inverting x . Since $K[[x; \alpha]]$ is embedded in the power series ring, it is also embedded in $K((x; \alpha))$, therefore, by the uniqueness of the field of fractions, so is $K(x; \alpha)$.

When $\delta \neq 0$, this method cannot be used, essentially because the multiplication by a is not continuous in the x -adic topology, as the equation

$$ax = xa^{\alpha} + a^{\delta} \quad (2)$$

shows. To overcome this difficulty we introduce $y = x^{-1}$ and rewrite the commutation formula (2) in terms of y . We thus obtain

$$ay^{-1} = y^{-1}a^{\alpha} + a^{\delta}, \quad (3)$$

$$ay = ya^{\beta} - ya^{\beta\delta}y. \quad (4)$$

In operator form (4) may be written as $\beta L_y = (1 + \beta\delta L_y)R_y$. If we solve this equation for R_y , we obtain

$$ay = ya^{\beta} - y^2a^{\beta\delta\beta} + y^3a^{(\beta\delta)^2\beta} - \dots \quad (5)$$

The result may be summed up as

PROPOSITION 7.1. *Let K be a field with automorphism α and α -derivation δ . Then the skew function field $K(x; \alpha, \delta)$ may be embedded in the field of Laurent series in $y = x^{-1}$ with commutation relations (3), (4), where $\beta = \alpha^{-1}$. ■*

So far we have assumed α to be an automorphism of K ; when α is merely an endomorphism, i.e. not surjective, it must be injective, as field homomorphism, so we have in any case an isomorphism between K and its image K^{α} . Let us write K_n for the image of K under α^n ($n \geq 0$); since α^n is injective, it provides an isomorphism between $K = K_0$ and K_n . Now for each

$m = 1, 2, \dots$ take an isomorphic copy K_{-m} of K and embed K_{-m} in K_{-m-1} by identifying it with the image of K under α . In this way we obtain a filtration

$$\dots K_n \subset K_{n-1} \subset \dots \subset K_1 \subset K_0 \subset K_{-1} \subset \dots,$$

whose union is again a field, which we shall write as $K^{[\alpha]}$. Since $\alpha: K_n \rightarrow K_{n+1}$ is an isomorphism, for all $n \in \mathbb{Z}$, α is an automorphism of $K^{[\alpha]}$.

We remark that $K^{[\alpha]}$ may also be obtained more directly as follows. In the skew function field $K(x; \alpha)$ we have the inner automorphism induced by x , which agrees with α on K , hence $K_{-m} = x^m K x^{-m}$ and $K^{[\alpha]} = \bigcup x^m K x^{-m}$. Thus in order to form Laurent series when α is not surjective (and $\delta = 0$), we may take $K^{[\alpha]}((x; \alpha))$.

When α is not surjective and $\delta \neq 0$, we can still form $K^{[\alpha]}$ but, of course, δ will not be defined on all of $K^{[\alpha]}$ and there is no natural way of doing so unless we have a permutation relation between α and δ , such as $\alpha\delta = \delta\alpha$.

In the commutation relation for a formal power series ring over K we may from the beginning allow all higher powers; the most general relation is then of the form

$$ax = xa^{\delta_0} + x^2 a^{\delta_1} + x^3 a^{\delta_2} + \dots, \quad (6)$$

where $\delta_0, \delta_1, \delta_2, \dots$ is a sequence of maps of K into itself. As in the case of skew polynomial rings (0.10) we can show that the δ 's are additive and satisfy $1^{\delta_0} = 1$, $1^{\delta_i} = 0$ for $i > 0$. Moreover, δ_0 must be an endomorphism of K , while the δ 's satisfy

$$(ab)^{\delta_i} = \sum a^{\Delta_i^j} b^{\delta_j}, \quad (7)$$

where Δ_i^n is the coefficient of t^{n+1} in $(\sum t^{k+1} \delta_k)^{i+1}$. Such a sequence $(\delta_0, \delta_1, \dots)$ with $\delta_0 = \alpha$ is sometimes called a *higher α -derivation*. We shall not need a special notation for the ring defined by such a higher derivation.

As an example of a ring with commutation formula (6) let us again take a skew polynomial ring $K[x; \alpha, \delta]$, where K is a field. In the skew function field $K(x; \alpha, \delta)$ consider the subring generated by K and x^{-1} ; here it will be convenient to write the coefficients on the left. Writing $y = x^{-1}$, we have as before $ya = a^\alpha y + y a^\delta y$, or in operator form, $L_y = \alpha R_y + \delta L_y R_y$, whence

$$ya = a^\alpha y + a^{\delta\alpha} y^2 + a^{\delta^2\alpha} y^3 + \dots, \quad (8)$$

and this is indeed of the form (6), with $\delta_n = \delta^n \alpha$, except for a change of sides. In particular, if δ is nilpotent, say $\delta^{r+1} = 0$, then (8) reduces to the polynomial formula

$$ya = a^\alpha y + a^{\delta\alpha} y^2 + \dots + a^{\delta^r \alpha} y^{r+1}. \quad (9)$$

Of course, not every higher α -derivation is of the special form $\delta_n = \delta^n \alpha$, but

it is a remarkable fact that if in (6) $\delta_i = 0$ for $i > r$, then (with another mild restriction) we do indeed have $\delta_n = \delta^n \alpha$.

THEOREM 7.2. *Let A be the ring of polynomials in an indeterminate y with coefficients in a field K , with the normal form*

$$f = a_0 + a_1 y + \dots + a_n y^n \quad (a_i \in K), \quad (10)$$

such that $o(f) = \min\{i \mid a_i \neq 0\}$ is an order function, and

$$ya = a^\alpha y + a^{\delta_1} y^2 + \dots + a^{\delta_r} y^{r+1}. \quad (11)$$

Assume further that (i) r is independent of a , (ii) α is an automorphism of K and (iii) $\alpha, \delta_1, \dots, \delta_r$ are right linearly independent over K , in the sense that

$$a^\alpha b_1 + a^{\delta_1} b_2 + \dots + a^{\delta_r} b_{r+1} = 0 \text{ implies } b_1 = b_2 = \dots = b_{r+1} = 0.$$

Then A is obtained from a skew polynomial ring $R = K[x; \alpha, \delta]$, where δ is a nilpotent α -derivation: $\delta^{r+1} = 0$, by adjoining $y = x^{-1}$ and taking polynomials in y . Conversely, every skew polynomial ring with a nilpotent α -derivation δ leads to a ring satisfying (11), with $\delta_i = \delta^i \alpha$.

Proof. The converse has already been established. To prove the direct part, we have from (11), by induction on n ,

$$y^n a = a_{nn} y^n + a_{nn+1} y^{n+1} + \dots + a_{nk} y^k, \quad (12)$$

where k depends on n but not on a . We shall write (12) as

$$y^n a = \sum a_{ni} y^i, \quad (13)$$

where the summation is over all i and $a_{ni} = 0$ for $i < n$ or $i > k$. Clearly $a \mapsto (a_{ni})$ is a matrix representation of K over itself. From (11) we find, by induction on n ,

$$a_{nn} = a^{\alpha^n}. \quad (14)$$

For $n = 1$ we have from (13),

$$\begin{aligned} y(ab) &= \sum (ab)_{1i} y^i, \\ (ya)b &= (\sum a_{1i} y^i)b = \sum a_{1i} b_{ij} y^j. \end{aligned}$$

Hence for $j > r+1$,

$$\sum a_{1i} b_{ij} = 0 \quad (j = r+2, \dots).$$

Now $a_{11} = a^\alpha$, $a_{1i} = a^{\delta_{i-1}}$, and all these elements are right linearly independent over K , by hypothesis, so

$$b_{ij} = 0 \quad \text{for } i = 1, \dots, r+1; j = r+2, \dots$$

Thus (12) takes the form

$$y^i a = a_{ii} y^i + a_{i+1} y^{i+1} + \dots + a_{i+r+1} y^{r+1}. \quad (15)$$

In particular, taking $i = r+1$ and remembering (14), we find

$$y^{r+1} a = a^{\alpha^{r+1}} y^{r+1}. \quad (16)$$

Similarly, for $i = r$, (15) becomes

$$y^r a = a^\alpha y^r + a_{rr+1} y^{r+1}.$$

Let us put $\alpha^{-1} = \beta$ and write a^γ for a_{rr+1} ; then, using (16) we may write this relation formally as

$$y^{-1} a = y^r a^{\beta^{r+1}} y^{-r-1} = a^\beta y^{-1} + a^{\beta^{r+1}} \gamma.$$

If we define δ by setting $a^\delta = -a^{\beta^r} \gamma$, this relation takes on the form

$$y^{-1} a = a^\beta y^{-1} - a^{\beta \delta} \quad (17)$$

or

$$ay = ya^\beta - ya^{\beta \delta} y.$$

Now a comparison with (4) shows that we have

$$ay^{-1} = y^{-1} a^\alpha + a^\delta.$$

Thus A is obtained from the skew polynomial ring $K[x; \alpha, \delta]$ by taking the subring generated by K and $y = x^{-1}$. To find the relation between δ and the δ_i in (11) we apply y^{-1} to (11) and use (17):

$$\begin{aligned} a &= y^{-1}(a^\alpha y + a^{\delta_1} y^2 + \dots + a^{\delta_r} y^{r+1}) \\ &= a + a^{\delta_1 \beta} y + \dots + a^{\delta_r \beta} y^r - a^\delta y - a^{\delta_1 \beta \delta} y^2 - \dots - a^{\delta_r \beta \delta} y^{r+1}. \end{aligned}$$

Equating coefficients, we find $\delta_1 \beta = \delta$, $\delta_i \beta = \delta_{i-1} \beta \delta$, $\delta_r \beta \delta = 0$, hence by induction, $\delta_i \beta = \delta^i$ or

$$\delta_i = \delta^i \alpha, \quad i = 1, \dots, r, \quad \delta^{r+1} = 0. \quad \blacksquare$$

The power series representation is often useful for rational functions, e.g. for determining the centre of the rational function field as we have seen already in 6.2. For simplicity we shall assume that $\delta = 0$:

PROPOSITION 7.3. *Let K be a field and α an endomorphism of K , no positive power of which is inner. Let C be the centre of K and C_0 the subfield fixed by α , then the centre of $K(x; \alpha)$ is C_0 .*

Proof. If α^{-r} is inner, so is α^r , hence no power α^i ($i \neq 0$) is inner. Consider a power series $f = \sum x^i a_i$. If this centralizes $K(x; \alpha)$, then $bf = fb$ for all $b \in K$,

hence $\sum x^i (b^{\alpha^i} a_i - a_i b) = 0$, so $a_i b = b^{\alpha^i} a_i$, but α^i is not inner for $i \neq 0$, hence $a_i = 0$ for $i \neq 0$ and $f = a_0 \in C$. Since $tf = ft$, we have $a_0^\alpha = a_0$, i.e. $f \in C_0$. ■

Sometimes it is useful to have a criterion for the rationality of a power series. Such a criterion is familiar for complex series, and this carries over to the general case:

Rationality criterion. A series $\sum x^i a_i \in K((x; \alpha))$ is a rational function of x if and only if there exist integers r, n_0 and elements $c_1, \dots, c_r \in K$ such that

$$a_n = a_{n-1}^\alpha c_1 + a_{n-2}^{\alpha^2} c_2 + \dots + a_{n-r}^{\alpha^r} c_r \quad \text{for all } n > n_0. \quad (18)$$

For this is just the condition that

$$(\sum x^i a_i) (1 - \sum_1^r x^j c_j)$$

should be a polynomial, except for a factor x^{-k} .

As an illustration we have

PROPOSITION 7.4. Let $K \subseteq L$ be fields and α an automorphism of L mapping K into itself. Then

$$K((x; \alpha)) \cap L(x; \alpha) = K(x; \alpha). \quad (19)$$

Proof. Any element of $L(x; \alpha)$ may be written uniquely as a Laurent series $\sum x^i a_i$ with coefficients a_i in L , and if it is a Laurent series over K , it follows that $a_i \in K$. If f is rational, the coefficients satisfy the above criterion, thus the equations (18) have a solution for the c_j in L . They are linear equations with coefficients in K and hence have a solution in K ; this means that $f \in K(x; \alpha)$, as claimed. The converse is clear. ■

The power series ring and the Laurent series ring are special cases of the following construction, which allows the group algebra of any ordered group to be embedded in a field.

Let M be a totally ordered monoid, i.e. a monoid with a total ordering ' $<$ ' such that $a_i < b_i$ ($i = 1, 2$) implies $a_1 a_2 < b_1 b_2$. By a *convex* submonoid we shall understand a submonoid S such that $a, b \in S$ and $a \leq x \leq b$ (or $a \geq x \geq b$) implies $x \in S$. Let K be any ring and consider the direct power K^M , regarded as a K -module. With each $f \in K^M$ we associate its *support*, defined as

$$D(f) = \{a \in M \mid f(a) \neq 0\}.$$

The elements with finite support may be written as finite sums $\sum f(a)a$ and just constitute the monoid ring KM of M over K , with the multiplication rule

$$fg = h, \quad \text{where } h(c) = \sum_{ab=c} f(a)g(b). \quad (20)$$

When M is a group, the latter sum may also be written $\sum f(x)g(x^{-1}c)$. Now let $R = K((M))$ be the set of all elements of K^M with well-ordered support; here the definition (20) for the product still makes sense, for if the terms $f(a)g(b)$ are ordered, so that $a_1 < a_2 < \dots$ for the arguments of f , then for $c = a_i b_i$ we have $b_1 > b_2 > \dots$ and this must break off by the well ordering. Thus each $h(c)$ is defined; let us show further that h again has well-ordered support and that R is in fact a ring. Let $c_n \in D(h)$, say $c_n = a_n b_n$, and suppose that

$$c_1 > c_2 > \dots$$

By the well-ordering of $D(f)$ there is a subsequence (a_{n_i}) which is increasing, hence (b_{n_i}) is decreasing and so must terminate. Thus $D(h)$ is well-ordered and this shows that $h = fg \in R$: clearly $f+g \in R$ and it is easily checked that R is a K -ring. We shall call $R = K((M))$ the *ordered series ring* of M over K . We remark that R has an order function defined on it, with values in M , by

$$o(f) = \min\{a \in M \mid f(a) \neq 0\}.$$

If $o(f) = a$, then the term $f(a)a$ is called the *leading term* of f . We claim that when the leading term of f is invertible, then f is invertible. For let $f = f_0 a_0 + \dots$, then $g = a_0^{-1} f_0^{-1} f = 1 - h$, where $h = \sum h(x)x$ has support consisting entirely of elements > 1 . Formally we can write

$$p = 1 + h + h^2 + \dots \quad (21)$$

If we can show that $p \in R$, then it is clear that p is indeed an inverse of g , and this will prove f to be invertible.

Thus we have to show that $D(p)$ is well-ordered, and that any element of M occurs in only finitely many of the $D(h^n)$. Consider the first point; if this were not so, we would have a descending chain

$$z_1 = u_{11} \dots u_{1n_1} > \dots > z_i = u_{i1} \dots u_{in_i} > \dots, \quad (22)$$

where $u_{ij} \in D(h)$. Let H_i be the convex submonoid generated by z_i ; since $u_{ij} > 1$, H_i is the convex submonoid generated by the greatest of the elements u_{ij} ($j = 1, \dots, n_i$), say t_i . Now $D(h)$ is well-ordered, so there is among the submonoids H_i a least one, $H = H_k$, say. By (22), $H_1 \supseteq H_2 \supseteq \dots \supseteq H_k = H_{k+1} = \dots$; omitting the first k terms from (22), we may take $k = 1$. Let t be the least among the elements of $D(h)$ generating H as a convex submonoid, then $t \leq t_1 \leq z_1$ and since t and z_1 generate the same convex submonoid, there exists $r \geq 1$ such that $z_1 < t^r$. We choose r minimal and observe that by (22),

$z_i < t^r$ for $i = 1, 2, \dots$. Now z_i has the form

$$z_i = a_i t_i b_i, \quad (23)$$

where a_i, b_i are (possibly empty) products of the u_{ij} . Since the t_i include no infinite descending subsequence, either (a_i) or (b_i) has an infinite descending subsequence, say

$$b_{i_1} > b_{i_2} > \dots$$

Here $b_i < z_i$ and the z_i are such that H is the least possible, therefore $t \leq b_i$; moreover, by (23) and the definition of r , $b_i < t^{r-1}$. So from the sequence z_i satisfying (23) we have obtained the sequence b_i with the same property but for smaller r . By induction on r we reach a contradiction, and this shows $D(p)$ to be well-ordered.

It remains to show that any given element of M occurs in at most finitely many of the $D(h^n)$. If not, let $a \in M$ occur in $D(h^n)$ for infinitely many n , thus

$$a = b_{i_1} \dots b_{i_{n_1}}, \quad \text{where } n_1 < n_2 < \dots \text{ and } b_{ij} \in D(h).$$

Since $D(h)$ is well-ordered, we can choose the least a with this property. The sequence $b_{11}, b_{21}, \dots$ may be assumed to be non-decreasing, without loss of generality, and so $a_i = b_{i2} \dots b_{i_{n_i}}$ is non-increasing and $a_i < a$, therefore the sequence becomes stationary: $a_k = a_{k+1} = \dots$. But this contradicts the minimality of a , hence p is well-defined. So we have proved

THEOREM 7.5. *Let K be a ring and M a totally ordered monoid. Then the set $K((M))$ of power series with well-ordered support is a ring, the ordered series ring of M over K , and an element of this ring is invertible whenever its leading term is invertible. ■*

When M is a group and K is a field, every non-zero element has an invertible leading term and we obtain the *Malcev–Neumann construction*:

COROLLARY 7.6. *Let K be a field and G a totally ordered group. Then the ordered series ring $K((G))$ is a field. ■*

We shall want to apply this result to free groups, so we shall need to prove that every free group G can be totally ordered. This follows, for example, by writing the elements of G as infinite power products of basic commutators and taking the lexicographic ordering of the exponents (cf. M. Hall [59], Ch. 11). Here is a more direct approach, suggested by G. M. Bergman.

Let G be the free group on $X = \{x_1, \dots, x_d\}$; the group algebra kG may be regarded as the universal localization at X of the free algebra $k\langle X \rangle$. Let us

write $x_i = 1 - y_i$, $Y = \{y_1, \dots, y_d\}$; then $k\langle X \rangle = k\langle Y \rangle$ and the latter has a totally inert embedding in $k\langle\langle Y \rangle\rangle$, by the inertia theorem 2.9.15. Now the natural mapping $k\langle X \rangle \rightarrow k\langle\langle Y \rangle\rangle$ is X -inverting, because $(1 - y_i)^{-1} = \sum y_i^n$, so we have an embedding of kG into $k\langle\langle Y \rangle\rangle$. This follows because any non-zero element of kG has an admissible matrix with full numerator, and by inertia this numerator remains full over $k\langle\langle Y \rangle\rangle$. We note that this proves again that kG is embeddable in a field.

If k is any ordered field, then $k\langle\langle Y \rangle\rangle$ may be ordered by ordering the monomials in Y in any way consistent with multiplication, e.g. by length, with monomials of the same length ordered lexicographically, and then taking $f > 0$ if its leading term has coefficient > 0 . In this way $k\langle\langle Y \rangle\rangle$ becomes an ordered ring, and G as a subgroup, becomes an ordered group.

We conclude this section with a remarkable result due to G. M. Bergman, on normal forms under conjugation in ordered series rings.

THEOREM 7.7. *Let K be a commutative ring and M an ordered monoid, and let f be any element of $K((M))$ with leading term $f_u \cdot u$. If $f_u \cdot u$ is invertible, then there exists an element q with leading term 1 such that the element $q^{-1}fq$ has support entirely in the centralizer of u in M .*

Further, q may be chosen so that no element of its support except the leading term 1 commutes with u . Under this hypothesis q is unique.

Proof. If $u = 1$, there is nothing to prove, so we may assume that $u \neq 1$, and replacing f by f^{-1} if necessary we may also suppose that $u > 1$. Further, we may assume without loss of generality that $f_u = 1$. We shall denote the centralizer of u in M by C_u .

The idea of the proof will be to show that any term αt of f such that $t \notin C_u$ can be got rid of by conjugating with an element $(1 + \alpha t u^{-1})$ or $(1 - \alpha u^{-1} t)$, at the expense of adding higher terms. The process is then repeated on the new leading term. We shall need to set up some machinery to show that the construction of q can be made to 'converge'.

We shall use the customary notation $[u, v]$, $[u, v)$, etc., for closed, half-open, etc., intervals in M , and let ∞ be such that $s < \infty$ for all $s \in M$. This will allow us to use the phrase: 'the leading term of f is αt ', even when $f = 0$, in which case t is taken to be ∞ and α undefined. If $u, s \in M$, where u is invertible, we put $s/u = \max\{su^{-1}, u^{-1}s\}$; it is easy to see that $x < y \Rightarrow x/u < y/u$.

Let X be the set of triples (t, g, p) , where $t \in (u, \infty]$, g is an element of $K((M))$ with leading term u and support in $C_u \cap [u, t)$ and p is an element with leading term 1 and support in $[1, t/u)$, such that $pgp^{-1} - f$ has a leading term of the form αt . By our conventions this means that $pgp^{-1} = f$ if and only if $t = \infty$. We partially order X by writing $(t, g, p) \leq (t', g', p')$ if $t \leq t'$, $D(g' - g) \subseteq [t, t')$

and $D(p' - p) \subseteq [t/u, t'/u]$; these conventions just mean that the series g' and p' 'extend' g and p , respectively.

If the leading term of $f - u$ is αt , then $(t, u, 1) \in X$; this shows X to be non-empty. Our aim will be to show that X is inductive, so that we can apply Zorn's lemma. Suppose then that we have a chain $\{(t_\lambda, g_\lambda, p_\lambda) \mid \lambda \in I\}$ in X . We can 'piece together' the g_λ to get a common extension g , with support in $\bigcup_\lambda [u, t_\lambda]$, and similarly from the p_λ get a common extension p with support in $\bigcup_\lambda [1, (t_\lambda/u)]$. It is clear that g and p again have well-ordered supports; we also see that $pgp^{-1} - f$ will have support in $[u, \infty) \setminus \bigcup_\lambda [u, t_\lambda]$. If we write the leading term of $pgp^{-1} - f$ as αt , we see that (t, g, p) majorizes the given chain; hence X is inductive and by Zorn's lemma, X has a maximal element.

To complete the proof we show that if $(t, g, p) \in X$ and $t < \infty$, then we can construct $(t', g', p') > (t, g, p)$. It will follow that a maximal element of X must have the form (∞, h, q) , so that $qhq^{-1} - f = 0$, i.e. $h = q^{-1}fq$, as claimed.

Let the leading term of $pgp^{-1} - f$ be αt , say. We shall describe an element g' extending g and having support in $[u, t]$ (recall that $D(g) \subseteq [u, t]$) and p' extending p with support in $[1, t/u]$, and show that $D(p'g'p'^{-1} - f) \subseteq (t, \infty)$. Hence if $\alpha t'$ is the leading term of $p'g'p'^{-1} - f$, we see that $(t, g, p) < (t', g', p') \in X$. We distinguish three cases.

(i) $tu^{-1} > u^{-1}t$. Then $t/u = tu^{-1}$; write $p' = p - \alpha tu^{-1}$; since p has leading term 1, we see that $p'^{-1} = p^{-1} + \alpha tu^{-1} + \text{higher terms}$. Take $g' = g$; on multiplying out $p'g'p'^{-1}$ we find that the new terms introduced are $(-\alpha tu^{-1}) \cdot u \cdot 1 = -\alpha t$, $1 \cdot u \cdot (\alpha tu^{-1}) = \alpha tu^{-1}$, and higher terms. Since $tu^{-1} > u^{-1}t$, we have $utu^{-1} > t$, so the lowest term introduced is $-\alpha t$; this cancels the leading term αt of $pgp^{-1} - f$, hence $D(p'g'p'^{-1} - f) \subseteq (t, \infty)$ as claimed.

(ii) $tu^{-1} < u^{-1}t$. Now $t/u = u^{-1}t$; we shall put $p' = p + \alpha u^{-1}t$ and again take $g' = g$. The lowest terms introduced are $(\alpha u^{-1}t) \cdot u \cdot 1 = \alpha u^{-1}tu$ and $1 \cdot u \cdot (-\alpha u^{-1}t) = -\alpha t$. Here the latter is the lower and again this cancels the leading term αt of $pgp^{-1} - f$.

(iii) $tu^{-1} = u^{-1}t$. Then t commutes with u ; in this case the terms arising under (i), (ii) would cancel and so be of no help in eliminating the leading term of $pgp^{-1} - f$. So we set $p' = p$ and $g' = g - \alpha t$, which is permissible because $t \in C_u$. Now the lowest term by which $pgp^{-1} - f$ has changed is $1 \cdot (-\alpha t) \cdot 1 = -\alpha t$ so, in this case too, $p'g'p'^{-1} - f$ has support in (t, ∞) .

This then proves the existence of q ; since in the above construction we never added a term from C_u to our p 's, we can clearly take q so that

$$D(q) \cap C_u = \{1\}. \quad (24)$$

Let q be so chosen and suppose that $q' \neq q$ is another element with leading term 1 such that $D(q'^{-1}fq') \subseteq C_u$. Write $q' = q(1 + h)$, $g = q^{-1}fq$, so that $g' =$

$q'^{-1}fq' = (1+h)^{-1}g(1+h)$. If αt is the leading term of h , suppose that $t \notin C_u$; then g' would have a term ut or tu (whichever is the smaller), but by hypothesis, $D(g') \subseteq C_u$, so $t \in C_u$. But this means that $q' = q(1+h)$ will have in its support a term $t \neq 1$ which commutes with u . This is a contradiction, hence q is the unique element with the desired property and satisfying (24). ■

We note that the assertion of Th. 7.7 is strongest when the centralizer C_u is small. In particular, in a free group the centralizer of any element $\neq 1$ is a cyclic group. This yields

COROLLARY 7.8. *Let G be a free group. In the ordered series ring $k((G))$ every element is conjugate to a Laurent series in a single variable.*

If f has leading term αu , $u \neq 1$ and the centralizer of u is the cyclic group on v , then f is conjugate to a Laurent series in v , by Th. 7.7. If f has leading term $\alpha \cdot 1$, we can apply the same argument to $f - \alpha$. ■

Exercises 8.7

1°. Obtain an analogue of Prop. 7.1 when $\delta \neq 0$ and α is not surjective. (Try the cases $\alpha\delta = \pm\delta\alpha$ first.)

2°. To what extent are the conditions (i)–(iii) in Th. 7.2 necessary?

3. Show that in a totally ordered group G any inner automorphism is order-preserving. Deduce that an element f of $k((G))$ commutes with $u \in G$ if and only if the support of f lies in the centralizer of u .

4*. (Makar-Limanov) Show that for any commutator $[a, b] = ab - ba$ in an ordered series ring $k((G))$ the coefficient of 1 is 0. Deduce that $[[a, b], b] = 0$ implies $[a, b] = 0$.

5. (Bergman [78]) Let $k((G))$ be an ordered series ring of a group G . If $\alpha \in k^\times$, $u \in G$, $\text{char } k \nmid n$, show that any series with leading term $\alpha^n u^n$ has an n th root, and that this n th root is unique up to an n th root of 1 in k . Show that if $a^n = b^n$ and a, b have the same leading term, then $a = b$.

6°. Show that in $k\langle\langle X \rangle\rangle$ every element $f \notin k$ is conjugate to a power series f^* in a monomial in X . If f is rational, is the same true of f^* ?

7. Show that for any ordered group G , $k((G))$ has a valuation with values in G .

8. Let G be an ordered group with a maximal archimedean class. Show that $k((G))$ has a real-valued valuation.

9. Let F be the free group on x, y . Show that F can be defined as an ordered group such that there is a maximal archimedean class. Deduce that F has a non-trivial order-homomorphism to $\mathbf{R}$.

8.8 Iterated skew polynomial rings

From any ring R with an automorphism α we can form the skew polynomial ring $R[x; \alpha]$, but this will not be a principal ideal domain, unless R was a field. However, when α is only required to be an endomorphism, we get a principal ideal domain under wider conditions. The exact class was found by Jategaonkar, who used it to give an ingenious construction of 'iterated skew polynomial rings' which form a useful source of counter-examples. We shall here follow Lenstra's presentation in showing more generally that the iterated skew polynomial rings of Jategaonkar type (' J -rings') form the precise class of integral domains with unique remainder algorithm.

Let R be a ring, α an endomorphism and δ an α -derivation of R . Then the skew polynomial ring $R[x; \alpha, \delta]$ is called a J -skew polynomial ring over R or simply a J -ring if α is injective and $R^\alpha \subseteq U(R) \cup \{0\}$.

Given a ring R , a subring K and an ordinal number τ , we shall say that R is an *iterated skew polynomial ring* of type τ over K if R contains an ascending chain of K -rings R_λ ($\lambda \leq \tau$), such that

- (i) $R_0 = K$,
- (ii) $R_{\lambda+1}$ is a skew polynomial ring over R_λ ($\lambda < \tau$),
- (iii) $R_\mu = \bigcup_{\lambda < \mu} R_\lambda$ for any limit ordinal $\mu \leq \tau$,
- (iv) $R_\tau = R$.

From the definition it is clear that every element of R can be written uniquely as

$$\sum x_{\lambda_1} \dots x_{\lambda_r} a_{\lambda_1 \dots \lambda_r}, \quad a_{\lambda_1 \dots \lambda_r} \in K, \lambda_1 \geq \dots \geq \lambda_r. \quad (1)$$

If K is a field and each $R_{\lambda+1}$ is a J -skew polynomial ring over R_λ , R will be called a J -ring of type τ over K .

We know from Prop. 0.10.4 that any skew polynomial ring over a right Ore domain (with an injective endomorphism) is again right Ore. Jategaonkar's basic observation was that a J -skew polynomial ring over a principal right ideal domain is again right principal (cf. Ex. 1). We now show that a J -ring over a Euclidean domain is again Euclidean. In a Euclidean ring we shall denote the function giving the least algorithm by θ ($\theta(0) = -\infty$). We also recall the multiplication of ordinals: $2 \cdot \omega = \omega$, $\omega \cdot 2 = \omega + \omega$.

PROPOSITION 8.1. *Let R be an integral domain with unique remainder algorithm, defined by θ_R and suppose that $S = R[x; \alpha, \delta]$ is a J -ring. Then S again has a*

unique remainder algorithm, with the function

$$\theta(\sum_0^n x^i a_i) = \lambda \cdot n + \theta_R(a_n) \quad (a_n \neq 0, \lambda \text{ ordinal}). \quad (2)$$

Proof. Let $a = \sum_0^n x^i a_i$, $b = \sum_0^m x^i b_i$ ($a_n, b_m \neq 0$), and note that $\theta = \theta_R$ on R . If $\theta(a) \geq \theta(b)$, then $n \geq m$. Either $n > m$, then $a' = a - bx(b_m^\alpha)^{-1}x^{n-m-1}a_n$ has degree less than n and so $\theta(a') < \theta(a)$, or $n = m$, $\theta(a_n) \geq \theta(b_n)$ and so by the algorithm in R , $\theta(a_n - b_n c) < \theta(a_n)$ for suitable $c \in R$, hence $\theta(a - bc) < \theta(a)$.

Thus S is Euclidean; further the conditions of Prop. 2.1.3 are easily checked, so we have a unique remainder algorithm on S . ■

Since a field has a unique remainder algorithm, we obtain by transfinite induction

COROLLARY 8.2. *For any ordinal τ , a J -ring of type τ over a field is an integral domain with unique remainder algorithm. ■*

It is of interest that the converse also holds:

THEOREM 8.3. *Any ring R is a J -ring of type τ over a field (for some ordinal τ) if and only if R is an integral domain with unique remainder algorithm.*

Proof. The direct part was proved in Cor. 8.2. For the converse, let R be an integral domain with unique remainder algorithm $\theta = \theta_R$; by Prop. 2.1.5, θ is a degree function. Take an ordinal τ bounding θ and define Λ as the set of ordinals $\lambda \leq \tau$ in the range of θ such that

$$\lambda > 0 \text{ and } \beta, \gamma < \lambda \Rightarrow \beta + \gamma < \lambda. \quad (3)$$

We claim

$$\lambda \in \Lambda, \beta < \lambda \Rightarrow \beta + \lambda = \lambda. \quad (4)$$

For the unique solution γ of $\beta + \gamma = \lambda$ cannot be $< \lambda$ by (3), but clearly $\gamma \leq \lambda$, hence $\gamma = \lambda$.

We index Λ by an initial segment of the ordinals such that $\lambda_\beta < \lambda_\gamma \Rightarrow \beta < \gamma$, and define

$$R_\beta = \{a \in R \mid \theta(a) < \lambda_\beta\} \quad (\beta < \tau).$$

Since θ is a degree function, R_β is a subring of R . Now $\lambda_0 = 1$, so $R_0 = \{a \in R \mid \theta(a) \leq 0\}$ and this is a field, K say. For any limit ordinal γ we have $\lambda_\gamma = \lim_{\beta < \gamma} \lambda_\beta$ and hence

$$R_\gamma = \bigcup_{\beta < \gamma} R_\beta,$$

while $R_\tau = R$, so it only remains to show that for any $\beta < \tau$, $R_{\beta+1}$ is a J -skew polynomial ring over R_β . Choose $x \in R$ such that $\theta(x) = \lambda_\beta$ and denote by $R_\beta[x]$ the set of all polynomials $f = \sum x^i a_i$ ($a_i \in R_\beta$). We claim that $R_\beta[x]$ is a J -skew polynomial ring. Given $f = \sum_0^n x^i a_i$ with $a_n \neq 0$, we have $\theta(f) = \lambda_\beta \cdot n + \theta(a_n)$, hence $f \neq 0$ and so the expression is unique. Given $0 \neq a \in R_\beta$, we have $ax = xa^\alpha + a^\delta$ for a unique $a^\alpha, a^\delta \in R$ and $\theta(a^\delta) < \theta(x) = \lambda_\beta$, so $a^\delta \in R_\beta$. Next we have $\theta(ax) = \theta(a) + \theta(x) = \theta(x)$ by (4), so $\theta(xa^\alpha) = \theta(x)$, whence $\theta(a^\alpha) = 0$ and it follows that $a^\alpha \in R_0 \subseteq R_\beta$. Therefore α, δ map R_β into itself (if we define $0^\alpha = 0^\delta = 0$) and $R_\beta[x]$ is indeed a J -skew polynomial ring.

To establish the equality $R_{\beta+1} = R_\beta[x]$ we note that $\lambda = \lambda_\beta \cdot \omega$ is the smallest ordinal of Λ which is $> \lambda_\beta$, hence $\lambda_{\beta+1} = \lambda_\beta \cdot \omega$. We take $f \in R_{\beta+1} \setminus R_\beta[x]$ such that $\theta(f)$ has its least value. By the Euclidean algorithm, $f = xa + b$ with $b \in R_\beta$. As before, $\theta(f) = \theta(x) + \theta(a) = \lambda_\beta + \theta(a)$, since $a \neq 0$. But $\theta(f) < \lambda_{\beta+1}$, so we have $\theta(f) = \lambda_\beta \cdot n + \gamma$ with $\gamma < \lambda_\beta$ and $1 \leq n < \omega$. Hence $\theta(a) = \lambda_\beta(n-1) + \gamma < \lambda_\beta n \leq \theta(f)$, therefore $a \in R_\beta[x]$, by the minimality of $\theta(f)$, and so $f = xa + b \in R_\beta[x]$, a contradiction. This shows that $R_{\beta+1} = R_\beta[x]$, as claimed. ■

There remains the problem of constructing J -rings of a given type. The J -rings of type 0 are fields, J -rings of type 1 are the usual skew polynomial rings over fields, J -rings of type 2 may be obtained by an *ad hoc* construction (cf. Ex. 9), but beyond that it is no easier to construct a J -ring of finite type than one of arbitrary type. In particular it is not possible to give a recursive construction, because the set $K^\times$ must contain an isomorphic copy of $R^\times$, so that K depends very much on τ . To construct K we shall need an embedding of the free algebra $k\langle X \rangle$ in a field, in which the elements have a simple normal form. Here it is more convenient to use instead of the method of 7.5 the Malcev–Neumann construction (Cor. 7.6), which embeds $k\langle X \rangle$ in the ordered series field.

Let $X = \{x_\alpha\}$ ($\alpha < \tau$) be a sequence of indeterminates and denote by E the ordered series field constructed from the free group on X by Cor. 7.6. Let K be the subfield generated over k by the elements

$$u_{\beta\alpha_1 \dots \alpha_r} = (x_{\alpha_1} \dots x_{\alpha_r})^{-1} x_\beta x_{\alpha_1} \dots x_{\alpha_r}, \quad \alpha_1 \geq \dots \geq \alpha_r, \alpha_1 > \beta, r \geq 1. \quad (5)$$

LEMMA 8.4. *In the field K as defined above, the centralizer of x_γ for any γ , is k .*

Proof. Let F be the free group on X and G the subgroup of F generated by the right-hand sides of (5). Each of these generators has odd length in the x 's, and in any expression of an element of G as a power product of the u 's, cancellation cannot affect the central factor of any u . It follows that each such

expression must begin with a factor x_α^{-1} and end in a factor x_β , even after all the cancellations have been made, for this is true of the u 's and their inverses. In particular, it follows that G does not contain x_γ^n , for any $n \neq 0$.

Consider any $a \in K$; this is a power series: $a = \sum u_\gamma u$, where u runs over G . Conjugation by x_γ maps G into itself; explicitly we have

$$x_\gamma^{-1} u_{\beta\alpha_1 \dots \alpha_r} x_\gamma = u_{\alpha_r \gamma}^{-1} \dots u_{\alpha_1 \gamma}^{-1} u_{\beta\alpha_1 \dots \alpha_{i-1} \gamma} u_{\alpha_i \gamma} \dots u_{\alpha_r \gamma},$$

where i is such that $\alpha_{i-1} \geq \gamma > \alpha_i$.

Since x_γ commutes only with the powers x_γ^n , it follows that conjugation by x_γ fixes only $1 \in G$ and moves all other elements in infinite orbits; to be precise, the elements in each orbit arise from the positive powers of x_γ , since G admits conjugation by x_γ , but not by x_γ^{-1} . Hence $x_\gamma^{-1} a x_\gamma = a$ is possible only if $a_u = 0$ for $u \neq 1$, i.e. $a = a_1 \in K$. ■

Let ϕ_γ be the endomorphism of G induced by conjugation with x_γ ; clearly this can be extended to an endomorphism of K , again denoted by ϕ_γ . Then

$$a x_\gamma = x_\gamma a^{\phi_\gamma} \quad (a \in K). \quad (6)$$

Denote by R the subring of E generated by K and the x_β ($\beta < \tau$). By (6), every element of R can be written as a finite sum

$$\sum x_{\alpha_1} \dots x_{\alpha_r} a_{\alpha_1 \dots \alpha_r}, \quad \text{where } a_{\alpha_1 \dots \alpha_r} \in K. \quad (7)$$

Fix any term in (7) and let α_i be the last suffix such that $\alpha_i < \alpha_{i+1} \geq \dots \geq \alpha_r$. Then we can use (5) to pull x_{α_i} through to the right; this will only change the coefficient. Repeating this process if necessary we can ensure that $\alpha_1 \geq \dots \geq \alpha_r$ in each term of (7). We assert that under this condition the expression (7) is unique. Thus let

$$\sum x_{\alpha_1} \dots x_{\alpha_r} a_{\alpha_1 \dots \alpha_r} = 0 \quad (\alpha_1 \geq \dots \geq \alpha_r), \quad (8)$$

and assume that not all coefficients are 0. Let α be the highest suffix such that x_α occurs in (8); then (8) may be written as

$$\sum x_\alpha^i c_i = 0, \quad (9)$$

where each c_i is a polynomial in the x_β ($\beta < \alpha$), with right coefficients in K , and not all c_i vanish. On conjugating by x_α we get

$$\sum x_\alpha^i c_i^{\phi_\alpha} = 0,$$

where the coefficients now lie in K ; hence x_α is right algebraic over K . Take a monic equation of least degree satisfied by x_α :

$$\sum x_\alpha^i b_i = 0.$$

This expression is unique, but conjugating with x_α we obtain another equation of the same degree with coefficients $b_i^{\phi_\alpha}$. The leading coefficient, being 1, is unchanged, hence each b_i is unchanged and by the lemma, $b_i \in k$. Thus x_α is right algebraic over k , which is a contradiction. Thus we have established the uniqueness of (7) and this proves

THEOREM 8.5. *Let k be any field and τ an ordinal. Then there exists a J -ring of type τ over k (with zero derivation). ■*

By Th. 8.3 the resulting J -ring R has a Euclidean algorithm, and hence is a principal right ideal domain. However, it is not atomic if $\tau \geq 2$, for then we have

$$x_2 = x_1 x_2 u_{12}^{-1} = x_1^n x_2 u_{12}^{-n},$$

so x_2 has factorizations of arbitrary length.

A closely related class of rings has been studied by Brungs [69]; he considers rings in which the set of all right ideals is well-ordered by inclusion. Actually it is enough to take the set of all principal right ideals. Such a ring contains a unique maximal principal proper right ideal which is clearly also the unique maximal proper right ideal and hence is the Jacobson radical $J = J(R)$ of R . Hence J is two-sided and R/J is a field, because it has no non-trivial right ideals. Thus R is a local ring with maximal ideal J . More precisely, we have the following structure theorem:

THEOREM 8.6. *Let R be a ring in which the set of all principal right ideals is well-ordered by inclusion. Then all right ideals of R are principal and are in fact two-sided; thus all non-zero elements of R are right invariant.*

Proof. Let a be a non-zero right ideal, bR the maximal principal right ideal properly contained in a , and $c \in a \setminus bR$. Then $cR \subseteq a$, but $cR \not\subseteq bR$, hence $bR \subset cR \subseteq a$, and so $cR = a$. Thus all right ideals are principal. We claim that R is right invariant. If not, let aR be a maximal right but not left ideal; then there exists $b \in R$ such that $baR \supset aR$, hence $a = bac$, where c is a non-unit and so lies in J . But baR is two-sided, by the maximality of aR , so $b \cdot ba = ba \cdot b'$, hence

$$ba = b \cdot bac = bab'c.$$

This shows that $ba(1-b'c) = 0$, where $c \in J$, and so $1-b'c$ is a unit. Hence $ba = 0$, a contradiction; thus each right ideal of R is two-sided, as claimed. ■

In order to examine these rings more closely, we note that for any ring R with the property of Th. 8.6, the following holds:

Let $a, b, c \in R$. If $ab \in cR$, $b \notin cR$, then $a^n \in cR$ for some n .

For let $\cap a^n R = dR$, then $adR = dR$, hence $d = adu$ for some $u \in R$. Now either $bR \subseteq dR$, say $b = dv$; then $b = dv = aduv = advu' = abu' \in cR$, which contradicts the hypothesis, or $dR \subset bR$, then $a^n R \subseteq bR$ for some n , and $a^{n+1} \in abR \subseteq cR$.

We now define elements p_α ($0 \leq \alpha$) of R as follows: $p_0 R$ is the maximal right ideal of R , and for any $\alpha > 0$, p_α is defined (up to a right unit factor) by

$$p_\alpha R = \begin{cases} \bigcap_{\beta < \alpha} p_\beta R & \text{if } \alpha \text{ is a limit ordinal,} \\ \bigcap p_{\alpha-1}^n R & \text{otherwise.} \end{cases} \quad (10)$$

Clearly, if $\alpha > \beta$, then $p_\alpha R \subset p_\beta R$; moreover, $p_\beta p_\alpha R \subseteq p_\alpha R$, and by definition of p_α , $p_\alpha = p_\beta c$ for some non-unit c . Since $p_\beta^n \notin p_\alpha R$ for all n , we have $c \in p_\alpha R$, so $p_\alpha \in p_\beta p_\alpha R$, and therefore $p_\alpha R = p_\beta p_\alpha R$, i.e.

$$p_\beta p_\alpha = p_\alpha u_{\alpha\beta} \quad \text{for } \beta < \alpha, \text{ where } u_{\alpha\beta} \in U(R). \quad (11)$$

We assert that each $a \in R^\times$ is expressible uniquely as

$$a = p_{\alpha_1} \dots p_{\alpha_r} u, \quad \text{where } u \in U(R), \alpha_1 \geq \dots \geq \alpha_r. \quad (12)$$

For let α be the least ordinal such that $a \notin p_\alpha R$, then α cannot be a limit ordinal, say $\alpha = \alpha_1 + 1$ and $a \in p_{\alpha_1}^n R$ for some n . Taking n as large as possible, we have

$$a = p_{\alpha_1}^{n_1} a_1, \quad \text{where } a_1 \notin p_{\alpha_1} R.$$

Repeating this process on a_1 we find by induction that a has the form (12) as claimed, and this expression is unique, from the way it was found.

We denote by σ the order type of the sequence of right ideals of R and express σ in the form

$$\sigma = \omega^{\tau_1} n_1 + \omega^{\tau_2} n_2 + \dots + \omega n_{k-1} + n_k + 1, \quad (13)$$

where $\tau_1 > \tau_2 > \dots > \tau_{k-2} > 1$. Then the p_α are indexed by all $\alpha < \tau_1$, and (13) corresponds to the relation

$$p_{\tau_1}^{n_1} p_{\tau_2}^{n_2} \dots p_{\tau_{k-1}}^{n_{k-1}} p_0^{n_k} = 0. \quad (14)$$

It only remains to construct a ring with these properties. To do this we take a J -ring of type τ_1 over any field, and localize at the set of all polynomials with non-zero constant term (using Prop. 3.2.7). Now add the relations (14), where the x_α are to be replaced by the p_α .

Exercises 8.8

1°. Show that a skew polynomial ring $R[x; \alpha]$ is a principal right ideal domain if and only if R is a principal right ideal domain, α is injective and $R^\alpha \subseteq U(R) \cup \{0\}$. Show that the 'if' part still holds for rings $R[x; \alpha, \delta]$; what about the 'only if' part?

2. Show that if $R[x; \alpha]$ is a semifir, then R is a semifir and α maps R into $U(R) \cup \{0\}$.

3. (Jategaonkar [69]) Let P be the localization of a J -ring of type τ at the set of all polynomials with non-zero constant term. Show that the Jacobson radical of P is $J = x_1 P$. If transfinite powers of J are defined by the equations $J^{\alpha+1} = J^\alpha J$, and $J^\alpha = \bigcap_{\beta < \alpha} J^\beta$ at a limit ordinal α , show that $J^\alpha \supseteq x_\alpha P$. Deduce that $J^\alpha \neq 0$ if and only if $\alpha < \tau$. Show also that J^α is a two-sided ideal.

4. (Jategaonkar [69]) Let R be a J -ring of type τ . Show that the elements $1 + x_\alpha$ ($\alpha < \tau$) are left linearly independent over R . If τ is a limit ordinal, show that every non-zero right ideal contains an ideal of the form $x_\alpha R$. Show also that $\alpha = \Sigma R(1 + x_\alpha)$ is a proper left ideal and that no maximal left ideal containing α can contain a non-zero ideal. Deduce that in this case R is left but not right primitive.

5. Determine the strong prime ideals in a J -ring of given type.

6. Show that a left Ore domain whose right ideals are well-ordered (under inclusion) is a principal ideal domain.

7. Show that a ring whose right ideals are well-ordered and which has no non-zero nilpotent elements is an integral domain.

8. (Brungs [69]) Let R be a ring whose right ideals are well-ordered of type σ . Show that R is an integral domain if and only if $\sigma = \omega^\tau$, and that R is left Noetherian if and only if $\sigma \leq \omega$.

9*. Let k be a commutative field with an endomorphism α such that k contains an element t transcendental over k^α , and denote by K the subring of $k(y)$ consisting of all fractions $f \cdot (1 + yg)^{-1}$, where $f, g \in k[y]$. Show that α can be extended to K by letting $y \mapsto t$, and verify that the resulting endomorphism maps $K^\times$ into $U(K)$. Show that $R = K[x; \alpha]$ is a principal right ideal domain in which the right ideals are well-ordered, and determine the order type of its chain of right ideals (cf. Cohn [67]).

10. Let R be a principal right ideal domain which is also left Ore. If all atoms of R are right associated to a single one p , say, show that $J(R) = pR$. Deduce that R is a right principal valuation ring.

11. Let F be the free group on x, y , with the ordering defined as in the text, and let S be the monoid of elements ≥ 1 . Show that S is conical and rigid, but not atomic. Deduce that S is locally free, i.e. every finitely generated submonoid is contained in a free submonoid of S .

Notes and comments on Chapter 8

There are many papers on the diagonal reduction of matrices. A weak form of Th. 1.1 was proved for (non-commutative) Euclidean domains by Wedderburn [32], and the full form by Jacobson [37]. This was generalized to principal ideal domains by Teichmüller [37] and a uniqueness statement added by Nakayama [38]. Our account in 8.1 and 2 follows Jacobson [43], with some simplifications (*cf.* Amitsur [63]). For a general study of elementary divisor rings see Kaplansky [49]. An example of a principal ideal domain for which the non-elementary operation (iv) is needed is the ring of integers in $\mathbb{Q}(\sqrt{-19})$, which is therefore not Euclidean, though principal, *cf.* Cohn [66]. Lemma 2.7 appears to be folklore; I learnt of it through Stafford and Schofield.

The results of 8.3 are for the most part well-known; for Th. 3.5 see Cohn [77']. Section 8.4 follows Jacobson [37] and 8.5 follows Cohn [73]; Cor. 5.4 is implicit in Jacobson [64], p. 241. Prop. 5.5 is new in the form given here; it generalizes a result of Wedderburn [21] (wrongly ascribed to L. E. Dickson in the first edition) which is the 1×1 case of Cor. 5.6. In Cohn [73] it is shown that every square matrix has left and right eigenvalues in a suitable extension field; for an interesting geometrical interpretation of eigenvalues see Riesinger [82]. The Cayley–Hamilton theorem has been generalized in the following form: If A is cyclic over K , in the sense that $xI - A$ is stably associated to $p(x)$ over $K_k\langle x \rangle$, then $p(A) = 0$; *cf.* Cohn [76].

The applications to skew field extensions in the first half of 8.6 are for the most part due to Gordon and Motzkin [65], generalizing earlier results by Richardson [27] and Herstein [56]. Our account follows Cohn [70]. Prop. 6.5 appears in Bray and Whaples [83] (where the question of multiple zeros discussed in the remark following Prop. 6.5 is raised but left open). Pseudo-linear extensions were introduced by Cohn [61'] to construct field extensions of different left and right dimensions (*cf.* Ex. 6.9 and 10). Prop. 6.8 and 9 are well known.

Polynomial rings with the commutation rule (9) in 8.7 have been studied by Smits [68], to whom Th. 7.2 is due. The use of generalized Laurent series has a long history; infinite series with support other than $\mathbb{N}$ were considered by Levi-Civita [1892] and were used by Hilbert [1899] to construct an ordered field that was not commutative. Prop. 7.1 is implicit in Schur [04]; Hahn [07] showed that every totally ordered abelian group Γ can be embedded in a (lexicographically ordered) ordinal power of $\mathbb{R}$, and in the same paper introduced the ordered series ring $\mathbb{R}((\Gamma))$. Th. 7.5 (for groups) was proved independently by Malcev [48] and Neumann [49]; our proof follows the former source, but is stated for monoids (the first edition gave no proof but referred to Cohn [81], where a proof is given, based on a more general result of Higman [52], on algebras with a divisibility ordering). The method of ordering a free group in 8.7 was suggested by Bergman [b], and Th. 7.7 is due to Bergman [78]. The notion of a unique remainder algorithm is described in Lenstra [74], to whom Prop. 8.1 and Th. 8.3 are due, but the latter result was also proved independently by Korotkov [76]. Jategaonkar [69] constructed his J -rings by transfinite induction; the more direct proof given here was new to the first edition. Th. 8.6 (for rings with well-ordered set of right ideals) is due to Brungs [69], who also described their structure.

Appendix

This appendix gives a brief summary of facts needed from lattice theory and homological algebra, with references to proofs or the proofs themselves, when no convenient reference was available. In each section some reference books are listed, with an abbreviation which is used in quoting them in the appendix.

1. Lattice theory

References

LT. G. Birkhoff, *Lattice Theory*, 3rd ed. (Providence, R.I. 1967).

UA. P. M. Cohn, *Universal Algebra*, 2nd ed. (Dordrecht 1981).

(i) We recall that a *lattice* is a partially ordered set in which any pair of elements a, b has a supremum (least upper bound), also called *join* and written $a \vee b$, and an infimum (greatest lower bound), also called *meet* and written $a \wedge b$. It follows that in a lattice L every finite non-empty subset has a sup and an inf; if every subset has a sup and an inf, L is said to be *complete*. It is possible to define lattices as algebras with two binary operations $\vee, \wedge$ satisfying certain identities, so that lattices form a variety of algebras (LT, p. 9, UA, p. 63).

If we reverse the ordering in a lattice, we again obtain a lattice, in which meet has been replaced by join, and vice versa. This is the basis of the principle of duality in lattice theory, by which we obtain from each theorem about lattices another one which is its dual.

Given any elements $x \leq y$ in a lattice L , we define the *interval* from x to y as the set $[x, y] = \{z \in L \mid x \leq z \leq y\}$; it is a sublattice of L , i.e. closed under the operations of join and meet in L . A minimal element $> x$ is said to *cover* x .

(ii) Nearly all the lattices we deal with will be *modular*, i.e. they satisfy the *modular law* (Dedekind [1900]):

$$(x \vee y) \wedge z = x \vee (y \wedge z) \quad \text{for all } x, y, z \text{ such that } x \leq z. \quad (1)$$

An important example of a modular lattice, which will much occupy us here (and which incidentally is responsible for the name) is the set $\text{Lat}_R(M)$ of

submodules of an R -module M , partially ordered by inclusion. A criterion for modularity is provided by

PROPOSITION A.1. *A lattice is modular if and only if it contains no sublattice isomorphic to the 5-element lattice of Fig. A1 (cf. LT, p. 13, UA, p. 66). ■*

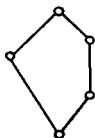


FIGURE A1

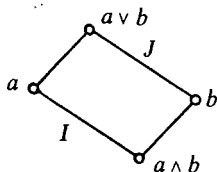


FIGURE A2

The following well-known consequence of (1) will be used without further reference:

Let M be a module and M' a submodule. Then any direct summand of M which is contained in M' is a direct summand of M' .

For if $M = P \oplus Q$ and $P \subseteq M'$, then by (1),

$$M' = (P \oplus Q) \cap M' = P \oplus (Q \cap M'). \quad \blacksquare$$

If a, b are any two elements of a modular lattice L , there is an isomorphism between the intervals $I = [a \wedge b, a]$ and $J = [b, a \vee b]$, given by the mapping $x \mapsto x \vee b$, with inverse $y \mapsto y \wedge a$. This fact is often referred to as the *parallelogram law* (the proof is an easy exercise, cf. LT, p. 13, UA, p. 69). Two intervals related in this way are said to be *perspective*; more generally, two intervals are said to be *projective* if we can pass from one to the other by a series of perspectivities.

The parallelogram law shows that projective intervals are isomorphic as lattices. This may not tell us much (e.g. when the intervals refer to simple modules), but in the module case a stronger assertion can be made: by the Noether isomorphism theorem perspective intervals give isomorphic module quotients: $(a+b)/b \cong a/(a \cap b)$, hence by induction, so do projective intervals.

Many basic theorems on modules are lattice-theoretic in nature, in the sense that the results can be stated and proved in terms of lattices. This is true of the next group of theorems, which are all used in the text (mainly Chapter 3):

SCHREIER REFINEMENT THEOREM. *In a modular lattice, two finite chains between the same end-points have refinements that are isomorphic, in the sense that their intervals can be paired off in such a way that corresponding intervals are projective (LT, p. 66, UA, p. 70). ■*

The result shows in particular that any two maximal chains between given end-points in a modular lattice have the same length. By the *length* of a lattice L one understands the supremum of the lengths of chains in L , the length of a finite chain being the number of its links. A lattice of finite length necessarily has a greatest element, written 1, and a least element, written 0. The length of the interval $[0, a]$ is also called the *height* of a . A modular lattice is of finite length whenever it satisfies the ascending and descending chain conditions, for then all its maximal chains are finite and of equal length. As in group theory, the Schreier refinement theorem has the following consequence:

JORDAN-HÖLDER THEOREM. *In a modular lattice of finite length, any chain can be refined to a maximal chain, and any two maximal chains are isomorphic (UA, p. 70). ■*

(iii) A lattice L is said to be *distributive*, if

$$(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z) \quad \text{for all } x, y, z \in L. \quad (2)$$

This is easily seen to be equivalent to its dual (obtained by interchanging $\vee$ and $\wedge$) and to imply (1), so that every distributive lattice is modular. Like modularity, distributivity can be characterized by the non-existence of a certain type of sublattice.

PROPOSITION A.2. *A lattice L is distributive if and only if it is modular and contains no 5-element sublattice of length 2 (LT, p. 14, UA, p. 69). ■*

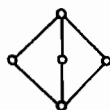


FIGURE A3

There is precisely one 5-element modular lattice of length 2, up to isomorphism; it is shown in Fig. A3. Prop. A.2 shows that a lattice is distributive if and only if relative complements, when they exist, are unique.

(iv) Distributive lattices are much more special than modular ones, as is clear from Chapters 3 and 4. This is also apparent from the special form taken

by the next two theorems in the distributive case. An element a in a lattice with least element 0 is said to be expressed as a *join of independent elements*: $a = a_1 \vee \dots \vee a_n$, if $a_i \neq 0$ and

$$a_i \wedge \left(\bigvee_{j \neq i} a_j \right) = 0 \quad \text{for } i = 1, \dots, n.$$

KRULL-SCHMIDT THEOREM. *In a modular lattice of finite length, let*

$$1 = a_1 \vee \dots \vee a_m$$

and

$$1 = b_1 \vee \dots \vee b_n$$

be two representations of 1 as a join of independent elements. Then $m = n$ and for some permutation $i \mapsto i'$ of $1, 2, \dots, n$, $[0, a_i]$ is projective with $[0, b_{i'}]$. In the case of a distributive lattice projectivity may be replaced by equality: $a_i = b_{i'}$. (LT, p. 168, UA, p. 73). ■

A finite decomposition $a = a_1 \vee \dots \vee a_n$ is called *irredundant*, if no a_i can be omitted. If $a \neq 0$ and no irredundant decomposition of a with more than one term exists, a is said to be *join-irreducible*.

KUROSH-ORE THEOREM. *In a modular lattice L , let*

$$c = p_1 \vee \dots \vee p_r$$

and

$$c = q_1 \vee \dots \vee q_s$$

be two irredundant decompositions of c into join-irreducible elements. Then $s = r$ and the p 's may be exchanged against the q 's, i.e. after suitable renumbering of the q 's we have

$$c = q_1 \vee \dots \vee q_i \vee p_{i+1} \vee \dots \vee p_r \quad (i = 1, \dots, r).$$

If, moreover, L is distributive, then the p 's and q 's are equal except for their order (LT, p. 58, UA, p. 76f.). ■

It is clear how an algebraic notion like 'homomorphism' is to be interpreted for lattices, namely as a join-and-meet-preserving mapping. An order-preserving mapping of lattices need not be a lattice-homomorphism, but an order-preserving bijection with an order-preserving inverse is a lattice-isomorphism, because the lattice structure can be defined in terms of the ordering.

A homomorphism of a modular lattice L which collapses an interval I (i.e. identifies its end-points) will clearly collapse all intervals perspective with I , and hence all intervals projective with I . Conversely, if we collapse all intervals projective with a given one, we obtain a homomorphic image of L . Thus each congruence on L (i.e. each collection of inverse image sets of a homomorphic image) is a union of projectivity classes of intervals of L .

(v) The *direct product* of lattices L_i is the Cartesian product $\prod L_i$ on which the operations $\vee, \wedge$ are defined componentwise. A sublattice of $\prod L_i$ which projects onto each factor L_i is called a *subdirect product* of the L_i . If L can be written as a subdirect product of the family (L_i) where none of the projections $L \rightarrow L_i$ is an isomorphism, L is said to be *subdirectly reducible*, otherwise it is *subdirectly irreducible*.

BIRKHOFF'S REPRESENTATION THEOREM. *Every modular lattice L is a subdirect product of a family (possibly infinite) of subdirectly irreducible modular lattices, which are homomorphic images of L . A corresponding result holds for distributive lattices (LT, p. 193, UA, p. 100). ■*

2. Categories and homological algebra

References

- H: S. Mac Lane, *Homology* (Berlin–Göttingen–Heidelberg 1963).
 C: S. Mac Lane, *Categories for the Working Mathematician* (New York–Heidelberg–Berlin 1971).
 A: P. M. Cohn, *Algebra II* (Chichester 1977).

(i) We shall not give a formal definition of a category here; let us just recall that if one merely looks at morphisms or maps (by identifying objects with their identity maps) a *category* is a class (i.e. a 'big' set) with a multiplication not necessarily everywhere defined, but where defined it is associative, with left and right neutrals for multiplication. In diagrams each map is represented by an arrow going from the *source* to the *target*. If the collection of objects is a set, we have a *small category*; a small category in which there is at most one map between any two objects is essentially a partially ordered set. The anti-isomorph of a category $\mathcal{C}$ (obtained by reversing all the arrows) is again a category, called the *opposite* of $\mathcal{C}$ and denoted by $\mathcal{C}^o$. Most of the categories in this book are categories of modules.

Let $\mathcal{C}$ be any category and write $\mathcal{C}(X, Y)$ for the set of all maps $X \rightarrow Y$ with source X and target Y . A *subcategory* $\mathcal{B}$ is defined (as in algebra) as a subclass closed under multiplication when defined, and containing with any map its left and right neutrals. If for any objects X, Y in $\mathcal{B}$, we have

$\mathcal{B}(X, Y) = \mathcal{C}(X, Y)$, $\mathcal{B}$ is said to be *full*; clearly a full subcategory is determined by its objects alone.

In any category a map is called an *isomorphism* if it has a 2-sided inverse, and two objects are *isomorphic* if there is an isomorphism between them. An object X_0 in a category $\mathcal{C}$ is said to be *initial* if for each object Y there is precisely one map $X_0 \rightarrow Y$; clearly any two initial objects are isomorphic. Dually, an initial object in the opposite category $\mathcal{C}^o$ is called a *final* object for $\mathcal{C}$. An object which is both initial and final is called a *zero* object and is written 0.

(ii) In any category consider a diagram consisting of two maps with the same target: $\alpha \downarrow \beta$. The different ways of completing this figure to a commutative square $\square$ form themselves the objects of a category, in which the 'maps' are maps between the new objects added to get a 'commutative wedge'. A final object in this category is called the *pullback* of α and β . Thus, the pullback consists of a pair of maps α', β' with the same source, such that $\alpha'\beta = \beta'\alpha$, and 'universal' with this property; in other words, the pullback of α and β is just their least common left multiple (when it exists). *Pushouts* are defined dually, as the least common right multiple of two maps with the same source (cf. C, p. 65ff., A, p. 90).

(iii) A category $\mathcal{A}$ is said to be *additive* if (i) $\mathcal{A}(X, Y)$ is an abelian group such that composition is distributive when defined: $\alpha(\beta + \beta') = \alpha\beta + \alpha\beta'$, $(\alpha + \alpha')\beta = \alpha\beta + \alpha'\beta$, (ii) there is a zero-object and (iii) to each pair of objects X_1, X_2 there corresponds an object S and maps as shown in (3)

$$\begin{array}{ccccc} X_1 & \xrightarrow{i_1} & S & \xrightarrow{p_2} & X_2 \\ & \searrow p_1 & & \swarrow i_2 & \\ & & & & \end{array} \quad (3)$$

such that $i_1 p_1 = 1$, $i_2 p_2 = 1$, $p_1 i_1 + p_2 i_2 = 1$. The object S is called the *direct sum* of X_1 and X_2 with injections i_v and projections p_v . By (i) and (ii) any finite family has a direct sum, and conversely this condition ensures (i) and (ii). If only (i) holds, $\mathcal{A}$ is called *preadditive*; a small preadditive category is called a *ringoid*, in the case of a single object we just have a ring.

Let $\mathcal{A}$ be an additive category; with each map $f: X \rightarrow Y$ we can associate a new category whose objects are maps α of $\mathcal{A}$ with target X and satisfying $\alpha f = 0$, and whose maps, between α and α' , are maps λ from the source of α to that of α' such that $\alpha = \lambda \alpha'$. A final object μ in this category (when one exists) is called a *kernel* of f , written $\ker f$; it is always a *monomorphism* or *monic*, i.e. $\xi \mu = \xi' \mu$ implies $\xi = \xi'$. Hence its source is unique up to isomorphism, and we shall sometimes use $\ker f$ to indicate this source. The dual notion is the *cokernel* of f , written $\operatorname{coker} f$. It is an *epimorphism* or *epic*, i.e. it can be cancelled whenever it appears as left-hand factor.

Given $f: X \rightarrow Y$, we define $\text{im} f = \ker \text{coker} f$, $\text{coim} f = \text{coker} \ker f$; of course, they need not exist, but when they do, we have the following picture, where we have identified $\ker f$ and $\text{im} f$ with their source, and $\text{coker} f$, $\text{coim} f$ with their target:

$$\begin{array}{ccccc} \ker f & \longrightarrow & X & \xrightarrow{f} & Y & \longrightarrow & \text{coker} f \\ & & \downarrow & \nearrow & \uparrow & & \\ & & \text{coim} f & \longrightarrow & \text{im} f & & \end{array}$$

Here the map $\text{coim} f \rightarrow \text{im} f$ follows from the definition of im and coim . If the additive category $\mathcal{A}$ is such that every map has a kernel and cokernel and the natural map $\text{coim} f \rightarrow \text{im} f$ is an isomorphism, then $\mathcal{A}$ is said to be *abelian*. We remark that the notion of an abelian category is again self-dual; observe also that in an abelian category a map which is both epic and monic is necessarily an isomorphism. The category of all modules over a ring is an abelian category, as is easily seen; of course, not every abelian category is of the form $\mathcal{M}_R$; for example, if R is any non-zero ring, then the opposite of $\mathcal{M}_R$ is not of the form $\mathcal{M}_S$ for any ring S .

We note the following easy consequence of the definitions:

PROPOSITION A.3. *Let $\mathcal{A}$ be an abelian category and $\mathcal{B}$ a full subcategory. Then $\mathcal{B}$ is abelian if and only if the direct sums, kernels and cokernels, taken in $\mathcal{A}$, of maps in $\mathcal{B}$ again lie in $\mathcal{B}$. ■*

We remark that in any 'concrete' category (where the objects have an underlying set structure) a monomorphism is necessarily injective; this simplifies the terminology. However, an epimorphism need *not* be surjective, thus in the category of rings the embedding $\mathbb{Z} \rightarrow \mathbb{Q}$ is an epimorphism.

(iv) A sequence of maps in an abelian category

$$\dots \longrightarrow X_{i-1} \xrightarrow{f_{i-1}} X_i \xrightarrow{f_i} X_{i+1} \dots$$

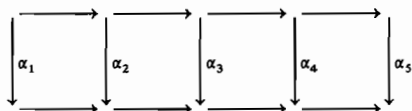
is *exact* at X_i if $\ker f_i = \text{im} f_{i-1}$; if it is exact at each object, it is called *exact*. An exact sequence beginning and ending at 0 cannot have just one non-zero term, and if it has two, they must be isomorphic. Thus the first non-trivial case is that of a 3-term exact sequence:

$$0 \rightarrow A' \rightarrow A \rightarrow A'' \rightarrow 0, \quad (4)$$

also called a *short exact sequence*. For example, in the case of modules, (4) represents an extension: A is an extension of A' by A'' .

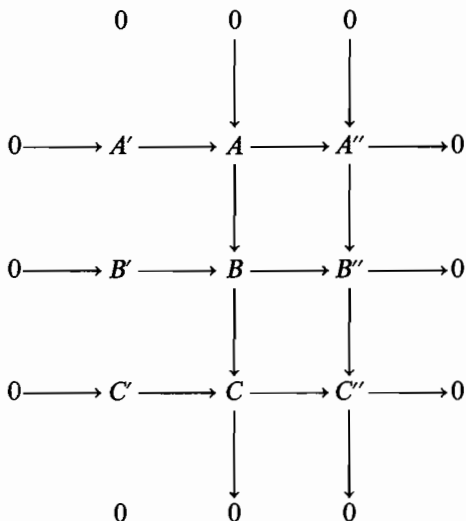
The following two results are easily proved for modules by a diagram chase (and this is all we need), but in fact they hold in any abelian categories:

FIVE-LEMMA. *Given a commutative diagram with exact rows*



if α_1 is epic and α_2, α_4 are monic, then α_3 is monic. Dually, if α_5 is monic and α_2, α_4 epic, then α_3 is epic. In particular, if $\alpha_1, \alpha_2, \alpha_4, \alpha_5$ are isomorphisms, then so is α_3 (H, p. 14, C, p. 201). ■

THREE-BY-THREE LEMMA. *Given a commutative diagram with exact rows and columns as shown,*



there is a unique way of filling in the first column so as to keep the diagram commutative and then the new column is exact too (H, p. 49). ■

(v) With every type of algebraic system there is associated the notion of a homomorphism, as a structure-preserving mapping. In the case of categories one speaks of functors. Thus a functor $F: \mathcal{A} \rightarrow \mathcal{B}$ is a mapping from one category, $\mathcal{A}$, to another, $\mathcal{B}$, which preserves neutrals and satisfies

$$F(\alpha\beta) = F\alpha \cdot F\beta. \quad (5)$$

If instead of (5) we have

$$F(\alpha\beta) = F\beta \cdot F\alpha, \quad (6)$$

F is said to be *contravariant*, in contrast to the sort defined by (5), which are also called *covariant*. Thus a contravariant functor from $\mathcal{A}$ to $\mathcal{B}$ may also be regarded as a covariant functor from $\mathcal{A}^0$ to $\mathcal{B}$, or from $\mathcal{A}$ to $\mathcal{B}^0$. If F and G are two covariant functors from $\mathcal{A}$ to $\mathcal{B}$, a *natural transformation* $t: F \rightarrow G$ is a function which assigns to each $\mathcal{A}$ -object X a $\mathcal{B}$ -map $t_X: FX \rightarrow GX$ such that for each map $\alpha: X \rightarrow Y$ the square

$$\begin{array}{ccc} FX & \xrightarrow{\quad} & FY \\ \downarrow t_X & & \downarrow t_Y \\ GX & \xrightarrow{\quad} & GY \end{array}$$

commutes. If each t_X is an isomorphism, t is called a *natural isomorphism*; if t_X is an inclusion map, F is called a *subfunctor* of G . A natural anti-isomorphism is also called a *duality*.

Given categories $\mathcal{A}$ and $\mathcal{B}$, if there is a pair of functors $F: \mathcal{A} \rightarrow \mathcal{B}$, $G: \mathcal{B} \rightarrow \mathcal{A}$ such that the compositions (from left to right) $FG: \mathcal{A} \rightarrow \mathcal{A}$ and $GF: \mathcal{B} \rightarrow \mathcal{B}$ are naturally isomorphic to the identity functor: $FG \simeq 1$, $GF \simeq 1$, then $\mathcal{A}$ and $\mathcal{B}$ are said to be *equivalent*, and F, G are *equivalences* between them.

(vi) Let $\mathcal{A}$ be a class of algebras and X a set. Then we can form the *comma category* $(X, \mathcal{A})$ whose objects are maps $\alpha: X \rightarrow G$ from X to an $\mathcal{A}$ -object and whose morphisms from $\alpha: X \rightarrow G$ to $\beta: X \rightarrow H$ are $\mathcal{A}$ -homomorphisms $f: G \rightarrow H$ such that $\beta = \alpha f$. An initial object in $(X, \mathcal{A})$, if it exists, is called a *universal $\mathcal{A}$ -object* for X ; as initial object it is unique up to isomorphism. For example, the free group on a set X is a universal object in this sense. We remark that if with each group G we associate its underlying set $\sigma(G)$ by means of the 'forgetful' functor from groups to sets, then we have a natural bijection of morphism sets:

$$\text{Gp}(F_X, G) \cong \text{Ens}(X, \sigma(G)), \quad (7)$$

where Gp , Ens denote the categories of groups and sets, respectively. More generally, a pair of functors $T: \mathcal{A} \rightarrow \mathcal{B}$, $S: \mathcal{B} \rightarrow \mathcal{A}$ is called an *adjoint pair*, and S *left adjoint*, T *right adjoint*, if for any $\mathcal{A}$ -object X and $\mathcal{B}$ -object V ,

$$\mathcal{A}(V^S, X) \cong \mathcal{B}(V, X^T). \quad (8)$$

It can be shown that each of S, T determines the other up to natural isomorphism by (8) (H, p. 266, A, p. 95). For example, the free group functor

is the left adjoint of the forgetful functor σ , by (7). More generally, almost every universal construction arises as the left adjoint of a forgetful functor.

The following lemma is often useful in discussing universal constructions; although we state it for rings, it clearly holds quite generally for all types of algebras:

LEMMA. Let A, B be any rings and f, g two homomorphisms from A to B . If f and g agree on a generating set of A , then $f = g$.

For the proof we note that the set $\{x \in A \mid xf = xg\}$ is easily seen to be a subring; since it contains a generating set, it must be the whole of A . ■

(vii) We shall mainly be concerned with additive (in particular abelian) categories; in that case all functors are assumed to be *additive*, i.e. $F(\alpha + \beta) = F\alpha + F\beta$. For example, $\text{Hom}_R(-, -)$ and $- \otimes -$ are additive functors. In what follows, all functors are tacitly assumed to be additive.

Any functor F transforms an exact sequence

$$\xrightarrow{\alpha} A \xrightarrow{\beta} \quad (9)$$

into a sequence

$$\xrightarrow{F\alpha} FA \xrightarrow{F\beta} \quad (10)$$

with composition zero: $F\alpha \cdot F\beta = 0$, but (10) need not be exact. If (9) is split exact, i.e. if $\text{im } \alpha = \ker \beta$ is a direct summand of A , then the same is true of (10). The functor F is said to be *exact* if it transforms any exact sequence into an exact sequence.

Exact functors are rare; thus $\text{Hom}_R(P, -)$ is exact if and only if P is projective. This may be taken as the definition of a projective module, and it can then be proved that P is projective if and only if it is a direct summand of a free module. Similarly $\text{Hom}_R(-, I)$ is exact if and only if I is injective. There is no simple description of injective modules, as there is for projectives, but in testing for injective modules, Baer's criterion is often useful. This states that a right R -module I is injective if and only if every homomorphism $a \rightarrow I$, where a is a right ideal, can be extended to a homomorphism $R \rightarrow I$ (A, p. 107).

Every module can be embedded in an injective module, and the least injective containing a given module M is unique up to isomorphism and is called the *injective hull* of M . There is an alternative definition of injective module and injective hull as a sort of algebraic closure, but this will not be needed here (cf. UA, p. 261, A, p. 152).

(viii) Although few functors are exact, many have a partial exactness property, which we now describe:

A covariant functor F is *left exact*, if the sequence

$$0 \rightarrow FA' \rightarrow FA \rightarrow FA'' \rightarrow 0, \quad (11)$$

obtained by applying F to the short exact sequence (4), is exact except possibly at FA'' , and F is *right exact* if (11) is exact except possibly at FA' . For a contravariant functor $D: \mathcal{A} \rightarrow \mathcal{B}$ these notions are defined by applying the definitions just given to the associated covariant functor from $\mathcal{A}^o$ to $\mathcal{B}$. A routine verification shows that $\text{Hom}_R(-, -)$ is left exact in each of its arguments.

With each functor F we can associate a series of derived functors F^i ($i \in \mathbb{Z}$) which measure the departure of F from exactness. When F is exact, $F^i = 0$ for $i \neq 0$ and F^0 is naturally isomorphic to F , in symbols $F^0 \cong F$. For a left exact functor F , $F^i = 0$ for $i < 0$ and (i) $F^0 \cong F$, (ii) $F^n I = 0$ for $n > 0$ and I injective, and (iii) to each short exact sequence (4) there corresponds a long exact sequence

$$0 \rightarrow F^0 A' \rightarrow F^0 A \rightarrow F^0 A'' \xrightarrow{\Delta} F^1 A' \rightarrow F^1 A \rightarrow \dots \quad (12)$$

with a 'connecting' homomorphism Δ which is a natural transformation. Moreover, F^i is uniquely determined up to natural isomorphism by these three properties.

As an example let us take $FA = \text{Hom}_R(M, A)$. The derived functor in this case is denoted by Ext . Thus $\text{Ext}_R^0(M, N) \cong \text{Hom}_R(M, N)$, $\text{Ext}_R^i(M, I) = 0$ for $i > 0$ and I injective, and we have an exact sequence

$$0 \rightarrow \text{Hom}_R(M, A') \rightarrow \text{Hom}_R(M, A) \rightarrow \text{Hom}_R(M, A'') \rightarrow \text{Ext}_R^1(M, A') \\ \rightarrow \text{Ext}_R^1(M, A) \rightarrow \text{Ext}_R^1(M, A'') \rightarrow \text{Ext}_R^2(M, A') \rightarrow \dots \quad (13)$$

Next let us take $\text{Hom}_R(M, N)$, regarded as a functor in M . This is contravariant, and we have to replace injectives in (ii) above by projectives. The derived functor is the same as before, namely Ext . Thus we have $\text{Ext}_R^i(P, N) = 0$ for $i > 0$ and P projective, and from any short exact sequence (4) we obtain

$$0 \rightarrow \text{Hom}_R(A'', N) \rightarrow \text{Hom}_R(A, N) \rightarrow \text{Hom}_R(A', N) \rightarrow \text{Ext}_R^1(A'', N) \\ \rightarrow \text{Ext}_R^1(A, N) \rightarrow \text{Ext}_R^1(A', N) \rightarrow \text{Ext}_R^2(A'', N) \rightarrow \dots \quad (14)$$

Whenever R is a field or, more generally, when R is a semi-simple Artinian ring, $\text{Ext}^i = 0$ for all $i \neq 0$, because for such rings all exact sequences split, so Hom is then exact in each argument. Most of the rings considered in this book are hereditary, which amounts to saying that Ext^i vanishes identically

for $i > 1$. Therefore (13) and (14) reduce to 6-term sequences in this case. For any module M , the *projective (injective) dimension* is defined as the least $n \geq 0$ such that $\text{Ext}_R^{n+1}(M, -) = 0$ ($\text{Ext}_R^{n+1}(-, M) = 0$), or ∞ if no such n exists; they are denoted by $\text{pd. } M$, $\text{id. } M$ respectively. Now the *right global dimension* of R is defined as $\text{r.gl.dim. } R = \sup\{\text{pd. } M\}$, where M ranges over all right R -modules, and the *left global dimension* is defined similarly. In a Noetherian ring these two dimensions are equal, but in general they may be different (cf. 8.8).

Let us briefly consider the connexion of Ext with module extensions. In (13) put $M = A''$ and consider the image of $1 \in \text{Hom}_R(A'', A'')$ under Δ ; this is an element θ of $\text{Ext}_R^1(A'', A')$. It can be shown (H, Ch. 3, A, p. 138) that two short exact sequences (4) and

$$0 \rightarrow A' \rightarrow B \rightarrow A'' \rightarrow 0 \quad (15)$$

give rise to the same θ if and only if the extensions A and B are isomorphic in the sense that there is an isomorphism $f: A \rightarrow B$ making the diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & A' & \longrightarrow & A & \longrightarrow & A'' \longrightarrow 0 \\ & & \downarrow 1 & & \downarrow f & & \downarrow 1 \\ 0 & \longrightarrow & A' & \longrightarrow & B & \longrightarrow & A'' \longrightarrow 0 \end{array}$$

commutative. Moreover, every element of $\text{Ext}_R^1(A'', A')$ gives rise to an extension in this way; so the isomorphism classes of extensions of A' by A'' may be classified by $\text{Ext}_R^1(A'', A')$. We remark that the element θ can also be obtained as the image of 1 under Δ by putting $N = A'$ in (14). This element θ is sometimes called the *obstruction*; it is zero if and only if the extension (4) splits.

(ix) Since every free module is projective, every module can be written as a homomorphic image of a projective module. Given a module M , we can write it as a homomorphic image of a projective module P_0 , say $f_0: P_0 \rightarrow M$. Similarly we can find a projective module P_1 mapping onto $\ker f_0$ and by continuing in this way we get an exact sequence of projective modules P_i :

$$\dots \rightarrow P_1 \rightarrow P_0 \rightarrow M \rightarrow 0, \quad (16)$$

called a *projective resolution* of M . Dually, M has an *injective resolution*

$$0 \rightarrow M \rightarrow I_0 \rightarrow I_1 \rightarrow \dots, \quad (17)$$

where the I_i are injective. If in (16) we do not insist on P_1 being projective, we get a short exact sequence

$$0 \rightarrow Q \rightarrow P \rightarrow M \rightarrow 0, \quad P \text{ projective.} \quad (18)$$

This is called a *presentation* of M ; different presentations of M are compared in

SCHANUEL'S LEMMA. *Let M be any module; if we have two presentations of M , say (18) and*

$$0 \rightarrow Q' \rightarrow P' \rightarrow M \rightarrow 0, \quad (19)$$

where P' is projective, then

$$P \oplus Q' \cong P' \oplus Q.$$

(cf. A, p. 123, and Th. 0.6.2, above.) ■

The same proof will give an extension which is sometimes useful:

EXTENDED SCHANUEL LEMMA. *If we are given two presentations of a module M , one of them by a projective module, say (18) and (19) where P is projective (but not necessarily P'), then there is an exact sequence*

$$0 \rightarrow Q \rightarrow P \oplus Q' \rightarrow P' \rightarrow 0. \quad \blacksquare$$

This shows in particular that if M is finitely presented, i.e. there is a presentation (18) with P, Q finitely generated, then in any presentation (19) of M in which P' is finitely generated, Q' is also finitely generated.

(x) The tensor product may be defined as the left adjoint of the Hom-functor. Thus, if modules $U, {}_R V_S$ are given, their tensor product $U \otimes_R V$ is determined up to isomorphism as right S -module by the equation

$$\text{Hom}_R(U, \text{Hom}_S(V, W)) \cong \text{Hom}_S(U \otimes_R V, W), \quad (20)$$

where W is a right S -module; the isomorphism (20) is also referred to as *adjoint associativity*.

Let $f: R \rightarrow S$ be a ring homomorphism. With each right R -module M we can associate an S -module, the *coinduced extension* $M^f = \text{Hom}_R(S, M)$ (cf. A, p. 102). If I is an injective right R -module, then I^f is an injective right S -module, for if A is any S -module, we have, by (20),

$$\text{Hom}_S(A, I^f) = \text{Hom}_R(A \otimes_S S, I) = \text{Hom}_R(A, I),$$

hence $\text{Hom}_S(-, I^f)$ is an exact functor, and the assertion follows.

The tensor product functor $U \otimes -$ is always right exact, but not necessarily exact. If $U \otimes -$ is an exact functor, the module U is said to be *flat*. For example, any free module, and more generally any projective module, is flat. We shall need the following characterization of flatness:

PROPOSITION A.4. A right R -module U is flat if and only if, whenever

$$uc = 0, \quad \text{where } u \in U^n, c \in {}^nR,$$

there exist $B \in {}^mR^n$ and $x \in U^m$ such that

$$u = xB, \quad Bc = 0$$

(A, p. 147). ■

More generally, it can be shown in the same way that if U is flat and V is any left R -module, and if

$$\sum u_j \otimes v_j = 0 \text{ in } U \otimes V,$$

where u_j, v_j are the components of $u \in U^n, v \in {}^nV$, respectively, then there exists $B \in {}^mR^n, x \in U^m$ such that

$$u = xB, \quad Bv = 0.$$

COROLLARY A.5. If U is a flat right R -module and α is a left ideal of R , then the natural mapping

$$v: U \otimes \alpha \rightarrow U\alpha \quad (21)$$

is an isomorphism.

For clearly v is surjective; if $\sum u_i c_i = 0$ for $u_i \in U, c_i \in \alpha$, then $u_i = \sum x_j a_{ji}$, $\sum a_{ji} c_i = 0$ for some $a_{ji} \in R, x_j \in U$ by Prop. A.4, hence in $U \otimes \alpha, \sum u_i \otimes c_i = \sum x_j a_{ji} \otimes c_i = \sum x_j \otimes a_{ji} c_i = 0$, therefore v is injective. ■

From Prop. A.4 we see that a module U is flat if and only if every finitely generated submodule of U is contained in a flat submodule; in other words, flatness is a local property. The definition shows that U is flat if and only if $\text{Tor}_1(U, -) = 0$, where Tor is the left derived of the tensor product. Here it is enough that $\text{Tor}_1(U, -)$ vanishes on finitely generated modules, by the remark just made. In fact, it is enough to require that $\text{Tor}_1(U, R/\alpha) = 0$ for every finitely generated left ideal α of R (because $\otimes$ and hence also Tor commutes with direct limits).

The weak dimension of a module M is defined as

$$\text{wd. } M = \sup\{i \mid \text{Tor}_i^R(-, M) \neq 0\};$$

thus if $\text{wd. } M = r$ and we take a projective resolution (16) of M , then $\ker(P_{r-1} \rightarrow P_{r-2})$ is flat. Now the weak global dimension of a ring R is defined as $\text{w.gl.dim. } R = \sup\{i \mid \text{Tor}_i^R \neq 0\}$. Clearly this is left-right symmetric and is

majorized by the left and right global dimensions of R , while for Noetherian rings all three global dimensions coincide (A, p. 142).

(xi) A ring R is said to be *right coherent* if every finitely generated right ideal is finitely related. It follows easily that when R is right coherent, then every finitely generated submodule of a free right R -module is finitely related. Some equivalent formulations of coherence are given in the next result.

THEOREM A.6. *For any ring R the following assertions are equivalent:*

- (a) R is right coherent,
- (b) the right annihilator of any row vector over R is finitely generated, i.e. given $u \in R^n$, if $uB = 0$ for some $B \in {}^n R^I$, then there exist $C \in {}^n R^r$, $D \in {}^r R^I$ such that $B = CD$, $uC = 0$,
- (c) the dual of every finitely presented left R -module is finitely presented,
- (d) the direct product of any family of flat left R -modules is flat,
- (e) every direct power R^I is left flat,
- (f) the right annihilator of any matrix over R is finitely generated, i.e. given $A \in {}^m R^n$, if $AB = 0$ for some $B \in {}^n R^I$, there exist $C \in {}^n R^r$, $D \in {}^r R^I$ such that $B = CD$, $AC = 0$,
- (g) every finitely generated submodule of a free right R -module is finitely presented.

Proof. (a) $\Rightarrow$ (b). By definition R is right coherent precisely if every finitely generated right ideal is finitely related. Consider $u = (u_1, \dots, u_n) \in R^n$ and let the family of all relations be $\sum u_i b_{i\alpha} = 0$. By (a) these relations can be generated by a finite subset; writing the $b_{i\alpha}$ as a matrix B with n rows and (possibly) infinitely many columns, we see that there exists $C \in {}^n R^r$ and a matrix D with r rows such that $B = CD$ and $uC = 0$.

(a) + (b) $\Leftrightarrow$ (c). Let V be finitely presented, say

$$R^m \xrightarrow{\alpha} R^n \longrightarrow V \longrightarrow 0,$$

where α is given by $A \in {}^m R^n$. Dualizing, we obtain the exact sequence

$$0 \rightarrow V^* \rightarrow {}^n R \rightarrow {}^m R,$$

and V^* is the annihilator of A in ${}^n R$. Now (b) is just the condition for V^* to be finitely generated; but then it is finitely presented by (a). Conversely, if V^* is finitely generated, (a) and so also (b) holds.

(b) $\Rightarrow$ (d). Let (V_α) be a family of flat left R -modules and consider $V = \prod V_\alpha$. If $av = 0$, where $a \in R^n$, $v \in {}^n V$, let $v = (v_\alpha)$, $v_\alpha \in {}^n V_\alpha$, then $av_\alpha = 0$; by flatness of V_α there exists $B_\alpha \in {}^n R^{m_\alpha}$, $v'_\alpha \in {}^{m_\alpha} V_\alpha$ such that $v_\alpha = B_\alpha v'_\alpha$, $aB_\alpha = 0$. By (b), $B_\alpha = CD_\alpha$, where $aC = 0$. Now $v = (D_\alpha v'_\alpha)$ and this shows V to be flat.

(d) $\Rightarrow$ (e) is clear. To prove (e) $\Rightarrow$ (f), suppose that $AB = 0$, where $A \in {}^m R^n$ and B has possibly infinitely many columns, then since R^I is left flat, there exists $C \in {}^n R^r$ such that $B = CD$ for some D with r rows in R^I and $AC = 0$.

(f) $\Rightarrow$ (g). Let U be a finitely generated submodule of a free module F . By omitting basis elements of F not involved in the generators of U , we can take F to be finitely generated. If U is generated by the columns of the $m \times n$ matrix A and B is the matrix of all relations between the columns of A , then by (e) there exist $C \in {}^n R^r$ and D with r rows such that $B = CD$, $AC = 0$. Hence U is finitely related, by the columns of C . Finally, (g) $\Rightarrow$ (a) is clear. ■

COROLLARY A.7. *Over a right coherent ring, the intersection of any two finitely generated submodules of a free right R -module is finitely generated.*

For if U, V are submodules of a free module F , generated by $u_1, \dots, u_m$ and $v_1, \dots, v_n$, respectively, then the elements of $U \cap V$ are all of the form

$$\sum u_i a_i = \sum v_j b_j, \quad \text{where } a_i, b_j \in R,$$

and they thus correspond to the relations $\sum u_i a_i - \sum v_j b_j = 0$. But the module of all these relations is finitely generated by (f), hence $U \cap V$ is indeed finitely generated. ■

We end by showing that over a coherent ring the finitely presented modules form an abelian category. We shall need a general lemma:

LEMMA A.8. *Let R be a ring and M a finitely presented R -module. Then any submodule M' of M is finitely generated if and only if M/M' is finitely presented.*

For we may take $M = F/K$, where F is free of finite rank and K is a finitely generated submodule, and take M' to be of the form G/K , where $K \subseteq G \subseteq F$. Since K is finitely generated, M' is finitely generated if and only if G is finitely generated, and this is the condition for $M/M' \cong F/G$ to be finitely presented. ■

We remark that if M is a finitely presented right R -module over a right coherent ring R , then any finitely generated submodule M' of M is finitely presented. For if we again write $M = F/K$, $M' = G/K$, then G , being a finitely generated submodule of F , is finitely presented by Th. A.6(g), and the lemma, applied to G and $K \subseteq G$ shows $M' = G/K$ to be finitely presented. This proves the first part of

THEOREM A.9. *If R is a right coherent ring, then the category of finitely presented right R -modules is closed under taking finitely generated submodules,*

finite direct sums, kernels and cokernels, hence it is an abelian subcategory of the category of all right R -modules.

To prove the remaining assertions, let $f: M \rightarrow N$ be an R -linear map between finitely presented right R -modules; then $\text{im} f$ is a finitely generated submodule of N and so is finitely presented. If we apply the lemma to $\text{im} f \cong M/\ker f$ and $\text{coker} f \cong N/\text{im} f$, we see that $\ker f$ is a finitely generated submodule of M , hence finitely presented, and $\text{coker} f$ is likewise finitely presented. The rest is clear. ■

Bibliography and author index

Apart from listing works referred to in the text, the bibliography includes papers related to the main topics, although there has been no attempt at complete coverage of such general topics as principal ideal domains. The page references at the end of the entries indicate the places in the text where the entry is quoted (or implicitly referred to); other references to an author are listed after his name.

ALBRECHT, F.

61. On projective modules over semihereditary rings. *Proc. Amer. Math. Soc.* **12** (1961), 638–639. [37]

ALMKVIST, G., DICKS, W. and FORMANEK, E.

85. Hilbert series of fixed free algebras and noncommutative classical invariant theory. *J. Algebra* **93** (1985), 189–214. [370]

ALPERIN, R. C.

79. Homology of the group of automorphisms of $k[x, y]$. *J. Pure and Applied Algebra* **15** (1979), 109–115. [379]

AMITSUR, S. A.

48. On unique factorization in rings (Hebrew). *Rivon Lematematika* **2** (1948), 28–29. [158]

58. Commutative linear differential operators. *Pacif. J. Math.* **8** (1958), 1–10. [63]

63. Remarks on principal ideal rings. *Osaka J. Math.* **15** (1963), 59–69. [539]

65. Generalized polynomial identities and pivotal monomials. *Trans. Amer. Math. Soc.* **114** (1965), 210–226. [282]

66. Rational identities, and applications to algebra and geometry. *J. Algebra* **3** (1966), 304–359. [487]

ANICK, D.

[343]

ARTIN, E.

57. "Geometric Algebra". Interscience, New York, 1957. [448]

ASANO, K.

38. Nichtkommutative Hauptidealringe. *Act. Sci. Ind.* No. 696. Hermann (Paris, 1938). [192]

AUSLANDER, M.

[297]

BARBILIAN, D.

56. "Teoria Aritmetica a Idealilor (in inele necomutative)". Ed. Acad. Rep. Pop. Romine, Bucuresti, 1956. [192]

- BASS, H. [226]
 64. Projective modules over free groups are free. *J. Algebra* **1** (1964), 367–373. [63]
 68. "Algebraic K-theory". Benjamin, New York, 1968. [5, 62f., 451]
- BASS, H., CONNELL, E. H. and WRIGHT, D.
 82. The Jacobian conjecture: reduction of degree and formal expansion of the inverse. *Bull. Amer. Math. Soc.* **7** (1982), 287–330. [380]
- BAUMSLAG, B. and BAUMSLAG, G.
 71. On ascending chain conditions. *Proc. London Math. Soc.* (3) **22** (1971), 681–704. [86]
- BAUMSLAG, G. (see also BAUMSLAG, B.) [379]
- BEAUREGARD, R. A.
 69. Infinite primes and unique factorization in a principal right ideal domain. *Trans. Amer. Math. Soc.* **141** (1969), 245–253. [192]
 71. Right LCM domains. *Proc. Amer. Math. Soc.* **30** (1971), 1–7.
 73. Overrings of Bezout domains. *Canad. Math. Bull.* **16** (1973), 475–477.
 80. Left versus right LCM domains. *Proc. Amer. Math. Soc.* **78** (1980), 464–466. [211]
- BEAUREGARD, R. A. and JOHNSON, R. E.
 70. Primary factorization in a weak Bezout domain. *Proc. Amer. Math. Soc.* **25** (1970), 662–665. [192]
- BECK, I.
 72. Projective and free modules. *Math. Zeits.* **129** (1972), 231–234. [23]
- BEDOYA, H. and LEWIN, J.
 77. Ranks of matrices over Ore domains. *Proc. Amer. Math. Soc.* **62** (1977), 233–236. [420]
- BEHRENS, E.-A.
 65. "Algebren". Bibliographisches Institut, Mannheim, 1965. [193]
- BERGMAN, G. M. [70f., 138, 151f., 164, 192, 226, 336, 341, 352, 371, 379, 395, 437, 487]
 67. "Commuting elements in free algebras and related topics in ring theory". Thesis, Harvard University, 1967.
 [50, 63, 85f., 121, 151f., 192, 204, 226, 232, 244, 251, 293f., 297, 340f., 378f., 487]
 69. Centralizers in free associative algebras. *Trans. Amer. Math. Soc.* **137** (1969), 327–344. [378]
 70. Skew fields of noncommutative rational functions, after Amitsur. *Séminaire Schützenberger–Lentin–Nivat, Année 1969/70*, No. 16. Paris, 1970. [487]
 70'. Notes on epimorphisms of rings. *Preprint*, Berkeley, 1970. [488]
 71. Groups acting on hereditary rings. *Proc. London Math. Soc.* (3) **23** (1971), 70–82. [124, 380]
 Corr. ibid. **24** (1972), 192.
 71'. Hereditary commutative rings and centres of hereditary rings. *Proc. London Math. Soc.* (3) **23** (1971), 214–236.
 72. Boolean rings of projection maps. *J. London Math. Soc.* (2) **4** (1972), 593–598. [297]
 72'. Hereditarily and cohereditarily projective modules. *Proc. Conf. on Ring Theory at Park City, Utah, 1971* (ed. R. Gordon), Academic Press, New York, 1972, pp. 29–62. [13, 15, 63, 297]

73. Infinite multiplication of ideals in $\aleph_0$ -hereditary rings. *J. Algebra* **24** (1973), 56–70. [297]
 74. Modules over coproducts of rings. *Trans. Amer. Math. Soc.* **200** (1974), 1–32. [151]
 - 74'. Coproducts and some universal ring constructions. *Trans. Amer. Math. Soc.* **200** (1974), 33–88. [85, 143, 152, 417]
 76. Rational relations and rational identities in division rings I, II. *J. Algebra* **43** (1976), 252–266, 267–297.
 78. The diamond lemma for ring theory. *Adv. in Math.* **29** (1978), 178–218. [148]
 - 78'. Conjugates and n th roots in Hahn–Laurent group rings. *Bull. Malaysian Math. Soc. (2)* **1** (1978), 29–41. Historical addendum *ibid.* **2** (1979), 41–42. [529, 531, 539]
 83. Sfields finitely right-generated over subrings. *Comm. in Algebra* **11** (17) (1983), 1893–1902.
 - a. Dependence relations and rank functions on free modules (to appear). [241, 297]
 - b. Ordering free groups and coproducts of ordered groups (to appear). [528, 539]
 - c. The class of free subalgebras of a free associative algebra is not closed under taking unions of chains, or pairwise intersections (to appear). [336]
- BERGMAN, G. M. and COHN, P. M.
69. Symmetric elements in free powers of rings. *J. London Math. Soc. (2)* **1** (1969), 525–534. [370]
 71. The centres of hereditary rings and 2-firs. *Proc. London Math. Soc. (3)* **23** (1971), 83–98. [379]
- BERGMAN, G. M. and DICKS, W.
75. Universal derivations. *J. Algebra* **36** (1975), 193–211. [110]
 78. Universal derivations and universal ring constructions. *Pacif. J. Math.* **79** (1978), 293–337. [58, 420, 459, 488]
- BERGMAN, G. M. and LEWIN, J.
75. The semigroup of ideals of a fir is (usually) free. *J. London Math. Soc. (2)* **11** (1975), 21–32. [379f.]
- BERSTEL, J., PERRIN, D. and SCHÜTZENBERGER, M.-P.
- a. Théorie des codes (to appear). [324, 379]
- BERSTEL, J. and REUTENAUER, C.
84. “Les Séries rationnelles et leurs Langages”. Masson, Paris, 1984. [379]
- BIRKHOFF, G.
67. “Lattice Theory” (3rd ed.). Amer. Math. Soc., New York, 1967. [80, 199, 226, 301, 540]
- BOKUT', L. A.
64. Some examples of rings without zero-divisors (Russian). *Algebra i Logika* **3** (5–6) (1964), 5–28. [63]
 65. Factorization theorems for certain classes of rings without zero-divisors I, II (Russian). *Algebra i Logika* **4** (1965). I. (4) 25–52, II. (5) 17–46.
 67. The embedding of rings in skew fields (Russian). *Dokl. Akad. Nauk SSSR* **175** (1967), 755–758. [486]

69. On Malcev's problem (Russian). *Sibirsk. Mat. Zh.* **10** (1969), 965–1005. [486]
81. "Associative Rings 1,2" (Russian). NGU, Novosibirsk, 1981. [486]
- BOURBAKI, N.
72. "Commutative Algebra", Ch. 1–7. Addison-Wesley, Reading, Mass., 1972. [86, 138, 303]
- BOWTELL, A. J.
67. "The multiplicative semigroup of a ring and the embedding of rings in skew fields". Thesis, London University, 1967. [86, 163, 192]
- 67'. On a question of Malcev. *J. Algebra* **9** (1967), 126–139. [150, 486]
- BOWTELL, A. J. and COHN, P. M.
71. Bounded and invariant elements in 2-firs. *Proc. Camb. Phil. Soc.* **69** (1971), 1–12. [379]
- BRAY, U. and WHAPLES, G.
83. Polynomials with coefficients from a division ring. *Canad. J. Math.* **35** (1983), 509–515. [539]
- BRENNER, J. L.
55. Quelques groupes libres de matrices. *C. R. Acad. Sci. Paris* **241** (1955), 1689–1691. [120]
- BRUNGS, H.-H.
69. Ringe mit eindeutiger Faktorzerlegung. *J. reine angew. Math.* **236** (1969), 43–66. [183, 192]
- 69'. Generalized discrete valuation rings. *Canad. J. Math.* **21** (1969), 1404–1408. [536, 538f.]
71. Overrings of principal ideal domains. *Proc. Amer. Math. Soc.* **28** (1971), 44–46.
73. Left Euclidean rings. *Pacif. J. Math.* **45** (1973), 27–33.
74. Right invariant right hereditary rings. *Canad. J. Math.* **26** (1974), 1186–1191.
76. Rings with a distributive lattice of right ideals. *J. Algebra* **40** (1976), 392–400.
78. Unique factorization in rings with right ACC_1 . *Glasgow Math. J.* **19** (1978), 167–171. [192]
- BRUNGS, H.-H. and TÖRNER, G.
81. Right chain groups and the generalized semigroup of divisibility. *Pacif. J. Math.* **97** (1981), 293–305. [204]
- BURMISTROVICH, I. E.
63. On the embedding of rings in skew fields (Russian). *Sibirsk. Mat. Zh.* **4** (1963), 1235–1240. [486]
- CAMILLO, V. P.
75. Distributive modules. *J. Algebra* **36** (1975), 16–25. [199, 226]
- CARCANAGUE, J.
71. Idéaux bilatères d'un anneau de polynômes non commutatifs sur un corps. *J. Algebra* **18** (1971), 1–18. [379]
- CARTAN, H. and EILENBERG, S.
56. "Homological Algebra". Princeton Univ. Press, Princeton, 1956. [63, 110]
- CHALABI, A. and KHURI, S. M.
80. Pure submodules of injective modules. *Rend. Ist. Mat. Univ. Trieste* **12** (1980), 15–19. [50]

CHASE, S. U.

61. A generalization of the ring of triangular matrices. *Nagoya Math. J.* **18** (1961), 13–25. [86]

62. On direct sums and products of modules. *Pacif. J. Math.* **12** (1962), 847–854. [175, 192]

CHEVALLEY, C.

[380]

CHOO, K. G.

74. Whitehead groups of twisted free associative algebras. *Pacif. J. Math.* **50** (1974), 399–402.

77. Grothendieck groups of twisted free associative algebras. *Glasgow Math. J.* **18** (1977), 193–196.

CLARK, W. E.

[336]

CLAUS, H. J.

55. Über die Partialbruchzerlegung in nicht notwendig kommutativen euklidischen Ringen. *J. reine angew. Math.* **194** (1955), 88–100.

COHN, P. M. (see also BERGMAN, G. M., BOWTELL, A. J.)

59. On the free product of associative rings. *Math. Zeits.* **71** (1959), 380–398.

60. On the free product of associative rings II. The case of skew fields. *Math. Zeits.* **73** (1960), 433–456. [151]

61. On a generalization of the Euclidean algorithm. *Proc. Camb. Phil. Soc.* **57** (1961), 18–30. [142, 151, 329]

- 61'. On the embedding of rings in skew fields. *Proc. London Math. Soc.* (3) **11** (1961), 511–530. [486f.]

- 61''. Quadratic extensions of skew fields. *Proc. London Math. Soc.* (3) **11** (1961), 531–556. [539]

62. Factorization in non-commutative power series rings. *Proc. Camb. Phil. Soc.* **58** (1962), 452–464. [152, 192, 378]

- 62'. On subsemigroups of free semigroups. *Proc. Amer. Math. Soc.* **13** (1962), 347–351. [379]

- 62''. "Free Rings". Yale University lecture notes, Yale University, New Haven, 1962. [380]

63. Noncommutative unique factorization domains. *Trans. Amer. Math. Soc.* **109** (1963), 313–331. Corr. *ibid.* **119** (1965), 552. [63, 86, 192]

- 63'. Rings with a weak algorithm. *Trans. Amer. Math. Soc.* **109** (1963), 332–356. [151f., 379]

64. Subalgebras of free associative algebras. *Proc. London Math. Soc.* (3) **14** (1964), 618–632. [379]

- 64'. Free ideal rings. *J. Algebra* **1** (1964), 47–69. [85]

66. Some remarks on the invariant basis property. *Topology* **5** (1966), 215–228. [7, 63, 85f., 150ff.]

- 66'. On the structure of the GL_2 of a ring. *Publ. Math. IHES* No. 30, Paris, 1966, 5–53. [86, 116, 121, 124, 152, 539]

- 66''. A remark on matrix rings over free ideal rings. *Proc. Camb. Phil. Soc.* **62** (1966), 1–4. [62f., 86]

- 66'''. Hereditary local rings. *Nagoya Math. J.* **27** (1) (1966), 223–230. [177, 192]

67. Torsion modules over free ideal rings. *Proc. London Math. Soc.* (3) **17** (1967), 577–599. [86, 297, 538]

68. Bezout rings and their subrings. *Proc. Camb. Phil. Soc.* **64** (1968), 251–264. [86]

69. Free associative algebras. *Bull. London Math. Soc.* **1** (1969), 1–39. [63, 86, 192, 225]
- 69'. Rings with a transfinite weak algorithm. *Bull. London Math. Soc.* **1** (1969), 55–59. [152]
- 69". Dependence in rings II. The dependence number. *Trans. Amer. Math. Soc.* **135** (1969), 267–279. [85, 151f.]
70. Factorization in general rings and strictly cyclic modules. *J. reine angew. Math.* **239/40** (1970), 185–200. [63, 183, 192, 539]
- 70'. Torsion and protorsion modules over free ideal rings. *J. Austral. Math. Soc.* **11** (1970), 490–498. [239]
- 70". On a class of rings with inverse weak algorithm. *Math. Zeits.* **117** (1970), 1–6. [152, 297]
71. The embedding of firs in skew fields. *Proc. London Math. Soc.* (3) **23** (1971), 193–213. [152, 192, 421, 452, 487]
- 71'. Rings of fractions. *Amer. Math. Monthly* **78** (1971), 596–615. [414]
- 71". The embedding of radical rings in simple radical rings. *Bull. London Math. Soc.* **3** (1971), 185–188. *Corr. ibid.* **4** (1972), 54; **5** (1973), 322.
72. Universal skew fields of fractions. *Symposia Math.* **VIII** (1972), 135–148. [487]
- 72'. Skew fields of fractions and the prime spectrum of a general ring. Lectures on rings and modules. *Springer Lecture Notes in Math.* No. 246, Springer-Verlag, Berlin, 1972, 1–71. [487]
- 72". "Rings of fractions". University of Alberta Lecture Notes, Edmonton, 1972. [152, 410, 412]
- 72". Generalized rational identities. *Proc. Conf. on Ring Theory at Park City, Utah, 1971* (ed. R. Gordon), Academic Press, New York, 1972, 107–115. [297]
73. The similarity reduction of matrices over a skew field. *Math. Zeits.* **132** (1973), 151–163. [421, 505, 539]
- 73'. Free radical rings. *Colloq. Math. Soc. J. Bolyai*, 6. Rings, modules and radicals. Keszthely, Hungary, 1971 (1973), 135–145. [152]
- 73". Unique factorization domains. *Amer. Math. Monthly* **80** (1973), 1–17. [183, 192]
- 73". Bound modules over hereditary rings. "Selected questions on Algebra and Logic" (volume dedicated to the memory of A. I. Malcev), *Izd. "Nauka"*, SO, Novosibirsk, 1973, pp. 131–141. [296]
74. Progress in free associative algebras. *Isr. J. Math.* **19** (1974), 109–151. [63]
- 74'. Localization in semifirs. *Bull. London Math. Soc.* **6** (1974), 13–20. [297, 488]
- 74". The class of rings embeddable in skew fields. *Bull. London Math. Soc.* **6** (1974), 147–148. [486]
76. "Morita equivalence and duality". Queen Mary College Lecture Notes, London, 1966, 1976. [5, 62]
- 76'. The Cayley–Hamilton theorem in skew fields. *Houston J. Math.* **2** (1976), 49–55. [281, 297, 539]
77. "Algebra II". J. Wiley & Sons, Chichester, 1977. [6, 20f., 23, 36, 50, 126, 139, 187, 241, 284, 303, 317, 330, 352, 364, 442, 510, 520, 544]
- 77'. "Skew field constructions". London Math. Soc. Lecture Notes No. 27, Cambridge University Press, Cambridge, 1977. [110, 297, 421, 487, 505, 510, 514]
- 77". A construction of simple principal right ideal domains. *Proc. Amer. Math. Soc.* **66** (1977), 217–222. *Corr. ibid.* **77** (1979), 40. [539]
- 77". Full modules over semifirs. *Publ. Math. Debrecen* **24** (1977), 305–310. [296f.]
79. The affine scheme of a general ring. Applications of sheaves: Durham Research

- Symposium 1977. Springer Lecture Notes in Math. No. 753. Springer-Verlag, Berlin (1979), 197–211. [152, 412, 487f.]
80. On semifir constructions. Word problems II, the Oxford Book (ed. S. I. Adian, W. W. Boone and G. Higman). North-Holland, Amsterdam, 1980, pp. 73–80.
81. "Universal Algebra" (2nd ed.). D. Reidel, Dordrecht, 1981. [7, 12, 35, 59, 82, 148, 286, 414, 419, 486, 539f.]
82. The universal field of fractions of a semifir I. Numerators and denominators. *Proc. London Math. Soc.* (3) **44** (1982), 1–32. [63, 85, 192, 293, 297, 488]
- 82'. The divisor group of a fir. *Publ. Sec. Mat. Univ. Aut. Barcelona* **26** (1982), 131–163. [451, 488]
- 82''. Ringe mit distributivem Faktorverband. *Abh. Braunschweig. Wiss. Ges.* **33** (1982), 35–40. [226]
- 82'''. Torsion modules and the factorization of matrices. Advances in non-comm. ring theory. *Proc. 12th George H. Hudson Symposium* (ed. P. J. Fleury). Springer Lecture Notes in Math. No. 951. Springer-Verlag, Berlin (1982), 1–11. [297]
83. Determinants on free fields. *Contemp. Math.* **13** (1983), 99–108. [488]
84. Fractions. *Bull. London Math. Soc.* **16** (1984), 561–574. [297, 487, 521]
- 84'. Embedding problems for rings and semigroups. Proc. "25. Arbeitstagung über allgemeine Algebra", Darmstadt 1983 (ed. P. Burmeister *et al.*). N. Heldermann, Berlin, 1984, 115–126. [148]
85. The universal field of fractions of a semifir II. The depth. *Proc. London Math. Soc.* (3) **50** (1985), 69–94. [192, 437, 488]
- 85'. The universal field of fractions of a semifir III. Centralizers and normalizers. *Proc. London Math. Soc.* (3) **50** (1985), 95–113. [226, 488]
- a. Right principal Bezout domains (to appear). [492]
- b. On coproducts of ordered systems (to appear). [521]
- COHN, P. M. and DICKS, W.
76. Localization in semifirs II. *J. London Math. Soc.* (2) **13** (1976), 411–418. [152, 488]
- COHN, P. M. and MAHDAVI-HEZAVEHI, M.
80. Extensions of valuations on skew fields. *Proc. Ring Theory Week*, Antwerp, 1980 (ed. F. Van Oystaeyen). Springer Lecture Notes in Math. No. 825. Springer-Verlag, Berlin, 1980, 28–41.
- COHN, P. M. and SCHOFIELD, A. H.
82. On the law of nullity. *Math. Proc. Camb. Phil. Soc.* **91** (1982), 357–374. [63, 265, 297, 488]
85. Two examples of principal ideal domains, *Bull. London Math. Soc.* **17** (1985), 25–28. [488]
- CONNELL, E. H. (see BASS, H.)
- CORNER, A. L. S.
69. Additive categories and a theorem of W. G. Leavitt. *Bull. Amer. Math. Soc.* **75** (1969), 78–82. [150]
- COXETER, H. S. M. [380]
- COZZENS, J. H.
72. Simple principal left ideal domains. *J. Algebra* **23** (1972), 66–75.

COZZENS, J. H. and FAITH, C.

75. "Simple Noetherian rings". Cambridge Tracts in Math. No. 69, Cambridge University Press, Cambridge, 1975. [192]

CURTIS, C. W.

52. A note on non-commutative polynomial rings. *Proc. Amer. Math. Soc.* **3** (1952), 965–969. [63]

CZERNIAKIEWICZ, A. J.

71. Automorphisms of a free associative algebra of rank 2. *Trans. Amer. Math. Soc.* I. **160** (1971), 393–401. II. **171** (1972), 309–315. [355, 379]

DAUNS, J.

70. Embeddings in division rings. *Trans. Amer. Math. Soc.* **150** (1970), 287–299.
82. "A concrete approach to division rings". Heldermann-Verlag, Berlin, 1982.

DEDEKIND, R.

1900. Über die von drei Moduln erzeugte Dualgruppe. *Math. Ann.* **53** (1900), 371–403. [*Ges. Werke II*, 236–271.] [540]

DICKER, R. M.

68. A set of independent axioms for a field and a condition for a group to be the multiplicative group of a field. *Proc. London Math. Soc.* (3) **18** (1968), 114–124. [487]

DICKS, W. (see also ALMKVIST, G., BERGMAN, G. M., COHN, P. M.)

- [151, 336, 379f., 444, 487f.]
72. On one-relator associative algebras. *J. London Math. Soc.* (2) **5** (1972), 249–252.
74. "Idealizers in free algebras". Thesis, London University, 1974. [151f., 341]
77. Meyer–Vietoris presentations over colimits of rings. *Proc. London Math. Soc.* (3) **34** (1977), 557–576. [62]
79. Hereditary group rings. *J. London Math. Soc.* (2) **20** (1979), 27–38.
81. An exact sequence for rings of polynomials in partly commuting indeterminates. *J. Pure Appl. Algebra* **22** (1981), 215–228.
82. A commutator test for two elements to generate the free algebra of rank two. *Bull. London Math. Soc.* **14** (1982), 48–51. [355, 380]
83. The HNN construction for rings. *J. Algebra* **81** (1983), 434–487. [152]
83'. Automorphisms of the polynomial ring in two variables. *Publ. Sec. Mat. Univ. Aut. de Barcelona* **27** (1983), 155–162. [379]
83''. A free algebra can be free as a module over a non-free subalgebra. *Bull. London Math. Soc.* **15** (1983), 373–377.
85. Homogeneous elements of free algebras have free idealizers, *Math. Proc. Camb. Phil. Soc.* **97** (1985), 7–26.

DICKS, W. and FORMANEK, E.

82. Poincaré series and a problem of S. Montgomery. *Linear and Multilin. Algebra* **12** (1982), 21–30. [370, 380]

DICKS, W. and LEWIN, J.

82. A Jacobian conjecture for free associative algebras. *Comm. in Algebra* **10** (12) (1982), 1285–1306. [362]

DICKS, W. and MENAL, P.

79. The group rings that are semifirs. *J. London Math. Soc.* (2) **19** (1979), 288–290.

DICKS, W. and SONTAG, E. D.

78. Sylvester domains. *J. Pure Appl. Algebra* **13** (1978), 243–275. [253, 256, 261, 297]

DIEUDONNÉ, J.

43. Les déterminants sur un corps non-commutatif. *Bull. Soc. Math. France* **71** (1943), 27–45. [448]

73. "Introduction to the Theory of Formal Groups". M. Dekker, New York, 1973. [192]

DIXMIER, J.

68. Sur les algèbres de Weyl. *Bull. Soc. Math. France* **96** (1968), 209–242. [357]

DLAB, V. and RINGEL, C. M.

85. A class of bounded hereditary Noetherian domains. *J. Algebra* **92** (1985), 311–321. [488]

DONKIN, S.

[380]

DOSS, R.

48. Sur l'immersion d'un semi-groupe dans un groupe. *Bull. Sci. Math.* **72** (1948), 139–150. [63]

DRESS, F.

71. Stathmes euclidiens et séries formelles. *Acta Arith.* **19** (1971), 261–265. [94]

DROGOMIZHSKA, M. M.

70. Augmentation of a matrix to an invertible one in an Ore ring (Ukrainian, Russian summary). *Visnik L'viv. Politekhn. Inst.* No. 44 (1970), 50–60, 210. [63]

DUBOIS, D. W.

66. Modules of sequences of elements of a ring. *J. London Math. Soc.* **41** (1966), 177–180. [297]

DUBREIL-JACOTIN, M.-L.

47. Sur l'immersion d'un semigroupe dans un groupe. *C. R. Acad. Sci. Paris* **225** (1947), 787–788. [379]

DUBROVIN, N. I.

82. Non-commutative valuation rings (Russian). *Trudy Mat. Obshch.* **45** (1982), 265–280.

EILENBERG, S. (see also CARTAN, H.)

[58]

ELIZAROV, V. P.

69. Rings of fractions (Russian). *Algebra i Logika* **8** (4) (1969), 381–424. [63]

EUCLID

- 300. Elements. [150]

FAITH, C. (see also COZZENS, J. H.)

73. "Algebra I, Rings, Modules and Categories". Springer-Verlag, Berlin, 1973. [191]

FAIZOV, S. K.

81. Free ideal categories (Russian). *Ukrainskii Mat. Zh.* **33** (1981), 626–630. [86]

FELLER, E. H.

60. Intersection irreducible ideals of a non-commutative principal ideal domain. *Canad. J. Math.* **12** (1960), 592–596. [192]

61. Factorization and lattice theorems for a bounded principal ideal domain. *Duke Math. J.* **28** (1961), 579–583.

FISHER, J. L.

71. Embedding free algebras in skew fields. *Proc. Amer. Math. Soc.* **30** (1971), 453–458. [394]
 74. The poset of skew fields generated by a free algebra. *Proc. Amer. Math. Soc.* **42** (1974), 33–35.
 74'. The category of epic R -fields. *J. Algebra* **28** (1974), 283–290.

FITTING, H.

35. Primärkomponentenzerlegung in nichtkommutativen Ringen. *Math. Ann.* **111** (1935), 19–41. [63]
 36. Über den Zusammenhang zwischen dem Begriff der Gleichartigkeit zweier Ideale und dem Äquivalenzbegriff der Elementarteilertheorie. *Math. Ann.* **112** (1936), 572–582. [25, 63, 163, 192]

FLIESS, M.

70. Inertie et rigidité des séries rationnelles et algébriques. *C. R. Acad. Sci. Paris Ser. A* **270** (1970), 221–223. [152]
 70'. Transductions algébriques. *R.I.R.O. R-1* (1970), 109–125.
 70''. Sur le plongement de l'algèbre des séries rationnelles non commutatives dans un corps gauche. *C. R. Acad. Sci. Paris Ser. A* **271** (1970), 926–927.
 71. Deux applications de la représentation matricielle d'une série rationnelle non-commutative. *J. Algebra* **19** (1971), 344–353.

FORMANEK, E. (see ALMKVIST, G., DICKS, W.)

FOX, R. H.

53. Free differential calculus I. Derivation in the free group ring. *Ann. Math.* **57** (1953), 547–560. [55]

GABEL, M. R.

[19]

GENTILE, E. R.

60. On rings with a one-sided field of quotients. *Proc. Amer. Math. Soc.* **11** (1960), 380–384. [50, 63]

GERASIMOV, V. N.

[486]

73. Rings that are nearly free (Russian). *Rings II* (Russian), pp. 9–19. Inst. Mat. Sibirsk. Otdel. Akad. Nauk SSSR, Novosibirsk, 1973.
 76. Distributive lattices of subspaces and the word problem for one-relator algebras (Russian). *Algebra i Logika* **15** (4) (1976), 384–435, 487.
 79. Inverting homomorphisms of rings (Russian). *Algebra i Logika* **18** (1979), 648–663. [411]
 82. Localizations in associative rings (Russian). *Sibirsk. Mat. Zh.* **23** (1982), 36–54. [488]

GERSTEN, S. M.

65. Whitehead groups of free associative algebras. *Bull. Amer. Math. Soc.* **71** (1965), 157–159.
 74. K -theory of free rings. *Comm. in Algebra* **1** (1974), 39–64.

GOLDIE, A. W.

58. The structure of prime rings under ascending chain conditions. *Proc. London Math. Soc.* (3) **8** (1958), 589–608. [63]

- GOLDSCHMIDT, D. [492]
- GOODEARL, K. R.
79. "von Neumann Regular Rings". Pitman, London, 1979. [63]
- GOODEARL, K. R. and WARFIELD, JR, R. B.
81. State spaces of K_0 of Noetherian rings. *J. Algebra* **71** (1981), 322–378. [63]
- GORBACHUK, O. L. and KOMARNICKII, M. YA.
77. Radical filters in a principal ideal domain (Russian, English summary). *Dokl. Akad. Nauk Ukrain. SSR Ser. A*, 1977, 103–104.
- GORDON, B. and MOTZKIN, T. S.
65. On the zeros of polynomials over division rings. *Trans. Amer. Math. Soc.* **116** (1965), 218–226. Corr. *ibid.* **122** (1966), 547. [539]
- GOURSAUD, J. M. and VALETTE, J.
75. Anneaux de groupe héréditaires et semi-héréditaires. *J. Algebra* **34** (1975), 205–212. [311]
- GOVOROV, V. E.
72. Graded algebras (Russian). *Mat. Zametki* **12** (1972), 197–204.
73. The global dimension of algebras (Russian). *Mat. Zametki* **14** (1973), 399–406.
81. Algebras of homological dimension 1 (Russian). *Mat. Sb. (N.S.)* **116** (158) (1981), 111–119.
- GRAMS, A.
74. Atomic rings and the ascending chain conditions for principal ideals. *Proc. Camb. Phil. Soc.* **75** (1974), 321–329. [154]
- GRÄTZER, G.
78. "General Lattice Theory". Birkhäuser, Basel, Stuttgart, 1978. [213]
- GUAZZONE, S.
62. Sui Λ -moduli liberi i alcuni teoremi di C. J. Everett. *Rend. Sem. Mat. Univ. Padova* **32** (1962), 304–312. [63]
- HAHN, H.
07. Über die nichtarchimedischen Grössensysteme. *S.-B. Akad. Wiss. Wien IIa* **116** (1907), 601–655. [539]
- HALL, M.
59. "The Theory of Groups". Macmillan, New York, 1959. [528]
- HASSE, H.
28. Über die eindeutige Zerlegung in Primelemente oder in Primhauptideale in Integritätsbereichen. *J. reine angew. Math.* **159** (1928), 3–12. [94, 100]
- HAUSKNECHT, A. O. [492]
- HELM, P. R.
83. Generators and relations for certain linear groups over rings of linear operators. *Comm. in Algebra* **11** (5) (1983), 551–565. [120]
- HENSEL, K.
27. Über eindeutige Zerlegung in Primelemente. *J. reine angew. Math.* **158** (1927), 195–198.
29. Über Systeme in einfachen Körpern. *J. reine angew. Math.* **160** (1929), 131–142.

HERSTEIN, I. N.

56. Conjugates in division rings. *Proc. Amer. Math. Soc.* **7** (1956), 1021–1022. [539]
 76. "Rings with involution". Chicago Lectures in Math., Chicago Univ. Press, Chicago and London, 1976. [282]

HIBLOT, J.-J.

75. Des anneaux euclidiens dont le plus petit algorithme n'est pas à valeurs finies. *C. R. Acad. Sci. Paris* **281** (1975), 411–414. [151]
 76. Sur les anneaux euclidiens. *Bull. Soc. Math. France* **104** (1976), 33–50.

HIGMAN, G.

40. The units of group rings. *Proc. London Math. Soc.* (2) **46** (1940), 231–248. [272]
 52. Ordering by divisibility in abstract algebras. *Proc. London Math. Soc.* (3) **2** (1952), 326–336. [539]

HIGMAN, G., NEUMANN, B. H. and NEUMANN, H.

49. Embedding theorems for groups. *J. London Math. Soc.* **24** (1949), 247–254. [151]

HILBERT, D.

1899. Grundlagen der Geometrie. Festschrift zur Feier der Enthüllung des Gauss-Weber-Denkmal in Göttingen. B. G. Teubner, Leipzig, 1899. [539]

HOCHSTER, M.

69. Prime ideal structure in commutative rings. *Trans. Amer. Math. Soc.* **142** (1969), 43–60. [410, 412]

HOTJE, H.

83. On Cohnfunctions. *Annals of Discrete Math.* **18** (1983), 467–468. [487]

HUDRY, A.

70. Quelques applications de la localisation au sens de Gabriel. *C. R. Acad. Sci. Paris Ser. A–B* **270** (1970), A8–A10. [488]

IKEDA, M.

69. Über die maximalen Ideale einer freien assoziativen Algebra. *Hamb. Abh.* **33** (1969), 59–66. [225]

INCE, E. L.

[58]

ISCHEBECK, F.

71. Die Zentren hereditärer Ringe. *Math. Ann.* **193** (1971), 83–88. [379]

JACOBSON, N.

[10]

34. A note on non-commutative polynomials. *Ann. Math.* **35** (1934), 209–210. [151]
 37. Pseudo-linear transformations. *Ann. Math.* **38** (1937), 484–507. [506, 539]
 43. "Theory of Rings". Amer. Math. Soc., New York, 1943. [192, 320, 379, 539]
 64. "Structure of Rings" (rev. ed.). Amer. Math. Soc., Providence, 1964. [380, 539]

JAFFARD, P.

60. "Les Systèmes d'Idéaux". Dunod, Paris, 1960. [86]

JAMES, G. D. and KERBER, A.

81. The representation theory of the symmetric group. "Encyclopedia of Mathematics and its Applications", vol. 16. Addison-Wesley, Reading, Mass., 1981. [364]

JATEGAONKAR, A. V.

69. A counter-example in ring theory and homological algebra. *J. Algebra* **12** (1969), 418–440. [532, 538f.]

- 69'. Rings with a transfinite left division algorithm. *Bull. Amer. Math. Soc.* **75** (1969), 559–561.
- 69". Ore domains and free algebras. *Bull. London Math. Soc.* **1** (1969), 45–46. [62]
71. Skew polynomial rings over semisimple rings. *J. Algebra* **19** (1971), 315–328.
- JENSEN, C. U.
63. On characterizations of Prüfer rings. *Math. Scand.* **13** (1963), 90–98. [193, 198]
66. A remark on semihereditary local rings. *J. London Math. Soc.* **41** (1966), 479–482.
69. Some cardinality questions for flat modules and coherence. *J. Algebra* **12** (1969), 231–241. [86]
- JOHNSON, R. E. (see also BEAUREGARD, R. A.)
63. Principal right ideal rings. *Canad. J. Math.* **15** (1963), 297–301.
65. Unique factorization in a principal right ideal domain. *Proc. Amer. Math. Soc.* **16** (1965), 526–528. [192]
67. The quotient domain of a weak Bezout domain. *J. Math. Sci.* **2** (1967), 21–22. [83]
- JØNDRUP, S.
70. On finitely generated flat modules. *Math. Scand.* **26** (1970), 233–240.
77. The centre of a right hereditary ring. *J. London Math. Soc.* (2), **15** (1977), 211–212.
- DE JONQUIÈRES, J. P. E. DE F. [379]
- JOOSTE, T. DE W.
71. "Derivations in free power series rings and free associative algebras". Thesis, London University, 1971. [137f.]
78. Primitive derivations on free associative algebras. *Math. Zeits.* **164** (1978), 15–23.
- JORDAN, D. A. [58]
- JUNG, H. W. E.
42. Über ganze birationale Transformationen der Ebene. *J. reine angew. Math.* **184** (1942), 161–174. [348, 379]
- KAPLANSKY, I.
49. Elementary divisors and modules. *Trans. Amer. Math. Soc.* **66** (1949), 464–491. [492, 539]
58. Projective modules. *Ann. Math.* **68** (1958), 372–377. [12, 14, 23, 63]
70. "Commutative Rings". Allyn & Bacon, Boston, 1970. [310]
- KAZIMIRSKII, P. S.
62. A theorem on elementary divisors for the ring of differential operators (Ukrainian, Russian summary). *Dopovidi Akad. Nauk Ukrain. RSR* (1962), 1275–1278.
- KAZIMIRSKII, P. S. and LUNIK, F. P.
72. Completion of a rectangular matrix over an associative ring to an invertible one (Ukrainian, English summary). *Dopovidi Akad. Nauk Ukrain. RSR, Ser. A* (1972), 505–506. [63]
- KERBER, A. (see JAMES, G. D.)
- KERR, J. W.
82. The power series ring over an Ore domain need not be Ore. *J. Algebra* **75** (1982), 175–177. [63]

- KHARCHENKO, V. K. [372, 376]
77. Galois theory of semiprime rings (Russian). *Algebra i Logika* 16 (1977), 313–363. [380]
78. On algebras of invariants of free algebras (Russian). *Algebra i Logika* 17 (1978), 478–487. [363, 380]
84. Noncommutative invariants of finite groups and Noetherian varieties. *J. Pure Appl. Algebra* 31 (1984), 83–90. [380]
- KHURI, S. M. (see CHALABI, A.)
- KIREZCI, M. [10, 150]
- KLEIN, A. A.
67. Rings nonembeddable in fields with multiplicative semigroups embeddable in groups. *J. Algebra* 7 (1967), 100–125. [150, 486]
69. Necessary conditions for embedding rings into fields. *Trans. Amer. Math. Soc.* 137 (1969), 141–151. [82f, 86]
70. Three sets of conditions on rings. *Proc. Amer. Math. Soc.* 25 (1970), 393–398.
72. A remark concerning embeddability of rings into fields. *J. Algebra* 21 (1972), 271–274. [251]
72'. Involution division rings with arbitrary centers. *Proc. Amer. Math. Soc.* 34 (1972), 38–42. [445]
80. Some ring-theoretic properties implied by embeddability in fields. *J. Algebra* 66 (1980), 147–155.
- KNIGHT, J. T.
70. On epimorphisms of non-commutative rings. *Proc. Camb. Phil. Soc.* 68 (1970), 589–600.
- KNUS, M. A.
68. Homology and homomorphisms of rings. *J. Algebra* 9 (1968), 274–284. [379]
- KOLOTOV, A. T.
78. On free subalgebras of free associative algebras (Russian). *Sibirsk. Mat. Zh.* 19 (1978), 328–335. [336, 341, 378]
- KOMARNICKII, M. YA. (see GORBACHUK, O. L.)
- KOROTKOV, M. V.
76. Description of rings with single-valued division with remainder (Russian). *Uspekhi Mat. Nauk* 31 (1976), No. 1 (187), 253–254. [539]
- KOSHEVOI, E. G.
66. On the multiplicative semigroup of a class of rings without zero-divisors (Russian). *Algebra i Logika* 5 (5) (1966), 49–54. [177, 192]
70. On certain associative algebras with transcendental relations (Russian). *Algebra i Logika* 9 (5) (1970), 520–529. [62]
71. Pure subalgebras of free associative algebras (Russian). *Algebra i Logika* 10 (1971), 183–187. [341]
- KOZHUKHOV, I. B.
82. Free left ideal semigroup rings (Russian). *Algebra i Logika* 21 (1) (1982), 37–59. [152]
- KRAFT, L. G.
49. "A device for quantizing, grouping and coding amplitude modulated pulses".

M.S. Thesis, Electrical Engineering Dept., M.I.T. Cambridge, Mass., 1949.

[379]

KRULL, W.

28. Zur Theorie der zweiseitigen Ideale in nichtkommutativen Bereichen. *Math. Zeits.* **28** (1928), 481–503.

54. Zur Theorie der kommutativen Integritätsbereiche. *J. reine angew. Math.* **192** (1954), 230–252.

KULK, W. VAN DER

53. On polynomial rings in two variables. *Nieuw Archief voor Wisk.* (3) **I** (1953), 33–41. [348, 379]

LALLEMENT, G.

79. "Semigroups and Combinatorial Applications". J. Wiley & Sons, New York, Chichester, 1979.

LAM, T. Y.

76. Series summation of stably free modules. *Quart. J. Math. Oxford Ser. (2)* **27** (1976), 37–46. [19]

78. "Serre's conjecture". Lecture Notes in Math., No. 635. Springer-Verlag, Berlin, 1978. [19, 63]

LANDAU, E.

02. Ein Satz über die Zerlegung homogener linearer Differentialausdrücke in irreduzible Faktoren. *J. reine angew. Math.* **124** (1902), 115–120. [191]

LANE, D. R.

[151, 336, 371]

75. Fixed points of affine Cremona transformations of the plane over an algebraically closed field. *Amer. J. Math.* **97** (1975), 707–732.

76. "Free algebras of rank two and their automorphisms". Thesis, London University, 1976. [357, 362f., 380]

LEAVITT, W. G.

57. Rings without invariant basis number. *Proc. Amer. Math. Soc.* **8** (1957), 322–328. [62]

60. The module type of a ring. *Trans. Amer. Math. Soc.* **103** (1960), 113–130.

LEISSNER, W.

71. Eine Charakterisierung der multiplikativen Gruppe eines Körpers. *Jber. Deutsch. Math.-Verein.* **73** (1971), 92–100.

LEMMLEIN, V.

54. On Euclidean rings and principal ideal rings (Russian). *Dokl. Akad. Nauk SSSR* **97** (1954), 585–587. [94]

LENSTRA JR, H. W.

74. "Lectures on Euclidean Rings". Universität, Bielefeld, 1974. [94, 151, 532, 539]

LESIEUR, L.

[58]

LEUTBECHER, A.

78. Euklidischer Algorithmus und die Gruppe GL_2 . *Math. Ann.* **231** (1978), 269–285. [151]

LEVI, F. W.

44. On semigroups. *Bull. Calcutta Math. Soc.* **36** (1944), 141–146. [379]

LEVI-CIVITA, T.

1892. Sugli infiniti ed infinitesimali attuali quali elementi analitici. *Atti Ist. Veneto di Sci. Lett. ed Arti, Ser. 7a* 4 (1892-93), 1765-1816 (= Op. mat. vol. primo, pp. 1-39, Bologna, 1954). [539]

LEVITZKI, J.

[63]

LEVY, L. S. and ROBSON, J. C.

74. Matrices and pairs of modules. *J. Algebra* 29 (1974), 427-454.

LEWIN, J. (see also BEDOYA, H., BERGMAN, G. M., DICKS, W.)

[109]

68. On Schreier varieties of linear algebras. *Trans. Amer. Math. Soc.* 132 (1968), 553-562. [379]

69. Free modules over free algebras and free group algebras: the Schreier technique. *Trans. Amer. Math. Soc.* 145 (1969), 455-465. [110, 151, 297]

72. A note on zero-divisors in group rings. *Proc. Amer. Math. Soc.* 31 (1972), 357-359.

73. On some infinitely presented associative algebras. *J. Austral. Math. Soc.* 16 (1973), 290-293.

74. A matrix representation for associative algebras. *Trans. Amer. Math. Soc.* 188 (1974). I. 293-308. II. 309-317. [110]

- 74'. Fields of fractions for group algebras of free groups. *Trans. Amer. Math. Soc.* 192 (1974), 339-346. [487]

77. Ranks of matrices over Ore domains. *Proc. Amer. Math. Soc.* 62 (1977), 233-236.

LEWIN, J. and LEWIN, T.

68. On ideals of free associative algebras generated by a single element. *J. Algebra* 8 (1968), 248-255.

78. An embedding of the group algebra of a torsion free one-relator group in a field. *J. Algebra* 52 (1978), 39-74.

LEWIN, T. (see LEWIN, J.)

LISSNER, D.

65. Outer product rings. *Trans. Amer. Math. Soc.* 116 (1965), 526-535. [63]

LOEWY, A.

03. Über reduzible lineare homogene Differentialausdrücke. *Math. Ann.* 56 (1903), 549-584. [191]

20. Begleitmatrizen und lineare homogene Differentialausdrücke. *Math. Zeits.* 7 (1920), 58-128.

LOPATINSKII, Y. B.

45. Linear differential operators (Russian). *Mat. Sb.* 17 (59) (1945), 267-288.

46. A theorem on bases (Russian). *Trudy Sek. Mat. Akad. Nauk Azerb. SSR* 2 (1946), 32-34.

47. On some properties of rings of linear differential operators (Russian). *Nauch. Zap. Lvov Univ. I Ser. fiz.-mat.* 2 (1947), 101-107.

LOTHAIRE, M.

83. Combinatorics on words. *Encyclopedia of Mathematics and its Applications*, vol. 17. Addison-Wesley, Reading, Mass., 1983. [324, 329, 379]

LUNIK, F. P. (see also KAZIMIRSKII, P. S.)

70. The intersection of principal ideals in a certain non-commutative ring (Ukrainian, Russian summary). *Visnik Lviv Politekhn. Inst.* No. 44 (1970), 24-26, 209. [86]

LYNDON, R. C. and SCHUPP, P. E.

77. "Combinatorial Group Theory". Springer-Verlag, Berlin-Heidelberg-New York, 1977. [152]

MCADAM, S. and RUSH, D. E.

78. Schreier rings. *Bull. London Math. Soc.* **10** (1978), 77-80. [252]

MACDONALD, I. G.

68. "Algebraic Geometry: Introduction to Schemes". Benjamin, New York, 1968. [412, 487]

MACINTYRE, A.

79. Combinatorial problems for skew fields. I. Analogue of Britton's lemma, and results of Adjan-Rabin type. *Proc. London Math. Soc.* (3) **39** (1979), 211-236. [152]

MAC LANE, S.

63. "Homology". Springer-Verlag, Berlin, 1963. [79, 544]
71. "Categories for the Working Mathematician". Springer-Verlag, Berlin, 1971. [544]

MCMILLAN, B.

56. Two inequalities implied by unique decipherability. *IRE Trans. Inform. Theory* **IT-2** (1956), 115-116. [379]

MAGNUS, W.

40. Über Gruppen und zugeordnete Liesche Ringe. *J. reine angew. Math.* **182** (1940), 142-149.

MAHDAVI-HEZAVEHI, M. (see also COHN, P. M.)

79. Matrix valuations on rings. *Antwerp Conf. on Ring Theory*. M. Dekker, New York, 1979, 691-703.
82. Matrix valuations and their associated skew fields. *Resultate d. Math.* **5** (1982), 149-156. [453, 488]

MAKAR-LIMANOV, L. G.

70. "On automorphisms of certain algebras" (Russian). Thesis, Moscow University, 1970. [45, 531]
70'. The automorphisms of the free algebra with two generators (Russian). *Funkcional. Anal. i Prilozhen.* **4** (1970), 107-108. [357, 379]
75. On algebras with one relation. *Uspekhi Mat. Nauk* **30**, No. 2 (182) (1975), 217. [355]
83. The skew field of fractions of the Weyl algebra contains a free non-commutative subalgebra. *Comm. in Algebra* **11** (17) (1983), 2003-2006.
85. Algebraically closed skew fields. *J. Algebra* **93** (1985), 117-135.

MALCEV, A. I.

37. On the immersion of an algebraic ring into a field. *Math. Ann.* **113** (1937), 686-691. [150, 486]
39. Über die Einbettung von assoziativen Systemen in Gruppen I (Russian, German summary), *Mat. Sb.* **6**(48) (1939), 331-336. [486]
40. Über die Einbettung von assoziativen Systemen in Gruppen II (Russian, German summary). *Mat. Sb.* **8**(50) (1940), 251-264. [486]
48. On the embedding of group algebras in division algebras (Russian). *Dokl. Akad. Nauk SSSR* **60** (1948), 1499-1501. [487, 539]
53. Nilpotent semigroups (Russian). *Uchen. Zap. Ivanovsk. Ped. In-ta* **4** (1953), 107-111. [487]

- MALCOLMSON, P. [486]
 78. A prime matrix ideal yields a skew field. *J. London Math. Soc.* (2) **18** (1978), 221–233.
80. On making rings weakly finite. *Proc. Amer. Math. Soc.* **80** (1980), 215–218. [63]
 80'. Determining homomorphisms into a skew field. *J. Algebra* **64** (1980), 399–413.
82. Construction of universal matrix localizations. Advances in non-comm. Ring theory. *Proc. 12th G. H. Hudson Symp.*, Plattsburgh, 1981. Lecture Notes in Math. No. 951, Springer-Verlag, Berlin, pp. 117–131.
84. Matrix localizations of firs. *Trans. Amer. Math. Soc.* **282** (1984), 503–518, 519–527. [485, 488]
- MARTINDALE III, W. S.
 82. The extended center of coproducts. *Canad. Math. Bull.* **25**(2) (1982), 245–248.
- MARTINDALE III, W. S. and MONTGOMERY, M. S.
 83. The normal closure of the coproducts of domains. *J. Algebra* **82** (1983), 1–17. [380]
- MENAL, P. (see also DICKS, W.)
 79. Remarks on the GL_2 of a ring. *J. Algebra* **61** (1979), 335–359. [152]
 81. The monoid rings that are Bezout domains. *Arch. Math.* **37** (1981), 43–47. [86]
- MILNOR, J. W.
 71. "Introduction to algebraic K-theory". Ann. Math. Studies No. 72, Princeton Univ. Press, Princeton, 1971. [19]
- MITCHELL, B.
 72. Rings with several objects. *Adv. Math.* **8** (1972), 1–161. [86]
- MOLIEN, T. [380]
- MONTGOMERY, M. S. (see also MARTINDALE III, W. S.)
 83. von Neumann finiteness of tensor products of algebras. *Comm. in Alg.* **11**(6) (1983), 595–610. [10, 150]
- MONTGOMERY, M. S. and PASSMAN, D. S.
 84. Galois theory of prime rings. *J. Pure Appl. Algebra* **31** (1984), 139–184. [372, 380]
- MORAN, S.
 73. Some subalgebra theorems. *J. Algebra* **27** (1973), 366–371. [379]
- MOTZKIN, T. (see also GORDON, B.)
 49. The euclidean algorithm. *Bull. Amer. Math. Soc.* **55** (1949), 1142–1146. [151]
- MOUFANG, R.
 37. Einige Untersuchungen über geordnete Schiefkörper. *J. reine angew. Math.* **176** (1937), 203–223. [487]
- NAGATA, M. [343]
 57. A remark on the unique factorization theorem. *J. Math. Soc. Japan* **9** (1957), 143–145. [157, 192]
 72. "On Automorphism Group of $k[x, y]$ ". Kinokuniya Bookstore & Co., Ltd., Tokyo, 1972. [352, 379]
- NAKAJIMA, H.
 83. Regular rings of invariants of unipotent groups. *J. Algebra* **85** (1983), 253–286. [380]

NAKAYAMA, T.

38. A note on the elementary divisor theory in non-commutative domains. *Bull. Amer. Math. Soc.* **44** (1938), 719–723. [539]

NEUMANN, B. H. (see also HIGMAN, G.)

49. On ordered division rings. *Trans. Amer. Math. Soc.* **66** (1949), 202–252.

[487, 539]

NEUMANN, H. (see HIGMAN, G.)

NIVAT, M.

68. Transduction des langages de Chomsky. *Ann. Inst. Fourier* **18**(1) (1968), 339–456. [151]

69. Séries rationnelles et algébriques en variables non commutatives. *Cours DEA* 1969/70. [487]

NÖBELING, G.

68. Verallgemeinerung eines Satzes von Herrn Specker. *Invent. Math.* **6** (1968), 41–55. [297]

NOETHER, E.

[380, 487]

NOETHER, E. and SCHMEIDLER, W.

20. Moduln in nichtkommutativen Bereichen, insbesondere aus Differential- und Differenzenausdrücken. *Math. Zeits.* **8** (1920), 1–35. [226]

NUÑEZ, P.

1567. "Libro de Algebra". Antwerp, 1567. [150]

OJANGUREN, M. and SRIDHARAN, R.

71. Cancellation of Azumaya algebras. *J. Algebra* **18** (1971), 501–505. [63]

ORE, O.

31. Linear equations in non-commutative fields. *Ann. Math.* **32** (1931), 463–477.

[63, 486]

32. Formale Theorie der linearen Differentialgleichungen. *J. reine angew. Math.* **167** (1932), 221–234. II. *ibid.* **168** (1932), 233–252. [58, 63, 192]

33. Theory of non-commutative polynomials. *Ann. Math.* **34** (1933), 480–508.

[63, 192]

- 33'. On a special class of polynomials. *Trans. Amer. Math. Soc.* **35** (1933), 559–584.

[55]

OSOFSKY, B. L.

71. On twisted polynomial rings. *J. Algebra* **18** (1971), 597–607.

OSTMAN, H.-H.

50. Euklidische Ringe mit eindeutiger Partialbruchzerlegung. *J. reine angew. Math.* **188** (1950), 150–161.

PARIMALA, S. and SRIDHARAN, R.

75. Projective modules over polynomial rings over division rings. *J. Math. Kyoto Univ.* **15** (1975), 129–148.

PASSMAN, D. S. (see MONTGOMERY, M. S.)

PAUL, Y.

73. Unique factorization in a 2-fir with right ACC_1 . *Riv. Mat. Univ. Parma* (3) **2** (1973), 115–119. [192]

PERRIN, D. (see BERSTEL, J.)

PIERCE, R. S.

82. "Associative Algebras". Springer-Verlag, New York, Heidelberg, Berlin, 1982. [193]

PIHKILKOV, S. A.

74. Primitivity of a free associative algebra with a finite number of generators (Russian). *Uspekhi Mat. Nauk* **29** (1974), No. 1 (175), 183-184.

POINCARÉ, H.

[192]

PONTRYAGIN, L.

39. "Topological Groups". Princeton Univ. Press, Princeton, 1939. [297]

PREST, M. Y.

83. Existentially complete prime rings. *J. London Math. Soc.* (2) **28** (1983), 238-246. [86]

REUTENAUER, C. (see also BERSTEL, J.)

83. Sulla fattorizzazione dei codici, *Ricerche di Mat.* XXXII (1983), 115-130. [124]

RÉVÉSZ, G.

81. "Universal fields of fractions, their orderings and determinants". Thesis, London University, 1981. [297]

83. On the abelianized multiplicative group of a universal field of fractions. *J. Pure Appl. Algebra* **27** (1983), 277-297. [451, 488]

- 83'. Ordering epic R -fields. *Manu. Math.* **44** (1983), 109-130. [488]

RICHARD, J.

70. Représentations matricielles des séries rationnelles en variables non-commutatives. *C. R. Acad. Sci. Paris Sér. A-B* **270** (1970), A224-A227.

RICHARDSON, A. R.

27. Equations over a division algebra. *Messenger of Math.* **57** (1927), 1-6. [539]

RIESINGER, R.

82. Geometrische Überlegungen zum rechten Eigenwertproblem für Matrizen über Schiefkörpern. *Geom. Dedicata* **12** (1982), 401-405. [539]

RINGEL, C. M. (see also DLAB, V.)

79. The spectrum of a finite-dimensional algebra. "Ring Theory" (ed. F. V. Oystaeyen). *Proc. 1978 Antwerp Conference*, M. Dekker, New York, Basel, 1979, pp. 535-597. [488]

ROBERTS, M. L.

[219, 379]

82. "Normal forms, factorizations and eigenrings in free algebras". Thesis, London University, 1982. [225, 281, 297]

84. Endomorphism rings of finitely presented modules over free algebras. *J. London Math. Soc.* (2) **30** (1984), 197-209. [297]

ROBSON, J. C. (see also LEVY, L. S.)

67. Rings in which finitely generated right ideals are principal. *Proc. London Math. Soc.* (3) **17** (1967), 617-628.

72. Idealizers and hereditary Noetherian prime rings. *J. Algebra* **22** (1972), 45-81. [34]

RODOSSKI, K. I.

80. On Euclidean rings (Russian). *Dokl. Akad. Nauk SSSR* **253** (1980), 819-822. [151]

ROOS, J.-E.

67. "Locally distributive spectral categories and strongly regular rings". *Reports of the Mid-West Category Seminar, Springer Lecture Notes in Math.* No. 47, Springer-Verlag, Berlin, 1967. [226]

81. Homology of loop spaces and of local rings. *Proc. 18th Scand. Cong. Math. Progress in Math.*, Birkhäuser, Basel, 1981, pp. 441–468. [151]

ROSSET, S.

[45]

RUSSELL, P.

80. Hamburger–Noether expansions and approximate roots of polynomials, *Manu. Math.* 31 (1980), 29–95. [362]

SAMUEL, P.

[94]

68. Unique factorization. *Amer. Math. Monthly* 75 (1968), 945–952. [142]

71. About Euclidean rings. *J. Algebra* 19 (1971), 282–301. [151]

SANOV, I. N.

67. Euclidean algorithm and onesided prime factorization for matrix rings (Russian). *Sibirsk. Mat. Zh.* 8 (1967), 846–852. [93]

SARGES, H.

76. Ein Beweis des Hilbertschen Basissatzes. *J. reine angew. Math.* 283/4 (1976), 436–437. [63]

SCHLESINGER, L.

1897. "Handbuch der Theorie der Differentialgleichungen". B. G. Teubner, Leipzig, 1897. [191]

SCHMEIDLER, W. (see NOETHER, E.)

SCHOFIELD, A. H. (see also COHN, P. M.) [73, 183, 235, 297, 371, 420, 462, 487, 539]

85. "Representations of rings over skew fields". London Math. Soc. Lecture Notes No. 92, Cambridge University Press, Cambridge, 1985. [362]

- 85'. Artin's problem for skew field extensions, *Math. Proc. Camb. Phil. Soc.* 97 (1985), 1–6.

SCHUPP, P. E. (see LYNDON, R. C.)

SCHUR, I.

04. Über vertauschbare lineare Differentialausdrücke. *Berl. Math. Ges. Sitzber.* 3 (Archiv. d. Math. Beilage (3) 8) (1904), 2–8. [340, 539]

SCHÜTZENBERGER, M.-P. (see also BERSTEL, J.)

59. Sur certains sous-demi-groupes qui interviennent dans un problème de mathématiques appliquées. *Publ. Sci. Univ. d'Alger. Ser. A* 6 (1959), 85–90. [379]

62. On a theorem of R. Jungen. *Proc. Amer. Math. Soc.* 13 (1962), 885–890. [487]

SCHWARZ, L.

47. Zur Theorie der nichtkommutativen rationalen Funktionen. *Ber. Math. Tagung Tübingen* 1946 (1947), 134–136.

49. Zur Theorie des nichtkommutativen Polynombereichs und Quotientenrings. *Math. Ann.* 120 (1947/49), 275–296. [486f.]

SERBIN, H.

38. Factorization in principal ideal rings. *Duke Math. J.* 4 (1938), 656–663.

SERRE, J.-P.

[380]

SHAFAREVICH, I. R.

66. On some infinite-dimensional groups. *Rend. Mat. e Appl.* (5) **25** (1966), 208–212. [379]

SHEPARD, G. C.

[380]

SHEPHERDSON, J. C.

51. Inverses and zero-divisors in matrix rings. *Proc. London Math. Soc.* (3) **1** (1951), 71–85. [62]

SIERPIŃSKA, A.

73. Radicals of rings of polynomials in non-commutative indeterminates (Russian summary). *Bull. Acad. Polon. Sci. Ser. Math. Astronom. Phys.* **21** (1973), 805–808.

SILVESTER, J. R.

73. On the K_2 of a free associative algebra. *Proc. London Math. Soc.* (3) **26** (1973), 35–56.

SINGH, S.

[94]

SIZER, W. S.

75. "Similarity of sets of matrices over a skew field". Thesis, London University, 1975.

SKORNYAKOV, L. A.

65. On Cohn rings (Russian). *Algebra i Logika* **4**(3) (1965), 5–30. [152]
 66. On left chain rings (Russian). *Izv. vyss. uchebn. zaved. Mat.* No. 53 (1966), 114–117.
 67. The homological classification of rings (Russian). *Mat. Vesnik* **4**(19) (1967), 415–434.

SMITH, K. C.

[521]

74. "Theory and application of noncommutative polynomials". Univ. of Oklahoma Lecture Notes, 1974.
 77. Algebraically closed noncommutative polynomial rings. *Comm. in Algebra* **5**(4) (1977), 331–346.

SMITS, T. H. M.

68. Skew polynomial rings. *Indag. Math.* **30** (1968), 209–224. [539]
 68'. Nilpotent S -derivations. *Indag. Math.* **30** (1968), 72–86.

SONTAG, E. D. (see DICKS, W.)

SPECKER, E.

50. Additive Gruppen von Folgen ganzer Zahlen. *Portugal. Math.* **9** (1950), 131–140. [297]

SPRINGER, T. A.

77. "Invariant theory". Springer Lecture Notes in Math. No. 585, Springer-Verlag, Berlin, 1977. [380]

SRIDHARAN, R. (see OJANGUREN, M., PARIMALA, S.)

STAFFORD, J. T.

85. Stably free projective right ideals. *Compositio Math.* **54** (1985), 63–78. [63]

STANLEY, R. P.

79. Invariants of finite groups and their applications to combinatorics. *Bull. Amer. Math. Soc.* **1** (1979), 475–511. [380]

STENSTRÖM, B.

75. Rings of quotients. *Grundl. Math. Wiss.* 217, Springer-Verlag, Berlin, 1975. [296]

STEPHENSON, W.

74. Modules whose lattice of submodules is distributive. *Proc. London Math. Soc.* (3) **28** (1974), 291–310. [199, 226]

STEVIN, S.

1585. "Arithmétique". Antwerp, 1585. (Vol. II of collected works 1958.) [150]

SUSLIN, A. A.

76. On a theorem of Cohn (Russian). "Rings and Modules". *Zap. Nauchn. Sem. Leningrad Otdel. Mat. Inst. Steklov (LOMI)* **64** (1976), 127–130.

SUSHKEVICH, A. K.

36. On the embedding of semigroups in groups of fractions (Ukrainian). *Zap. Khark. Mat. O-va* **4**(12) (1936), 81–88. [486]

SYLVESTER, J. J.

1884. On involutants and other allied species of invariants to matrix systems. *Johns Hopkins Univ. Circ.* **III** (1884), 9–12, 34–35. [227, 253, 297]

TARASOV, G. V.

67. On free associative algebras. *Algebra i Logika* **6**(4) (1967), 93–105. [152]

TAYLOR, J. L.

73. Functions of several noncommuting variables. *Bull. Amer. Math. Soc.* **79** (1973), 1–34.

TAZAWA, M.

33. Bemerkung über den Elementarteilersatz. *Proc. Imper. Acad. Japan* **9** (1933), 468–471.

TEICHMÜLLER, O.

37. Der Elementarteilersatz für nichtkommutative Ringe. *S.-Ber. Preuss. Akad. Wiss.* (1937), 169–177. [539]

TODD, J. A.

[380]

TÖRNER, G. (see BRUNGS, H.-H.)

TYLER, R.

75. On the lower central factors of a free associative ring. *Canad. J. Math.* **27** (1975), 434–438.

VALETTE, J. (see GOURSAUD, J. M.)

VALITSKAS, A. I.

82. The absence of a finite basis for the quasi-identities for the quasi-variety of rings embeddable in radical rings (Russian). *Algebra i Logika* **21**(1) (1982), 13–36. [485, 488]

84. The representation of finite-dimensional Lie algebras in radical rings (Russian). *Doklady Akad. Nauk SSSR* **279** (1984), 1297–1300.

VAMOS, P.

76. Test modules and cogenerators. *Proc. Amer. Math. Soc.* **56** (1976), 8–10. [192]

78. Finitely generated artinian and distributive modules are cyclic. *Bull. London Math. Soc.* **10** (1978), 287–288. [199]

WARFIELD, JR, R. B. (see GOODEARL, K. R.)

VAN DE WATER, A.

70. A property of torsion-free modules over left Ore domains. *Proc. Amer. Math. Soc.* **25** (1970), 199–201.

VAN DER WAERDEN, B. L.

30. "Moderne Algebra". Springer-Verlag, Leipzig, 1930. [486]

WEBB, P.

[370]

WEBBER, D. B.

70. Ideals and modules in simple Noetherian hereditary rings. *J. Algebra* **16** (1970), 239–242. [5]

WEDDERBURN, J. H. M.

21. On division algebras. *Trans. Amer. Math. Soc.* **22** (1921), 129–135. [538f.
32. Noncommutative domains of integrity. *J. reine angew. Math.* **167** (1932), 129–141. [152]

WEYL, H.

[487]

WHAPLES, G. (see BRAY, U.)

WILLIAMS, R. E.

[336]

68. A note on weak Bezout rings. *Proc. Amer. Math. Soc.* **19** (1968), 951–952. [86
68'. Sur une question de P. M. Cohn. *C. R. Acad. Sci. Paris, Ser. A* **267** (1968), 79–80. [164
69. A note on rings with weak algorithm. *Bull. Amer. Math. Soc.* **75** (1969), 959–961. (cf. Correction in *Math. Rev.* **40** (1970), No. 7308.)
69'. On the free product of rings with weak algorithm. *Proc. Amer. Math. Soc.* **23** (1969), 596–597.

WOLF, M. C.

36. Symmetric functions of noncommuting elements. *Duke Math. J.* **2** (1936), 626–637.

WONG, R.

78. Free ideal monoid rings. *J. Algebra* **53** (1978), 21–35. [86
a. Hereditary monoid rings (to appear).

WRIGHT, D. (see BASS, H.)

ZAKS, A.

70. Hereditary local rings. *Michigan Math. J.* **17** (1970), 267–272.
72. Restricted left principal ideal rings. *Isr. J. Math.* **11** (1972), 190–215.
82. Atomic rings without ACC on principal right ideals. *J. Algebra* **74** (1982), 223–231. [154]

Subject index

A

α -adic filtration, 125
 Abelian category, 546
 Abelianizing a ring (or group), 341, 448f.
 Absolute property, 205
 Abstract atomic factor, 215
 ACC = ascending chain condition, 5, 154
 ACC_n = ACC on n -generator submodules, 5, 71f., 287
 ACC_{ds} = ACC on finitely generated submodules with dense inclusions, 289
 Additive category, 545
 Additive functor, 549
 Adjoint associativity, 552
 Adjoint functor, pair, 548
 Admissible system, matrix, 384, 421
 Affine automorphism, 342
 Affine scheme, 487
 Algebraic algebra, 220
 Algebraic element (left, right), 515
 Algebraic matrix, 506
 Algebraic power series, 134, 291
 Amitsur's theorem (on generalized polynomial identities), 281ff.
 Anti-ideal, 322, 334, 376, 379
 Artin's problem, 521
 Artinian ring = ring with (left, right) DCC, 22
 Associated, 25, 153
 Atom, (n -) atomic, 73, 154, 168
 Augmentation ideal, 61
 Augmentation preserving automorphism, 61, 342

B

Baer's criterion, 80, 549

Basis theorem for abelian groups, 494
 Bergman's centralizer theorem, 340, 378
 Bergman's conjugacy theorem, 529
 Bergman-Dicks localization theorem, 459
 Bezout domain (left, right), 12, 69
 Bezout identity, 86
 Bezout ring (weak), 86
 Biased, 357
 Biassociated, 422
 Bicentralizer, 282
 Bidegree, 344
 Binomial extension, 518
 Birkhoff's representation theorem, 180, 544
 Bordered matrix, 405, 421
 Bound component, module, 228f.
 Bound, strongly, 234
 Bounded element, module, 312, 379
 Bounded linear transformation, matrix, 505

C

Cancellation monoid, 35
 Capacity, 22
 $\text{Card}(I)$ = cardinal of I , also $|I|$, xx
 Category, 544
 Cayley-Hamilton theorem, 505, 539
 Central extension, 518
 Chain ring, 170
 Characteristic of a module, 24
 Chase's lemma, 175
 Cleavage, cleft, 189
 Closed submodule, closure, 245, 251
 Closed subring, 102
 Code, 324, 379
 Cofinal, 73, 289
 Cohen's theorem, 310

Coherent family of matrix groups, 74
 Coherent ring, 258, 554
 Coinduced extension, 552
 Cokernel, 545
 Column rank, 247
 Comaximal relation, 28, 161
 Comaximal transposition, 178, 270
 Commaximally transposable, 171
 Comma category, 548
 Comma-free code, 379
 Commensurable, xix
 Companion matrix, 505
 Comparison theorem for numerators
 and denominators, 429
 Completable matrix, 15
 Complete direct decomposition, 186
 Complete inversely filtered ring, 126
 Complete lattice, 540
 Completely primary ring, 187
 Completely reducible module, 189
 Completion (of a filtered ring), 126
 Complex-skew polynomial ring, 54, 164,
 179f., 191, 210, 225
 Conductor, 299, 304
 Conical monoid, 19, 153, 321
 Conjugate, 504
 Connected (inversely filtered ring), 129
 Connecting homomorphism, 550
 Constants, 40
 Content, 107
 Continuant polynomial, 117
 Convex, 219, 526
 Coprime relation, 161
 Core, 384
 Cover (in a lattice), 540
 Cramer's rule, 384
 Cyclic matrix, 504, 539

D

DCC = descending chain condition, 10
 Decomposable (left, right), 184, 316
 Dedekind lemma, 372f.
 Defect theorem, 329
 Degenerate matrix, 400
 Degree, 60, 100, 111, 344, 349
 formal, 104

Degree-function, 51, 139
 Degree of filtration, 95
 Denominator, 36, 384
 Dense subcategory, 314
 Dense submodule, 245
 Dependence number, $\lambda_v(-)$, 96, 151
 inverse, $\mu_v(-)$, 125
 Dependence relation, 64, 100, 475
 Dependence relative to a filtration, 95
 Dependence relative to a transfinite
 function, 139
 Depth, 386, 432ff.
 Derivation, 39
 higher, 137, 358, 523
 Derived set, 88, 151
 Determinantal sum, 396
 DFL = distributive factor lattice, 199,
 207
 Diagonal reduction, 489, 538
 Diagonal sum, xix, 395
 Diamond lemma, 148
 Dicks commutator test, 355f.
 Dieudonné determinant, 382, 448
 Differential equations, 63
 Differential operators, 54, 58
 Differential transformations, 503
 Dimension, injective, projective, 258, 551
 Direct product, 544
 Direct sum, 545
 Directed set, 292, 398
 Distributive factor lattice, DFL, 199
 Distributive lattice, 542
 Distributive module, 193
 Division algorithm, 87ff.
 Division closure, 335, 387
 Divisor group, $\mathbf{D}(-)$, 451
 DL = category of finite distributive
 lattices, 211
 Duality = category anti-equivalence, 548
 factorial, 166
 Duality in lattices, 540
 Duality for modules, xviii, 165, 232ff.

E

E-related, 119
 Eigenring, $\mathbf{E}(-)$, 30, 63, 192
 scalar, 220

Eigenvalue, 504, 510
 Elementary divisor, 495
 Elementary divisor ring, 492, 539
 Elementary embedding, 287
 Elementary matrix, operation, 490
 End of a graph, 352
 Epic, epimorphism, 545
 Epic R -field, 388
 Equivalence of categories, 548
 Equivalence of factors, 215
 Essential extension, 235
 Essential left factor, 243
 Euclidean algorithm, 89, 121f., 151
 Euclidean ring, 88
 Euler's theorem (homogeneous functions), 62
 Exact functor (left, right), 549
 Exact sequence, 546
 Exchange principle, 126
 Ext, extension of modules, 551
 Extremity of a graph, 351

F

Factor closed, 415, 454
 Factor complete, 454
 Factorial duality, 166
 Fibonacci numbers, 117
 Field of fractions, 38, 251, 388
 Field spectrum, $X(-)$, 410, 487
 Filtered ring, 95
 truncated, 111
 Filtration, 94
 α -adic, 125
 Final object, 545
 Finitely generated (presented, related), 24
 Fir (left, right), n -fir, 67, 71, 85
 Fir, one-sided, 142, 176, 536
 Firoid, 86
 Fitting's lemma, 187, 197, 240
 Five-lemma, 547
 Fixed ring, 362
 Flat module, 49, 552
 Formal degree, 104
 Formal power series, 55
 Fox derivative, 55

Fractions (monoid, group, ring, field), 34ff.
 Free algebra, 59, 106
 Free ideal ring (=fir), 71
 Free monoid, 59
 Free product of groups, 349
 Frobenius inequality, 261
 Full matrix, 159, 192
 Full operation, 422
 Full subcategory, 545
 Fully atomic, 168
 Fully invariant, 195
 Fully inverting homomorphism, 415
 Fully reducible, 189, 314
 Functor, 547

G

Galois correspondence for free algebras, 376ff., 380
 Gauss's lemma, 85f., 107, 124
 GE_2 , standard form, 115f.
 GE-related, 119
 GE-ring, 76, 86
 Generalized polynomial identity, GPI, 282
 Gerasimov-Malcolmson localization theorem, 484
 GL_n = general linear group, xviii
 GL-related, 119, 162
 Global dimension, 551
 Global section (rational, integral), 487
 Graded ring, 100
 Grothendieck category, 239
 Group of fractions, 34

H

Hasse's criterion, 94
 HCF, HCLF, HCRF = highest common (left, right) factor (= greatest common divisor), 122, 150, 154
 Height (of lattice element), 542
 Hereditary ring, 11
 Hermite ring, 15, 63

Higher derivation, 137, 358, 523
 Higman's trick, 272
 Hilbert basis theorem, 53, 63
 Hilbert Theorem 90, 521
 Hilbert series, 107, 364ff.
 HNN-construction, 114, 151f., 421
 Hollow matrix, 160, 397
 Homogeneous, 345
 Honest homomorphism, 250, 415

I

I-atom, *I*-prime, 156, 301
I-decomposable, 316
 IBN = invariant basis number, 6, 63
 Idealizer, *I*(-), 30, 63
 Indecomposable, 186, 316
 Index of a matrix, *i*(-), 28
 Inert (totally, *n*-), 83, 86
 Inertia theorem, 133, 152
 Initial object, 545
 Injective dimension, 551
 Injective hull, 549
 Inner derivation, 40
 Inner rank, 248, 487
 Integral element, closure, 298f., 303
 Intersection theorem for firs, 295, 297
 Interval, 540
 Invariant basis number (IBN), 6, 63
 Invariant, 362, 379
 Invariant elements, *I*(-), 154, 156, 300, 379
 Invariant factors, 495, 504
 Invariant ring (monoid), 154, 156
 Inverse dependence number, $\mu_v(-)$, 125
 Inverse filtration, 125
 Inverse weak algorithm (*n*-term), 125
 Inverting, 34, 382
 Involution, 294
 Irreducible topological space, 411
 Irredundant decomposition, 411, 543
 Isomorphism in a category, 545
 Isomorphism of factorizations, 164
 Isomorphism of idempotent matrices, 20
 Isotone (= order-preserving) mapping, 211
 Iterated skew polynomial ring, 532

J

Jacobian matrix, problem, 355, 380
 Jacobson radical, *J*(-), 20
 Join, 540, 543
 Join-irreducible, 212, 543
 de Jonquières automorphism, 342, 379
 Jordan-Hölder theorem, 168, 198, 214, 542
J-ring, 532
 Jung-van der Kulk theorem, 348, 379

K

K_0 , 19, 63
 K_1 , 449
 Kaplansky's theorem (projective modules), 12, 63
 Kernel, 45, 545
 Kharchenko-Galois correspondence, 376
 Kharchenko-Lane theorem, 363f., 380
 Klein's theorem, 82, 86
 Kraft-McMillan inequality, 325, 379
 Kronecker functional ring, 308
 Krull domain, 302
 Krull-Schmidt theorem, 186, 241, 297, 494, 543
 Kurosh-Ore theorem, 185, 543

L

$L(cR, R)$ = lattice of principal right ideals between cR and R , 80, 179
 Lane's lemma, 357, 380
 $\text{Lat}_R(M)$ = lattice of submodules of M , 3
 Large (left, right), 36
 Laurent series, 56, 521ff.
 LCM, LCRM, LCLM = least common (right, left) multiple, 81, 154
 Leading form, 344
 Leading term, 101, 140, 527
 Leapfrog construction, 117
 Leibniz's formula, 46
 Length of a lattice, 542
 Length of an element, monomial, 60, 322, 494

Level, 174
 Lifting property, 77
 Linear automorphism, 342
 Linear differential operator, 191, 506
 Linearization by enlargement, 272
 Link in a lattice, 180
 Local homomorphism, 388
 Local rank, 255
 Local ring (scalar, matrix), 22
 Localization, 37, 81, 390
 Localization theorems, 457ff., 472, 484
 Locally free module, 48
 Lower multiplicative set, 382
 Lower segment, 211, 333

M

Magnus' theorem, 296
 Makar-Limanov-Czerniakiewicz theorem, 355, 379
 Malcev conditions, 35, 82
 Malcev-Neumann construction, 528
 Malcolmson's criterion, 485
 Matrix algebra, 220
 Matrix basis, 1
 Matrix (pre-) ideal, 396ff.
 Matrix local ring, 22
 Matrix object, 463
 Matrix reduction functor, $\mathfrak{M}_n$, 147f., 152
 Maximal code, 325
 Meet, 540
 Meta (semi-) fir, 70, 73
 Metro-equation, 515
 Minimal admissible matrix, 428
 Minimal factorization, 248
 Mixed free algebra, 62
 Modular lattice, law, 540
 Monic matrix over a free algebra, 272
 Monomial, 60
 Monomial term, 105
 Monomorphism, 545
 Morita equivalence, 4, 62, 78
 Multiplicative set, 382

N

Nagata's formula, 364

Nagata's theorem, 157
 Nakayama's lemma, 21
 Natural isomorphism, transformation, 548
 Negative module, Neg, 236, 297
 Noetherian ring = ring with (left, right) ACC, 15
 Non-singular matrix, xix
 Normal form in a free algebra, 59
 Normal closure, 443
 Normalizer, 443
 Null matrix, 3
 Nullity, 161
 Numerator, 384

O

One-sided fir, 142, 176, 536
 Opposite category, 544
 Order, 55, 60, 421
 Ordered series ring, 527
 Ore condition, domain, 38
 Ore set, 35
 Outer derivation, 40
 Outer rank, 252

P

Parallelogram law, 541
 Partition lemma, 67
 Persistent property, 205, 442
 Perspective, 541
 Place permutation, 364
 Polycyclic module, 244
 Pontryagin's theorem, 297
 Positive module, Pos, 236, 297
 Pos = category of finite partially ordered sets, 211
 Power series D-ring, 61
 Preadditive category, 545
 Prefix set, code, 324
 Presentation, 552
 Primal ring, 252
 Primary decomposition, 184
 Primary element, 191
 Prime, 449
 Prime element, 154, 156

Prime matrix, 160, 425
 Prime matrix ideal, 401
 Prime module, 236
 Principal ideal domain, ring, PID, 71
 Principal valuation ring, PVR, 130, 173
 Product of matrix ideals, 400
 Projective, in a lattice, 541
 Projective dimension, 551
 Projective free ring, 17, 63
 Projective module, 10, 549
 Projective trivial ring, 18
 Proper factorization, 268
 Proper matrix relation, 161, 266
 Proper specialization, 393
 Protorsion module, 239
 Pseudolinear extension, 517
 Pseudolinear mapping, 503
 Pseudo-Sylvester domain, 263, 297, 418
 Pseudo-valuation, 95
 Pullback, 545
 Pure extension, 518
 Pure subalgebra, 341
 Pushout, 545
 PVR = principal valuation ring, 130, 173

Q

Quasi-Frobenius ring, 304
 Quasi-identities, 414, 419, 486
 Quasi-variety, 419, 486
 Quaternions, 23, 93, 158, 225f., 514
 Quotient monoid, 154

R

R-field (universal), 388f.
 Radical of a matrix ideal, 402
 Jacobson, 20, 171
 tertiary, 321
 Rank formula (modules over semifirs), 68
 Rank of a free algebra, 60
 Rank of a free module, 6
 Rank of a matrix, 245ff., 424, 431
 Rank of a module, 48, 251
 Rational closure, 382

Rationality criterion, 526
 Rational power series, 134
 Rational section, 487
 Reduced admissible matrix, 429
 Refinement of a factorization, 178
 Regular, xix
 Regularly embedded, 330
 Relation (*n*-term, trivial), 64, 100
 Relevant *N*-value, 460
 Resolution (injective, projective), 258, 551
 Retract of a ring, 70, 221
 Right invariant, 203, 301
 Rigid domain, UFD, 170, 192
 Rigid element, 41, 170
 Rigid factorization, 190
 Rigid monoid, 41
 Ring of fractions, 37
 Ringlike matrix object, 463
 Ringoid, 86, 545
 Root, 323, 339
 Row rank, 247

S

S-ring, 263, 418
 Σ -inverting homomorphism, 382
 Σ -inverting ring (universal), 390
 Σ -rational closure, 382
 Saturated set, saturation, 394, 467
 SAut, Special automorphism group, 357
 Schanuel's lemma, 24f., 192, 552
 Schreier-Lewin formula, 109, 334
 Schreier refinement theorem, 178, 542
 Schreier set, 333
 Schur's lemma, 167, 222
 Section, global (integral, rational), 487
 Segment (lower, upper), 211
 Semi-Artinian, -Noetherian, 195
 Semifir, 67, 85
 Semifree module, 48
 Semihereditary (left, right, weakly), 11, 13, 63
 Semilinear mapping, 503
 Semiprime matrix ideal, 402
 Semisimple Artinian ring, 193, 550
 Shear, 342
 Similar elements, matrices, $\sim$, 25, 162, 192

Similar right ideals, 158
 Simple, $\mathcal{T}$ -simple, 167
 Simple N-value, 450
 Singular kernel, 391
 Singular matrix, xix
 Singularity support, 286, 410
 Skew formal Laurent series, 522f.
 Skew Laurent polynomial, 54
 Skew polynomial ring, 52, 497ff.
 iterated, 532
 Skew rational function field, 53, 521
 Source, 544
 Spacial module, 254
 Special automorphism, 357
 Specialization, 388
 Specialization lemma, 285, 297
 Spectral space, 412
 Spectrum of a matrix, 510
 Split exact, 549
 Squarefree module, 196
 Stabilize, stab, 342
 Stable atom, 17
 Stable GL, 449
 Stable rank, 261
 Stably associated, 26, 63, 162
 Stably free, 15
 Stably full, 262
 Stably honest, 418
 Standard basis in R^n , ^nR , 3
 Standard form in GE_2 , 115f.
 Strictly bordered, 426
 Strictly cyclic module, 192
 Strong DFL-property, 201, 299f.
 Strong G-ring, G_n -ring, 76
 Strong prime ideal, 177
 Strongly regular, 419
 Subdirect product, 544
 Subdirectly irreducible, 544
 Suffix set, 324
 Support, 59, 354, 526
 Sylvester domain, 253, 297, 417
 Sylvester's law of nullity, 253, 297

T

Tame automorphism, 343
 Target, 544

Tensor D -ring (on a D -bimodule), 61
 Tensor product of modules, 552
 Tertiary radical, 321
 Test module, 192
 Three-by-Three lemma, 547
 Top component, 344
 Topological fir, 128
 Torsion element, submodule, 46
 Torsion class, torsion free class, 228, 296
 Torsion free module, 46
 Torsion module (over semifir), 164, 239, 297
 Total divisor, 489
 Totally coprime, 315, 435
 Totally inert, 84
 Totally unbounded, 316
 Tr_n = group of triangular matrices, 74
 Trace form, 372
 Transcendental matrix, 506
 Transduction, 105
 Transfinite degree function, 139
 Transfinite division algorithm, 90
 Transfinite weak algorithm, 139, 152
 Translation, 342
 Transpose of a module, $Tr(-)$, 233, 297
 Transvection, 74
 Tree, 349
 Trivial filtration, 95
 Trivial operation, 405, 422, 469
 Trivial relation, 64, 475
 Trivializable, 65, 101, 475
 Truncated filtered ring, 111

U

$U(R)$ = group of units of R , xviii
 UFD = unique factorization domain, 153, 164, 192
 UF-monoid, 153f.
 UGN (= unbounded generating number), 6, 63
 Ultraproduct theorem, 286
 Unbound, 228
 Unbounded generating number (UGN), 6, 63
 Unbounded, totally, 316
 Uncleft, totally, 189

Unfactorable matrix, 269
 Unique factorization domain (UFD),
 153, 164, 192
 Unique remainder algorithm, 90, 532, 539
 Unitriangular matrix, 74
 Universal denominators, theorem on,
 434
 Universal field of fractions, R -field, 281,
 389
 Universal group of a monoid, 35
 Universal localization, 390
 Universal mapping property of free
 algebras, 59
 Universal T -inverting monoid, ring, 34,
 36
 Universal object, 548
 Upper multiplicative set, 382

V

Valuation (ring), 302, 452, 531
 Value, 445
 Vandermonde's identity, 360

W

Weak algebra basis, 104, 129

Weak algorithm (n -term), 96, 273
 Weak algorithm (n -term) in a graded
 ring, 101f.
 Weak algorithm (n -term), inverse, 125
 Weak algorithm, transfinite, 139
 Weak v -basis of a right ideal, 98, 128
 Weak global dimension, 256ff., 553
 Weakly finite ring, 6, 63
 Weakly semihereditary, 13
 Weight, 342f.
 Weyl algebra, A_1 , 55, 265
 Whitehead group, 449
 Whitehead lemma, 451, 453
 Width of a graph, 351
 Wild automorphism, 342

X

X -inner, X -outer automorphism, 372

Z

Zero delay code, 324
 Zero (left, right) of a polynomial, 514
 Zero object, 545