Introductory Lectures on Rings and Modules

JOHN A. BEACHY

London Mathematical Society
Student Texts **47**

CAMBRIDGE

more information - www.cambridge.org/9780521643405

LONDON MATHEMATICAL SOCIETY STUDENT TEXTS

Managing editor: Professor C.M. Series, Mathematics Institute
University of Warwick, Coventry CV4 7AL, United Kingdom

3 Local fields, J.W.S. CASSELS
4 An introduction to twistor theory: Second edition, S.A. HUGGETT & K.P. TOD
5 Introduction to general relativity, L.P. HUGHSTON & K.P. TOD
7 The theory of evolution and dynamical systems, J. HOFBAUER & K. SIGMUND
8 Summing and nuclear norms in Banach space theory, G.J.O. JAMESON
9 Automorphisms of surfaces after Nielsen and Thurston, A. CASSON & S. BLEILER
11 Spacetime and singularities, G. NABER
12 Undergraduate algebraic geometry, MILES REID
13 An introduction to Hankel operators, J.R. PARTINGTON
15 Presentations of groups: Second edition, D.L. JOHNSON
17 Aspects of quantum field theory in curved spacetime, S.A. FULLING
18 Braids and coverings: selected topics, VAGN LUNDSGAARD HANSEN
19 Steps in commutative algebra, R.Y. SHARP
20 Communication theory, C.M. GOLDIE & R.G.E. PINCH
21 Representations of finite groups of Lie type, FRANÇOIS DIGNE & JEAN MICHEL
22 Designs, graphs, codes, and their links, P.J. CAMERON & J.H. VAN LINT
23 Complex algebraic curves, FRANCES KIRWAN
24 Lectures on elliptic curves, J.W.S. CASSELS
26 An introduction to the theory of L-functions and Eisenstein series, H. HIDA
27 Hilbert Space: compact operators and the trace theorem, J.R. RETHERFORD
28 Potential theory in the complex plane, T. RANSFORD
29 Undergraduate commutative algebra, M. REID
31 The Laplacian on a Riemannian manifold, S. ROSENBERG
32 Lectures on Lie groups and Lie algebras, R. CARTER, G. SEGAL & I. MACDONALD
33 A primer of algebraic *D*-modules, S.C. COUTINHO
34 Complex algebraic surfaces, A. BEAUVILLE
35 Young tableaux, W. FULTON
37 A mathematical introduction to wavelets, P. WOJTASZCZYK
38 Harmonic maps, loop groups and integrable systems, M. GUEST
39 Set theory for the working mathematician, K. CIESIELSKI
40 Ergodic theory and dynamical systems, M. POLLICOTT & M. YURI
41 The algorithmic resolution of diophantine equations, N.P. SMART
42 Equilibrium states in ergodic theory, G. KELLER
44 Classical invariant theory, P. OLVER
45 Permutation groups, P. CAMERON

London Mathematical Society Student Texts 47

Introductory Lectures on Rings and Modules

John A. Beachy
Northern Illinois University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
Sao Paulo, Delhi, Dubai, Tokyo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on title: www.cambridge.org/9780521643405

First published 1999

A catalogue record/or this publication is available/rom the British Library

Library o/Congress Cataloguing in Publication data

Beachy, John A.
Introductory lectures on rings and modules / John A. Beachy
p. cm. - (London Mathematical Society student texts; 47)
ISBN 0 521 643406 (hbk.) ISBN 0 521 644070 (pbk.)
1. Noncommutative rings (Algebra) 2. Modules (Algebra)
I. Title. II. Series
QA251.4.B43 1999
512'A-dc21 9854417 CIP

ISBN 978-0-521-64340-5 Hardback
ISBN 978-0-521-64407-5 Paperback

Transferred to digital printing 2009

Contents

PREFACE

This set of lecture notes is focused on the noncommutative aspects of the study of rings and modules. It is intended to complement the book *Steps in Commutative Algebra*, by R. Y. Sharp, which provides excellent coverage of the commutative theory. It is also intended to provide the necessary background for the book *An Introduction to Noncommutative Noetherian Rings*, by K. R. Goodearl and R. B. Warfield.

The core of the first three chapters is based on my lecture notes from the second semester of a graduate algebra sequence that I have taught at Northern Illinois University. I have added additional examples, in the hope of making the material accessible to advanced undergraduate students. To provide some variety in the examples, there is a short section on modules over the Weyl algebras. This section is marked with an asterisk, as it can be omitted without causing difficulties in the presentation. (The same is true of Section 1.5 and Section 3.4.) Chapter 4 provides an introduction to the representation theory of finite groups. Its goal is to lead the reader into an area in which there has been a very successful interaction between ring theory and group theory.

Certain books are most useful as a reference, while others are less encyclopedic in nature, but may be an easier place to learn the material for the first time. It is my hope that students will find these notes to be accessible, and a useful source from which to learn the basic material. I have included only as much material as I have felt it is reasonable to try to cover in one semester. The role of an encyclopedic text is played by any one of the standard texts by Jacobson, Hungerford, and Lang. My personal choice for a reference is *Basic Algebra* by N. Jacobson.

There are many possible directions for subsequent work. To study noncommutative rings the reader might choose one of the following books: *An Introduction to Noncommutative Noetherian Rings*, by K. R. Goodearl and R. B. Warfield, *A First Course in Noncommutative Rings*, by T. Y. Lam, and *A Course in Ring Theory*, by D. S. Passman. After finishing Chapter 4 of this text, the reader should have the background necessary to study *Representations and Characters of Finite Groups*, by M. J. Collins. Another possibility is to study *A Primer of Algebraic D-Modules*, by S. C. Coutinho.

I expect the reader to have had prior experience with algebra, either at the advanced undergraduate level, or in a graduate level course on Galois theory and the structure of groups. Virtually all of the prerequisite material can be found in undergraduate books at the level of Herstein's *Abstract Algebra*. For

the sake of completeness, two definitions will be given at this point. A *group* is a nonempty set G together with a binary operation $\cdot$ on G such that the following conditions hold: (i) for all $a, b, c \in G$, we have $a \cdot (b \cdot c) = (a \cdot b) \cdot c$; (ii) there exists $1 \in G$ such that $1 \cdot a = a$ and $a \cdot 1 = a$ for all $a \in G$; (iii) for each $a \in G$ there exists an element $a^{-1} \in G$ such that $a \cdot a^{-1} = 1$ and $a^{-1} \cdot a = 1$. The group G is said to be *abelian* if $a \cdot b = b \cdot a$ for all $a, b \in G$, and in this case the symbol $\cdot$ for the operation on G is usually replaced by a $+$ symbol, and the identity element is denoted by the symbol 0 rather than by the symbol 1. The definition of an abelian group is fundamental, since the objects of study in the text (rings and modules) are constructed by endowing an abelian group with additional structure.

I sincerely hope that the reader's prior experience with algebra has included the construction of examples. Good examples provide the foundation for understanding this material. I have included a variety of them, but it is best if additional ones are constructed by the reader. A good example illustrates the key ideas of a definition or theorem, but is not so complicated as to obscure the important points. Each definition should have several associated examples that will help in understanding and remembering the conditions of the definition. It is helpful to include some that do *not* satisfy the stated conditions.

I would like to take this opportunity to thank my colleagues for many helpful conversations: Bill Blair, Harvey Blau, Harald Ellers (who made corrections in the last chapter), and George Seelinger (who class-tested the manuscript). I would also like to thank Svetlana Butler, Sonia Edghill, Lauren Grubb, Suzanne Riehl, and Adam Slagell, who gave me lists of misprints. I would like to dedicate the book to my daughter Hannah, with thanks for her patience while I was writing, and for her help in proofreading.

John A. Beachy
DeKalb, Illinois
August, 1998

Chapter 1

RINGS

The abstract definition of a ring identifies a set of axioms that underlies some familiar sets: the set $\mathbf{Z}$ of integers, the set of polynomials $\mathbf{Q}[x]$ with coefficients in the field $\mathbf{Q}$ of rational numbers, and the set $\mathrm{M}_n(\mathbf{R})$ of $n \times n$ matrices with coefficients in the field $\mathbf{R}$ of real numbers. Much of the interest in what we now call a ring had its origin in number theory. In the field $\mathbf{C}$ of complex numbers, Gauss used the subset

$$\mathbf{Z}[\mathrm{i}] = \{a + b\mathrm{i} \in \mathbf{C} \mid a, b \in \mathbf{Z}\}$$

to prove facts about the integers, after first showing that unique factorization into 'primes' still holds in this context. The next step was to consider subsets of the form $\mathbf{Z}[\zeta]$, where ζ is a complex root of unity. These rings were used to prove some special cases of Fermat's last theorem.

Kummer was interested in higher reciprocity laws, and found it necessary to investigate unique factorization in subrings of the field $\mathbf{C}$ of complex numbers. He realized that the analog of the prime factorization theorem in $\mathbf{Z}$ need not hold in all such subrings, but he was able to prove such a theorem in enough cases to obtain Fermat's last theorem for exponents up to 100. The modern definition of an ideal was given in 1871 by Dedekind, who proved that in certain subrings of $\mathbf{C}$ every nonzero ideal can be expressed uniquely as a product of

prime ideals. We will see that prime ideals play a crucial role even in the case of noncommutative rings.

The search for a way to extend the concept of unique factorization motivated much of the early work on commutative rings. An integral domain with the property that every nonzero ideal can be expressed uniquely as a product of prime ideals is now called a Dedekind domain. Lasker developed a parallel theory for polynomial rings, in which an ideal is written as an intersection of primary ideals (rather than as a product of prime ideals). Both of these theories were axiomatized by Emmy Noether, who worked with rings that satisfy the ascending chain condition for ideals. The commutative theories are beyond the scope of this text; the interested reader can consult the texts by Sharp [23], Matsumura [21], and Eisenbud [8].

The term 'number ring' was used in 1897 in a paper by Hilbert. The current definition of an abstract ring seems to have first appeared in the 1921 paper "Theory of ideals in rings" published by Emmy Noether. She played a prominent role in the early (1920–1940) development of commutative ring theory, along with Krull.

Section 1.1 and Section 1.2 introduce some of the basic definitions that will be used later in the book. For our purposes, integral domains and rings of matrices form two of the most important classes of rings. Section 1.3 and Section 1.4 provide the tools for a study of unique factorization in integral domains. Section 1.5 introduces general matrix rings, and several other noncommutative examples. This section, together with Sections 2.8 and 3.4, carries forward the noncommutative emphasis of the text, but is not crucial to the development in other sections.

1.1 Basic definitions and examples

Working with polynomials and matrices leads naturally to a study of algebraic systems with two operations, similar to ordinary addition and multiplication of integers. Although we assume that the reader has some knowledge of commutative rings, especially integral domains and fields, we will give the definition of a ring in full detail. Including the study of even the set of 2×2 matrices over the real numbers means that we cannot impose the commutative law for multiplication.

We recall that a *binary operation* on a set S is a function from the Cartesian product $S \times S$ into S. If $*$ is a binary operation on S, then for all ordered pairs $a, b \in S$, the value $a * b$ is uniquely defined and belongs to S. Thus the operation is said to satisfy the *closure* property on S.

Definition 1.1.1 *Let R be a set on which two binary operations are defined, called* addition *and* multiplication, *and denoted by $+$ and $\cdot$, respectively. Then R is called a* ring *with respect to these operations if the following properties hold.*

(i) Associative laws: *For all $a, b, c \in R$,*

$$a + (b + c) = (a + b) + c \quad \text{and} \quad a \cdot (b \cdot c) = (a \cdot b) \cdot c \, .$$

(ii) Commutative law (for addition): *For all $a, b \in R$,*

$$a + b = b + a \, .$$

(iii) Distributive laws: *For all $a, b, c \in R$,*

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{and} \quad (a + b) \cdot c = a \cdot c + b \cdot c \, .$$

(iv) Identity elements: *The set R contains elements 0 and 1 (not necessarily distinct) called, respectively, an* additive identity element *and a* multiplicative identity element, *such that for all $a \in R$,*

$$a + 0 = a \quad \text{and} \quad 0 + a = a \, ,$$

and

$$a \cdot 1 = a \quad \text{and} \quad 1 \cdot a = a \, .$$

(v) Additive inverses: *For each $a \in R$, the equations*

$$a + x = 0 \quad \text{and} \quad x + a = 0$$

have a solution x in R, called the additive inverse *of a, and denoted by $-a$.*

The definition of a ring R can be summarized by stating that R is an abelian group under addition, with a multiplication that satisfies the associative and distributive laws. Furthermore, R is assumed to have a multiplicative identity element. (Note that if $1 = 0$, then R is said to be *trivial*.) Many authors choose not to require the existence of a multiplicative identity, and there are important examples that do not satisfy the associative law, so, strictly speaking, we have given the definition of an 'associative ring with identity element.'

The distributive laws provide the only connection between the operations of addition and multiplication. In that sense, they represent the only reason for studying the operations simultaneously instead of separately. For example, in the following list of additional properties of a ring, any property using the additive identity or additive inverses in an equation involving multiplication must depend on the distributive laws.

Our observation that any ring is an abelian group under addition implies that the cancellation law holds for addition, that the additive identity element is unique, and that each element has a unique additive inverse. The remaining properties in the following list can easily be verified.

Let R be a ring, with elements $a, b, c \in R$.
(a) If $a + c = b + c$, then $a = b$.
(b) If $a + b = 0$, then $b = -a$.
(c) If $a + b = a$ for some $a \in R$, then $b = 0$.
(d) For all $a \in R$, $a \cdot 0 = 0$.
(e) For all $a \in R$, $(-1) \cdot a = -a$.
(f) For all $a \in R$, $-(-a) = a$.
(g) For all $a, b \in R$, $(-a) \cdot (-b) = a \cdot b$.

A ring can have only one multiplicative identity element. If $1 \in R$ and $e \in R$ both satisfy the definition of a multiplicative identity, then $1 \cdot e = e$ since 1 is an identity element, and $1 \cdot e = 1$ since e is an identity element. Thus $e = 1 \cdot e = 1$, showing that 1 is the unique element that satisfies the definition.

Various sets of numbers provide the most elementary examples of rings. In these sets the operation of multiplication is commutative. The set $\mathbf{Z}$ of integers should be the first example of a ring. In this ring we have the additional property that if $a \neq 0$ and $b \neq 0$, then $ab \neq 0$. The set $\mathbf{Q}$ of rational numbers, the set $\mathbf{R}$ of real numbers, and the set $\mathbf{C}$ of complex numbers also form rings, and in each of these rings every nonzero element has a multiplicative inverse.

We next review the definitions of some well-known classes of rings.

Definition 1.1.2 *Let R be a ring.*

(a) *The ring R is called a* commutative ring *if $a \cdot b = b \cdot a$ for all $a, b \in R$.*

(b) *The ring R is called an* integral domain *if R is commutative, $1 \neq 0$, and $a \cdot b = 0$ implies $a = 0$ or $b = 0$, for all $a, b \in R$.*

(c) *The ring R is called a* field *if R is an integral domain such that for each nonzero element $a \in R$ there exists an element $a^{-1} \in R$ such that $a \cdot a^{-1} = 1$.*

According to the above definition, $\mathbf{Z}$ is an integral domain, but not a field. The sets $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{C}$ are fields, since in each set the inverse of a nonzero element is again in the set. We assume that the reader is familiar with the ring $F[x]$ of polynomials with coefficients in a field F. This ring provides another example of an integral domain that is not a field.

If the cancellation law for multiplication holds in a commutative ring R, then for any elements $a, b \in R$, $ab = 0$ implies that $a = 0$ or $b = 0$. Conversely, if this condition holds and $ab = ac$, then $a(b - c) = 0$, so if $a \neq 0$ then $b - c = 0$ and $b = c$. Thus in a ring R with $1 \neq 0$, the cancellation law for multiplication holds

if and only if R is an integral domain. It is precisely this property that is crucial in solving polynomial equations. We note that any field is an integral domain, since the existence of multiplicative inverses for nonzero elements implies that the cancellation law holds.

It is possible to consider polynomials with coefficients from any commutative ring, and it is convenient to give the general definition at this point.

Example 1.1.1 (Polynomials over a commutative ring)

Let R be any commutative ring. We let $R[x]$ denote the set of infinite tuples

$$(a_0, a_1, a_2, \ldots)$$

such that $a_i \in R$ for all i, and $a_i \neq 0$ for only finitely many terms a_i. Two elements of $R[x]$ are equal (as sequences) if and only if all corresponding entries are equal. We introduce addition and multiplication as follows:

$$(a_0, a_1, a_2, \ldots) + (b_0, b_1, b_2, \ldots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots)$$

and

$$(a_0, a_1, a_2, \ldots) \cdot (b_0, b_1, b_2, \ldots) = (c_0, c_1, c_2, \ldots) ,$$

where

$$c_k = \textstyle\sum_{i+j=k} a_i b_j = \sum_{i=0}^{k} a_i b_{k-i} .$$

Showing that $R[x]$ is a ring under these operations is left as Exercise 2, at the end of this section.

We can identify $a \in R$ with $(a, 0, 0, \ldots) \in R[x]$, and so if R has an identity 1, then $(1, 0, 0, \ldots)$ is a multiplicative identity for $R[x]$. If we let $x = (0, 1, 0, \ldots)$, then $x^2 = (0, 0, 1, 0, \ldots)$, $x^3 = (0, 0, 0, 1, 0, \ldots)$, etc. Thus the elements of $R[x]$ can be expressed uniquely in the form

$$a_0 + a_1 x + \cdots + a_{m-1} x^{m-1} + a_m x^m ,$$

allowing us to use the standard notation for the *ring of polynomials over R in the indeterminate x*. Note that two elements of $R[x]$ are equal if and only if the corresponding *coefficients* a_i are equal. We refer to R as the *ring of coefficients* of $R[x]$.

If n is the largest nonnegative integer such that $a_n \neq 0$, then we say that the polynomial has *degree* n, and a_n is called the *leading coefficient* of the polynomial.

Once we know that $R[x]$ is a ring, it is easy to work with polynomials in two indeterminates x and y. We can simply use the ring $R[x]$ as the coefficient ring, and consider all polynomials in the indeterminate y, with coefficients in $R[x]$. For example, by factoring

out the appropriate terms in the following polynomial in x and y we have

$$\begin{aligned} &2x - 4xy + y^2 + xy^2 + x^2y^2 - 3xy^3 + x^3y^2 + 2x^2y^3 \\ &\quad = 2x + (-4x)y + (1 + x + x^2 + x^3)y^2 + (-3x + 2x^2)y^3 , \end{aligned}$$

showing how it can be regarded as a polynomial in y with coefficients in $R[x]$. The ring of polynomials in two indeterminates with coefficients in R is usually denoted by $R[x, y]$, rather than by $(R[x])[y]$, which would be the correct notation.

Now let D be any integral domain. Then the ring $D[x]$ of all polynomials with coefficients in D is also an integral domain. To show this we note that if $f(x)$ and $g(x)$ are nonzero polynomials with leading coefficients a_m and b_n, respectively, then since D is an integral domain, the product $a_m b_n$ is nonzero. This shows that the leading coefficient of the product $f(x)g(x)$ is nonzero, and so $f(x)g(x) \neq 0$. In actuality, we have $f(x)g(x) \neq 0$ because the degree of $f(x)g(x)$ is equal to $\deg(f(x)) + \deg(g(x))$.

We note that in the definition of the polynomial ring $R[x_1, x_2, \ldots, x_n]$, the coefficients could just as easily have come from a noncommutative ring. In that case, in defining the notion of a polynomial ring we generally require that the indeterminates commute with each other and with the coefficients.

Many familiar rings fail to be integral domains. Let R be the set of all functions from the set of real numbers into the set of real numbers, with ordinary addition and multiplication of functions (not composition of functions). It is not hard to show that R is a commutative ring, since addition and multiplication are defined pointwise, and the addition and multiplication of real numbers satisfy all of the field axioms. To show that R is not an integral domain, let $f(x) = 0$ for $x < 0$ and $f(x) = x$ for $x \geq 0$, and let $g(x) = 0$ for $x \geq 0$ and $g(x) = x$ for $x < 0$. Then $f(x)g(x) = 0$ for all x, which shows that $f(x)g(x)$ is the zero function.

We next consider several noncommutative examples.

Example 1.1.2 (2×2 Matrices over a field)

We assume that the reader is familiar with the ring $\mathrm{M}_2(F)$ of 2×2 matrices with entries in the field F. A review of the properties of matrices is provided in Section A.3 of the appendix. We recall that for matrices $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ in $\mathrm{M}_2(F)$, their product is given by the following matrix.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

The matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is invertible if and only if its determinant $\det(A) = a_{11}a_{22} - a_{12}a_{21}$ is nonzero. This follows from the next equation, found by multiplying A by its adjoint $\mathrm{adj}(A)$.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} = \begin{bmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix}$$

An interesting particular case is that of a lower triangular matrix with nonzero entries on the main diagonal. For a matrix of this type we have the following formula.

$$\begin{bmatrix} a & 0 \\ b & c \end{bmatrix}^{-1} = \begin{bmatrix} a^{-1} & 0 \\ -a^{-1}bc^{-1} & c^{-1} \end{bmatrix}$$

We recall the definition of a linear transformation. If V and W are vector spaces over the field F, then a *linear transformation* from V to W is a function $f : V \to W$ such that $f(c_1v_1+c_2v_2) = c_1f(v_1)+c_1f(v_2)$ for all vectors $v_1, v_2 \in V$ and all scalars $c_1, c_2 \in F$. For a vector space V of dimension n, there is a one-to-one correspondence between $n \times n$ matrices with entries in F and linear transformations from V into V. (See Theorem A.4.1 of the appendix.)

To give the most general such example, we need to recall the definition of a homomorphism of abelian groups. If A and B are abelian groups, with the operation denoted additively, then a *group homomorphism* from A to B is a function $f : A \to B$ such that $f(x_1 + x_2) = f(x_1) + f(x_2)$, for all $x_1, x_2 \in A$. If the group homomorphism maps A into A, it is called an *endomorphism* of A. We will use the notation $\mathrm{End}(A)$ for the set of all endomorphisms of A.

We will show in Proposition 1.2.8 that rings of the form $\mathrm{End}(A)$ are the generic rings, in the same sense that permutation groups are the generic examples of groups.

Example 1.1.3 (Endomorphisms of an abelian group)

Let A be an abelian group, with its operation denoted by $+$. We define addition and multiplication of elements of $\mathrm{End}(A)$ as follows: if $f, g \in \mathrm{End}(A)$, then

$$[f + g](x) = f(x) + g(x) \quad \text{and} \quad [f \cdot g](x) = f(g(x))$$

for all $x \in A$. Thus we are using pointwise addition and composition of functions as the two operations.

Since A is an abelian group, it is easy to check that addition of functions is an associative, commutative, binary operation. The identity element is the zero function defined by $f(x) = 0$ for all

$x \in A$, where 0 is the identity of A. The additive inverse of an element $f \in \mathrm{End}(A)$ is defined by $[-f](x) = -(f(x))$ for all $x \in A$.

Multiplication is well-defined since the composition of two group homomorphisms is again a group homomorphism. The associative law holds for composition of functions, but, in general, the commutative law does not hold. There is a multiplicative identity element, given by the identity function 1_A.

The most interesting laws to check are the two distributive laws. (We must check both since multiplication is not necessarily commutative.) If $f, g, h \in \mathrm{End}(A)$, then $(f + g) \cdot h = (f \cdot h) + (g \cdot h)$ since

$$\begin{aligned} ((f+g)\cdot h)(x) &= (f+g)(h(x)) \\ &= f(h(x)) + g(h(x)) \\ &= (f\cdot h)(x) + (g\cdot h)(x) \\ &= ((f\cdot h) + (g\cdot h))(x) \end{aligned}$$

for all $x \in A$. Furthermore, $h \cdot (f + g) = (h \cdot f) + (h \cdot g)$ since

$$\begin{aligned} [h\cdot(f+g)](x) &= h(f(x)+g(x)) \\ &= h(f(x)) + h(g(x)) \\ &= [(h\cdot f) + (h\cdot g)](x) \end{aligned}$$

for all $x \in A$. The last argument is the only place we need to use the fact that h is a group homomorphism. This completes the proof that $\mathrm{End}(A)$ is a ring.

The list of axioms for a ring is rather exhausting to check. In many cases we will see that the set we are interested in is a subset of a known ring. If the operations on the subset are the same as those of the known ring, then only a few of the axioms need to be checked. To formalize this idea, we need the following definition.

Definition 1.1.3 *A subset S of a ring R is called a* subring *of R if S is a ring under the operations of R, and the multiplicative identity of S coincides with that of R.*

Let F and E be fields. If F is a subring of E, according to the above definition, then we usually say (more precisely) that F is a *subfield* of E, or that E is an *extension field* of F. Of course, there may be subrings of fields that are not necessarily subfields.

Any subring is a subgroup of the underlying abelian group of the larger ring, so the two rings must have the same zero element. If F is any field contained in E, then the set of nonzero elements of F is a subgroup of the multiplicative group of nonzero elements of E, and so the multiplicative identity elements of F and E must coincide. Thus F must be a subfield of E.

Suppose that D is a subring of a field F, and that $ab = 0$ for some $a, b \in D$. If a is nonzero, then the definition of a field implies that there exists an element $a^{-1} \in F$ with $a \cdot a^{-1} = 1$. Multiplying both sides of the equation $ab = 0$ by a^{-1} shows that $b = 0$. We conclude that D is an integral domain. In Section 1.3 we will show that the converse holds: if D is any integral domain, then it is possible to construct a field F in which D can be considered to be a subring. Thus integral domains can be characterized as subrings of fields.

The next proposition is useful in constructing subrings.

Proposition 1.1.4 *Let S be a ring, and let R be a nonempty subset of S. Then R is a subring of S if and only if*

(i) *R is closed under the addition and multiplication of S,*

(ii) *if $a \in R$, then $-a \in R$,*

(iii) *the multiplicative identity of S belongs to R.*

Proof. If R is a subring, then the closure axioms must certainly hold. Condition (ii) holds since R is a subgroup of S under addition. Condition (iii) holds by definition of a subring.

Conversely, suppose that the given conditions hold. The first condition shows that condition (i) of Definition 1.1.1 is satisfied. Conditions (i)–(iii) of Definition 1.1.1 are inherited from S. Finally, since R is nonempty, it contains some element, say $a \in R$. Then $-a \in R$, so $0 = a + (-a) \in R$ since R is closed under addition. By assumption the multiplicative identity 1 of S belongs to R, and this serves as a multiplicative identity for R. Since identity elements are unique, this shows that 1 is the multiplicative identity for R. Thus conditions (iv) and (v) of Definition 1.1.1 are also satisfied. □

The first example of a subring should be to consider $\mathbf{Z}$ as a subset of $\mathbf{Q}$. Another interesting subring is found in the field $\mathbf{C}$ of complex numbers. The set $\mathbf{Z}[\mathrm{i}]$ is called the ring of *Gaussian integers*. It is by definition the set of complex numbers of the form $m + n\,\mathrm{i}$, where $m, n \in \mathbf{Z}$. Since

$$(m + n\,\mathrm{i}) + (r + s\,\mathrm{i}) = (m + r) + (n + s)\,\mathrm{i}$$

and

$$(m + n\,\mathrm{i})(r + s\,\mathrm{i}) = (mr - ns) + (nr + ms)\,\mathrm{i}\,,$$

for all $m, n, r, s \in \mathbf{Z}$, we see that $\mathbf{Z}[\mathrm{i}]$ is closed under addition and multiplication of complex numbers. Since the negative of any element in $\mathbf{Z}[\mathrm{i}]$ again has the

correct form, as does $1 = 1 + 0\,\mathrm{i}$, it follows that $\mathbf{Z}[\mathrm{i}]$ is a commutative ring by Proposition 1.1.4. Since it is a subring of the field of complex numbers, it is an integral domain.

Example 1.1.4 (Some subrings of matrix rings)

The ring $\mathrm{M}_2(\mathbf{C})$ of 2×2 matrices with complex entries is interesting in its own right, but in this example we consider two of its subrings.

Let R be the subring

$$\mathrm{M}_2(\mathbf{Z}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| a, b, c, d \in \mathbf{Z} \right\}$$

of $\mathrm{M}_2(\mathbf{C})$ consisting of all matrices with integer entries. This provides an interesting example, with a much more complex structure than $\mathrm{M}_2(\mathbf{C})$. For example, fewer matrices are invertible, since a matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in R has a multiplicative inverse if and only if $ad - bc = \pm 1$. Even though the entries come from an integral domain, there are many examples of nonzero matrices $A, B \in \mathrm{M}_2(\mathbf{Z})$ with $AB = 0$, such as $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

Let S be the subring of R consisting of all matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ such that $b = 0$. The ring S is called the ring of *lower triangular* matrices with entries in $\mathbf{Z}$. It will provide an interesting source of examples. For instance, in S the matrix $\begin{bmatrix} a & 0 \\ c & d \end{bmatrix}$ is invertible if and only if $a = \pm 1$ and $d = \pm 1$.

We need to define several classes of elements.

Definition 1.1.5 *Let R be a ring, and let $a \in R$.*

(a) *If $ab = 0$ for some nonzero element $b \in R$, then a is called a* left zero divisor. *If $ba = 0$ for some nonzero element $b \in R$, then a is called a* right zero divisor.

(b) *If a is neither a left zero divisor nor a right zero divisor, then a is called a* regular *element.*

(c) *The element $a \in R$ is said to be* invertible *if there exists an element $b \in R$ such that $ab = 1$ and $ba = 1$. The element a is also called a* unit *of R, and its multiplicative inverse is usually denoted by a^{-1}.*

(d) *The set of all units of R will be denoted by $R^\times$.*

For any ring R, the set $R^\times$ of units of R is a group under the multiplication of R. To see this, if $a, b \in R^\times$, then a^{-1} and b^{-1} exist in R, and so $ab \in R^\times$ since $(ab)(b^{-1}a^{-1}) = 1$ and $(b^{-1}a^{-1})(ab) = 1$. The element $1 \in R$ is an identity element for $R^\times$, and $a^{-1} \in R^\times$ since $(a^{-1})^{-1} = a$. The associative law for multiplication holds because R is assumed to be a ring.

Now suppose that $1 \neq 0$ in R. Since $0 \cdot b = 0$ for all $b \in R$, it is impossible for 0 to have a multiplicative inverse. Furthermore, if $a \in R$ and $ab = 0$ for some nonzero element $b \in R$, then a cannot have a multiplicative inverse since multiplying both sides of the equation by the inverse of a (if it existed) would show that $b = 0$. Thus no zero divisor is a unit in R.

An element a of a ring R is called *nilpotent* if $a^n = 0$ for some positive integer n. For example, any strictly lower triangular matrix A in $\mathrm{M}_n(F)$ is nilpotent, with $A^n = 0$. Our next observation gives an interesting connection between nilpotent elements and units. If a is nilpotent, say $a^n = 0$, then

$$(1 - a)(1 + a + a^2 + \cdots + a^{n-1}) = 1 - a^n = 1 .$$

Since $1 - a$ commutes with $1 + a + a^2 + \cdots + a^{n-1}$, this shows that $1 - a$ is a unit.

The next example provides a construction in which it is informative to consider zero divisors and units.

Example 1.1.5 (The direct sum of rings)

Let R_1, R_2, ..., R_n be rings. The set of n-tuples $(r_1, r_2, \ldots, r_n)$ such that $r_i \in R$ for all i is a group under componentwise addition. It is clear that componentwise multiplication is associative, and $(1, 1, \ldots, 1)$ serves as a multiplicative identity. It is easy to check that the distributive laws hold, since they hold in each component. This leads to the following definition.

The *direct sum* of the rings R_1, R_2, ..., R_n is defined to be the set

$$R_1 \oplus \cdots \oplus R_n = \{(r_1, r_2, \ldots, r_n) \mid r_i \in R \text{ for all } i\} ,$$

with the operations of componentwise addition and multiplication.

If R_1, R_2, ..., R_n are nontrivial rings, then zero divisors are easily found in the direct sum $R_1 \oplus \cdots \oplus R_n$. For example,

$$(1, 0, \ldots, 0)(0, 1, \ldots, 0) = (0, 0, \ldots, 0) .$$

An element $(a_1, \ldots, a_n) \in R_1 \oplus \cdots \oplus R_n$ is a unit if and only if each component is a unit. This can be shown by observing that $(a_1, \ldots, a_n)(x_1, \ldots, x_n) = (1, \ldots, 1)$ and $(x_1, \ldots, x_n)(a_1, \ldots, a_n) = (1, \ldots, 1)$ if and only if $x_i = a_i^{-1}$ for $1 \leq i \leq n$.

In the next definition we give the noncommutative analogs of integral domains and fields.

Definition 1.1.6 *Let R be a ring in which $1 \neq 0$.*

(a) *We say that R is a* noncommutative domain *if R is a noncommutative ring and $ab = 0$ implies $a = 0$ or $b = 0$, for all $a, b \in R$.*

(b) *We say that R is a* division ring *or* skew field *if every nonzero element of R is invertible.*

Strictly speaking, any field is a skew field, but we will generally use the term 'skew field' (or 'division ring') only when there is a chance that the ring is actually noncommutative. The next example is definitely noncommutative, and is probably the most familiar example of a skew field. Note that by Wedderburn's theorem (see [4], Theorem 8.5.6), it is not possible to give a finite example of a division ring that is not a field.

Example 1.1.6 (The quaternions)

The following subset of $\mathrm{M}_2(\mathbf{C})$ is called the set of *quaternions.*

$$\mathbf{H} = \left\{ \left[\begin{array}{cc} z & w \\ -\overline{w} & \overline{z} \end{array} \right] \middle| z, w \in \mathbf{C} \right\}$$

If we represent the complex numbers z and w as $z = a + b\mathrm{i}$ and $w = c + d\mathrm{i}$, then

$$\left[\begin{array}{cc} z & w \\ -\overline{w} & \overline{z} \end{array} \right] = a \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] + b \left[\begin{array}{cc} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{array} \right] + c \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right] + d \left[\begin{array}{cc} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{array} \right].$$

If we let

$$1 = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right], \ \mathbf{i} = \left[\begin{array}{cc} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{array} \right], \ \mathbf{j} = \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right], \ \mathbf{k} = \left[\begin{array}{cc} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{array} \right],$$

then we can write

$$\mathbf{H} = \{a \cdot 1 + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \mid a, b, c, d \in \mathbf{R}\}.$$

Direct computations with the elements $\mathbf{i}, \mathbf{j}, \mathbf{k}$ show that we have the following identities:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1;$$

$$\mathbf{ij} = \mathbf{k}, \quad \mathbf{jk} = \mathbf{i}, \quad \mathbf{ki} = \mathbf{j}; \qquad \mathbf{ji} = -\mathbf{k}, \quad \mathbf{kj} = -\mathbf{i}, \quad \mathbf{ik} = -\mathbf{j}.$$

These identities show that $\mathbf{H}$ is closed under matrix addition and multiplication, and it is easy to check that we have defined a subring of $M_2(\mathbf{C})$.

The determinant of the matrix corresponding to $a\cdot 1+b\,\mathbf{i}+c\,\mathbf{j}+d\,\mathbf{k}$ is $z\overline{z}+w\overline{w}=a^2+b^2+c^2+d^2$, and this observation shows that each nonzero element of $\mathbf{H}$ has a multiplicative inverse. The full name for $\mathbf{H}$ is the *division ring of real quaternions.* The notation $\mathbf{H}$ is used in honor of Hamilton, who discovered the quaternions after about ten years of trying to construct a field using 3-tuples of real numbers. He finally realized that if he would sacrifice the commutative law and extend the multiplication to 4-tuples he could construct a division ring.

The next example has its origins in analysis. It gives an example of a noncommutative domain that is not a division ring.

Example 1.1.7 (Differential operator rings)

Consider the homogeneous linear differential equation

$$a_n(z)\frac{d^n f}{dz^n}+\cdots+a_1(z)\frac{df}{dz}+a_0(z)f=0\,,$$

where the solution $f(z)$ is a polynomial with complex coefficients, and the terms $a_i(z)$ also belong to $\mathbf{C}[z]$. The equation can be written in compact form as

$$L(f)=0\,,$$

where L is the differential operator

$$L=a_n(z)\partial^n+\cdots+a_1(z)\partial+a_0(z)\,,$$

with $\partial=d/dz$. Thus the differential operator can be thought of as a polynomial in the two indeterminates z and ∂, but in this case the indeterminates do not commute, since

$$\partial(zf(z))=f(z)+z\partial(f(z))\,,$$

yielding the identity $\partial z=1+z\partial$. Repeated use of this identity makes it possible to write the composition of two differential operators in the standard form

$$a_n(z)\partial^n+\cdots+a_1(z)\partial+a_0(z)\,,$$

and we denote the resulting ring, called the *ring of differential operators*, by $\mathbf{C}[z][\partial]$ or $\mathbf{C}[z;\partial]$.

In taking the product of the terms $z^j\partial^n$ and $z^k\partial^m$, we obtain $z^{j+k}\partial^{n+m}$, together with terms having lower degree in z or ∂. This can be shown via an inductive argument using the identity $\partial z = 1 + z\partial$. In taking the product of two arbitrary elements

$$a_n(z)\partial^n + \cdots + a_1(z)\partial + a_0(z)$$

and

$$b_m(z)\partial^m + \cdots + b_1(z)\partial + b_0(z) \ ,$$

it is not difficult to show that the leading term is $a_n(z)b_m(z)\partial^{n+m}$, and so the product of two nonzero elements must be nonzero. This implies that $\mathbf{C}[z;\partial]$ is a noncommutative domain.

This construction can be made for any field F, and it can also be generalized to include polynomials in more than one indeterminate. The construction provides interesting and important noncommutative examples.

We next define an entire class of examples, in which the construction begins with a given group and a given field. This construction provides the motivation for many of the subsequent results in the text.

Example 1.1.8 (Group rings)

Let F be a field, and let G be a finite group of order n. We assume that the identity of G is denoted by 1, and that the elements of G are $g_1 = 1, g_2, \ldots, g_n$. The *group ring* FG determined by F and G is defined to be the n-dimensional vector space over F having the elements of G as a basis.

Vector addition is used as the addition in the ring. Elements of FG can be described as sums of the form

$$\textstyle\sum_{i=1}^n c_i g_i \ ,$$

where the coefficient c_i belongs to F. With this notation, the addition in FG can be thought of as componentwise addition, similar to addition of polynomials.

To define the multiplication in FG, we begin with the basis elements $\{g_i\}_{i=1}^n$, and simply use the multiplication of G. This product is extended by linearity (that is, with repeated use of the distributive law) to linear combinations of the basis elements. With this notation, the multiplication on FG is defined by

$$(\textstyle\sum_{i=1}^n a_i g_i)(\sum_{j=1}^n b_j g_j) = \sum_{k=1}^n c_k g_k \ ,$$

where $c_k = \sum_{g_i g_j = g_k} a_i b_j$.

Note that the elements of FG are sometimes simply written in the form $\sum_{g \in G} c_g g$. With this notation, the multiplication on FG is defined by

$$(\textstyle\sum_{g \in G} a_g g)(\sum_{h \in G} b_h h) = \sum_{k \in G} c_k k,$$

where $c_k = \sum_{gh=k} a_g b_h$.

Note that each of the basis elements is invertible in FG, since they have inverses in G. On the other hand, zero divisors are also easy to find. If $g \in G$ has order $m > 1$, then $1, g, \ldots, g^{m-1}$ are distinct basis elements, and we have

$$(1-g)(1+g+\cdots+g^{m-1}) = 1 - g^m = 0\,.$$

Thus if G is a finite nonabelian group, then FG is a noncommutative ring that is not a domain.

Let R be a commutative ring. For any $a \in R$, let

$$aR = \{x \in R \mid x = ar \text{ for some } r \in R\}\,.$$

This set is nonempty, since $a \in aR$, and it is closed under addition, subtraction, and multiplication since $ar_1 \pm ar_2 = a(r_1 \pm r_2)$ and $(ar_1)(ar_2) = a(r_1 a r_2)$, for all $r_1, r_2 \in R$. We note that $1 \in aR$ if and only if a is invertible, and that is the case if and only if $aR = R$. Thus aR is almost a subring of R, except for the fact that it does not generally contain an identity element. However, it has an important additional property: if $x \in aR$ and $r \in R$, then $xr \in aR$. This property plays a crucial role in constructing factor rings, and so in that sense the notion of an ideal of a ring (as defined below) corresponds to the notion of a normal subgroup of a group. We can give a general definition of an ideal that does not require the ring R to be commutative, but our proof that aR is an ideal definitely depends on commutativity.

Definition 1.1.7 *Let R be a ring.*

(a) *A nonempty subset I of R is called an* ideal *of R if*

(i) *$a + b \in I$ for all $a, b \in I$ and*

(ii) *$ra, ar \in I$ for all $a \in I$ and $r \in R$.*

(b) *If R is a commutative ring, then the ideal*

$$aR = \{x \in R \mid x = ar \text{ for some } r \in R\}$$

is called the principal ideal *generated by a. The notation (a) will also be used.*

(c) *An integral domain in which every ideal is a principal ideal is called a* principal ideal domain.