1
BASICS

This first chapter introduces the fundamental definitions and properties of rings and modules. Our starting point is that the reader knows something of arithmetic and of linear algebra, and our explanations and examples will often invoke such knowledge.

In arithmetic, we use the fact that unique factorization holds in the ring of integers, and we also use the division algorithm and the elementary properties of residue classes. In linear algebra, we call upon the standard results on finite-dimensional vector spaces and matrices. We also take for granted that the reader is acquainted with the basic language of set theory and group theory, and that he or she is happy to carry out 'routine' verifications to confirm that an object does possess some properties as claimed.

All these topics are met in a standard undergraduate mathematics course and in many expository texts, such as [Allenby 1991] and [Higgins 1975].

1.1 RINGS

In this section, we introduce rings, ideals, residue rings and homomorphisms of rings, and we discuss the relationships between these objects. We show how to construct two types of ring: one is a field of fractions, the other, a noncommutative polynomial ring in several variables. Our illustrations and examples are provided by the ring of integers, and by matrix rings and polynomial rings (in one variable), which we assume the reader has met before.

In this text we usually prefer to work with rings that have an identity element, but we sometimes make an excursion to examine rings that do not, which we call nonunital rings.

The abstract definition of a ring was first formulated by Fraenkel in 1914 [Kleiner 1996], although the term 'ring' had been introduced previously by Hilbert. Before then, the various types of ring that we encounter later –

polynomial rings, noncommutative algebras, rings of algebraic integers – had each been considered separately. Perhaps surprisingly, the idea of an ideal is much older, since it originates in number theory, as we shall see in Chapter 5. However, the explicit distinction between left and right ideals and the formal construction of a residue ring modulo a twosided ideal first occur in the work of Emmy Noether in the 1920s.

1.1.1 The definition

Informally, a ring is a set of elements which can be added and multiplied in such a way that most of the expected rules of arithmetic are obeyed. Familiar examples of rings are the integers $\mathbb{Z}$, the rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C}$.

However, we also wish to work with rings such as the ring $M_n(\mathbb{R})$ of $n \times n$ real matrices, so we cannot assume that multiplication is commutative, that is, that $xy = yx$ always.

The formal definition is as follows. A *ring* is a nonempty set R on which there are two operations, addition and multiplication. Under addition, R must be an *abelian group*, which means that the following axioms hold.

(A1) Closure:
if $r, s \in R$, then the sum $r + s \in R$.

(A2) Associativity:
$(r + s) + t = r + (s + t)$ for all r, s and $t \in R$.

(A3) Commutativity:
$r + s = s + r$ for all $r, s \in R$.

(A4) Zero:
there is a zero element 0 in R with $r + 0 = r$ for all $r \in R$.

(A5) Negatives:
if $r \in R$, then there is a negative $-r$ with $r + (-r) = 0$.

(As usual, we write $r + (-s) = r - s$ and $(-r) + s = -r + s$.)

Under multiplication, we require that R is a *monoid*, which means that the following axioms must hold.

(M1) Closure:
if $r, s \in R$, then the product $rs \in R$.

(M2) Associativity:
$(rs)t = r(st)$ for all r, s and $t \in R$.

(M3) Identity:
there is an identity element 1 in R with $r1 = r = 1r$ for all r in R.

The addition and multiplication in a ring are related by

(D) Distributivity:
for all r, s, t and $u \in R$,

$$(r+s)t = rt + st \quad \text{and} \quad r(t+u) = rt + ru.$$

If there are several rings under consideration, we sometimes indicate the zero and identity elements of R by 0_R and 1_R respectively.

We allow the possibility that $0 = 1$. In that event we have the *trivial* or *zero* ring 0 that has only one element.

Given any ring R, the set $M_n(R)$ of all $n \times n$ matrices over R is again a ring under the usual matrix addition and multiplication. Similarly, the set $R[T]$ of polynomials $f(T) = f_0 + f_1T + \ldots + f_nT^n$, with $f_0, f_1, \ldots, f_n \in R$ and T an indeterminate, is a ring under the standard addition and multiplication of polynomials.

We assume that the reader is familiar with these constructions, at least when the coefficient ring R is $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{Z}$. They are considered in more detail in sections 2.2 and 3.2.

In everday arithmetic, one takes for granted that products can be computed in any order, that is, $rs = sr$ always. However, this property does not hold for many of the rings that we wish to consider in this text, so we make a formal definition that distinguishes the rings that do have this property.

A ring R is *commutative* if the following condition holds.

(CR) Commutativity:
$rs = sr$ for all r and s in R.

It is well known and easy to verify that the polynomial ring $R[T]$ is commutative precisely when R is commutative; for $n > 1$, the matrix ring $M_n(R)$ is not commutative except in the trivial case $R = 0$. Naturally enough, a ring is said to be *noncommutative* if it is not commutative.

1.1.2 Nonunital rings

If the axiom of the identity, (M3) above, is omitted, R is called a *nonunital ring* or a *pseudoring*. (In this case, R is a *semigroup* under multiplication.) Thus every ring is necessarily a nonunital ring. Many of the definitions and constructions that we make for rings have evident counterparts for nonunital rings, which we usually do not state explicitly. A systematic way of passing from nonunital rings to rings is indicated in Exercise 1.1.5.

Some authors extend the definition of a ring to include nonunital rings as

rings, so that a 'ring with identity' becomes a special type of ring. However, for the purposes of algebraic K-theory our definition is the more convenient.

1.1.3 Subrings

A *subring* of R is a subset S of R with the following properties:

(SR1) $0, 1 \in S$;
(SR2) if $r, s \in S$, then $r + s, -r$ and rs are also in S.

Clearly, a subring of a ring is itself a ring with the same operations of addition and multiplication.

An important example is the *centre* $Z(R)$ of R:

$$Z(R) = \{z \in R \mid zr = rz \text{ for all } r \in R\}.$$

Note that $Z(R)$ is commutative. A method for the computation of the centre of a matrix ring is indicated in Exercise 1.1.4.

Sometimes it is more natural to focus attention on the subring S, for instance when the ring R is constructed from S in some way. We then say that R is an *extension* of S.

1.1.4 Ideals

A *right ideal* of R is a subset $\mathfrak{a}$ of R which satisfies the following requirements:

(Id1) $0 \in \mathfrak{a}$;
(Id2) if $a, b \in \mathfrak{a}$, then $a + b \in \mathfrak{a}$ and $-a \in \mathfrak{a}$ also;
(Id^{r}3) if $a \in \mathfrak{a}$ and $r \in R$, then $ar \in \mathfrak{a}$ also.

A *left ideal* has instead

(Id$^\ell$3) if $a \in \mathfrak{a}$ and $r \in R$, then $ra \in \mathfrak{a}$ also.

If $\mathfrak{a}$ is at the same time a left and a right ideal, we call it simply an *ideal* of R, although we sometimes refer to it as a *twosided ideal* if we need to avoid ambiguity.

The ring R is always an ideal of itself. We say that the ideal $\mathfrak{a}$ is *proper* if $\mathfrak{a} \subset R$, where we use the symbol $\subset$ to denote strict inclusion, that is, $\mathfrak{a} \subseteq R$ and $\mathfrak{a} \neq R$. At the other extreme, $\{0\}$ is always an ideal, the *zero ideal* of R, which we usually denote by 0. Of course, $0 \subseteq \mathfrak{a}$ for any ideal $\mathfrak{a}$.

Observe that any left or right ideal $\mathfrak{a}$ of R is a nonunital ring – one might call it a sub-nonunital-ring or nonunital subring of R – but that $\mathfrak{a}$ will not be a subring of R unless $\mathfrak{a} = R$.

If $\mathfrak{a}$ and $\mathfrak{b}$ are each right ideals of R, then evidently their *sum*

$$\mathfrak{a} + \mathfrak{b} = \{a + b \mid a \in \mathfrak{a},\ b \in \mathfrak{b}\}$$

is also a right ideal of R. Similarly, if $\mathfrak{a}$ and $\mathfrak{b}$ are both left ideals, then so too is $\mathfrak{a} + \mathfrak{b}$. Moreover, the procedure can obviously be iterated (by associativity (A2)) to give the definition of a finite sum of ideals $\mathfrak{a}_1 + \mathfrak{a}_2 + \cdots + \mathfrak{a}_n$.

1.1.5 Generators

A convenient method of defining an ideal is to specify a set of generators. Given a subset $\{r_i \mid i \in I\}$ of R, where I is some index set, possibly infinite, we say that $r_i = 0$ for *almost all* i, or for *all except a finite set of indices*, if the set of indices i with $r_i \neq 0$ is finite.

Let $X = \{x_i \mid i \in I\}$ be a subset of R. Then the right ideal $\mathfrak{a}$ *generated* by X is the set of all expressions

$$\sum_{i \in I} x_i r_i,$$

where $r_i = 0$ for all except a finite set of indices; X is then called a set of *generators* for $\mathfrak{a}$.

When $X = \{x_1, \ldots, x_n\}$ is finite, we write

$$\mathfrak{a} = x_1 R + \cdots + x_n R;$$

if $X = \{x\}$, then $\mathfrak{a} = xR$ is the *principal right* ideal generated by x. The left ideal generated by X is defined in a similar way.

When R is commutative, we use the notation $(x_1, \ldots, x_n)$ for the ideal generated by $\{x_1, \ldots, x_n\}$. Although this notation is the same as that for a sequence, the context should prevent any confusion.

Suppose that both $\mathfrak{a}$ and $\mathfrak{b}$ are right ideals of R. The *product* $\mathfrak{ab}$ is defined to be the right ideal generated by all products ab with $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$. The product of a pair of left ideals or a pair of twosided ideals is defined similarly.

Suppose that $\mathfrak{a}$ is a twosided ideal. If $\mathfrak{b}$ is a right ideal, then $\mathfrak{ab}$ is a twosided ideal; on the other hand, if $\mathfrak{b}$ is a left ideal, then $\mathfrak{ab}$ is a left ideal which need not be twosided.

1.1.6 Homomorphisms

Let R and S be rings with zero elements 0_R and 0_S and identity elements 1_R and 1_S respectively. A *ring homomorphism* $f : R \to S$ is a map from R to S with the properties that

(RH1) for all r and s in R,

$$f(r+s) = fr + fs \text{ and } f(rs) = f(r)f(s),$$

(RH2) $f1_R = 1_S$.

The facts that $f(-r) = -f(r)$ and $f(0_R) = 0_S$ are consequences of (RH1), but (RH2) is not. This follows from the observation that the obvious map from the zero ring 0 to R is not a ring homomorphism (unless R itself happens to be 0).

Note that there is a unique homomorphism from any ring R to the zero ring 0. In the other direction, there is, for any ring, a unique ring homomorphism, the *characteristic homomorphism*

$$\chi : \mathbb{Z} \to R,$$

given by $\chi(a) = a1_R$, where 1_R is the identity element of R. Since an integer a need not be a member of the ring R, we should say what we mean by $a1_R$. We take $0_{\mathbb{Z}}1_R = 0_R$ and $1_{\mathbb{Z}}1_R = 1_R$. Then, for $a > 1$, we make the inductive definition $a1_R = (a-1)1_R + 1_R$, and for $a < 0$ we put $a1_R = -(-a)1_R$. The fact that χ is a homomorphism can now be verified by an induction argument.

Given a homomorphism f, we associate with it its *kernel*

$$\operatorname{Ker} f = \{r \in R \mid fr = 0\}$$

and its *image*

$$\operatorname{Im} f = \{s \in S \mid s = fr \text{ for some } r \in R\}.$$

Then $\operatorname{Ker} f$ is an ideal in R and $\operatorname{Im} f$ is a subring of S.

An argument familiar from elementary linear algebra shows that f is injective precisely when $\operatorname{Ker} f = 0$; it is a tautology to say that f is surjective if and only if $\operatorname{Im} f = S$.

A homomorphism $f : R \to S$ of rings is an *isomorphism* if has an *inverse*, that is, there is a ring homomorphism $g : S \to R$ such that $fg = id_S$ and $gf = id_R$, where id_R is the identity map on R and id_S is the identity map on S. It is not hard to see that f is an isomorphism precisely when it is both injective and surjective. We then write $R \cong S$.

Here is a result, used in [BK: CM], which shows how a homomorphism into a ring may be used to promote an abelian group to a ring. We give the proof in some detail as it is our first.

1.1.7 Lemma

Let $(S, +)$ be an additive group, let T be a ring, and suppose that $\theta : S \to T$ is an injective group homomorphism whose image is a subring of T. Then

there is a unique multiplication $\cdot$ on S which makes $(S,+,\cdot)$ a ring and θ a ring homomorphism.

Proof

Since $\operatorname{Im}\theta$ is a subring of T, for any s_1, s_2 in S we have $\theta(s_1)\theta(s_2) \in \operatorname{Im}\theta$. Because θ is injective, we may define

$$s_1 \cdot s_2 = \theta^{-1}(\theta(s_1)\theta(s_2)),$$

since this element is uniquely determined. So Axiom (M1) holds. Evidently,

$$\begin{aligned}\theta((s_1\cdot s_2)\cdot s_3) = \quad \theta(s_1\cdot s_2)\theta(s_3) \quad &= (\theta(s_1)\theta(s_2))\theta(s_3) \\ = \theta(s_1)(\theta(s_2)\theta(s_3)) = \quad \theta(s_1)\theta(s_2\cdot s_3) \quad &= \quad \theta(s_1\cdot(s_2\cdot s_3)),\end{aligned}$$

so that Axiom (M2) holds in S because θ is injective. Likewise Axiom (D) holds in S. There remains Axiom (M3). Now $1_T \in \operatorname{Im}\theta$, so let $1_S = \theta^{-1}(1_T)$. Then again (M3) in S may be deduced from its counterpart in T. Hence $(S,+,\cdot)$ is a ring. By construction, θ is a ring homomorphism. If $*$ is another ring multiplication making θ a ring homomorphism, then

$$\theta(s_1 * s_2) = \theta(s_1)\theta(s_2) = \theta(s_1\cdot s_2),$$

so that, by injectivity again, $s_1 * s_2 = s_1 \cdot s_2$. □

1.1.8 Residue rings

Given an ideal $\mathfrak{a}$ of a ring R, we can construct the *residue ring* $R/\mathfrak{a}$ of R modulo $\mathfrak{a}$. (The residue ring is also called the *quotient* or *factor* ring in some texts.) The definition goes as follows.

We say that $r, s \in R$ are *congruent modulo* $\mathfrak{a}$ if $r - s \in \mathfrak{a}$; this is sometimes denoted by $r \equiv s \pmod{\mathfrak{a}}$. Congruence is easily seen to be an equivalence relation on R, and the equivalence class of an element $r \in R$ is called its *residue class* or *congruence class*.

The residue class of r is denoted $\overline{r}$, so that $\overline{r} = \{r + x \mid x \in \mathfrak{a}\}$, and the residue ring $R/\mathfrak{a}$ is defined to be the set of all such classes. Addition and multiplication in $R/\mathfrak{a}$ are given by

$$\overline{r} + \overline{s} = \overline{r+s} \text{ and } \overline{r}\cdot\overline{s} = \overline{rs};$$

then $R/\mathfrak{a}$ is a ring with zero $\overline{0}$ and identity $\overline{1}$.

The *canonical* or *standard* ring homomorphism $\pi : R \to R/\mathfrak{a}$ is defined simply by $\pi r = \overline{r}$. It is not hard to verify that π is a surjective ring homomorphism, with $\operatorname{Ker}\pi = \mathfrak{a}$.

The basic example of the construction of a residue ring, which also explains

the name, occurs when we take R to be the ring of integers $\mathbb{Z}$ and $\mathfrak{a} = n\mathbb{Z}$ for some $n > 0$.

It is wellknown that the *division algorithm* holds in $\mathbb{Z}$: any integer a can be written in the form $a = qn + r$ with $0 \leq r < n$. The integer r is the *residue* (or *remainder*) and q the *quotient*. Thus the residue ring $\mathbb{Z}/n\mathbb{Z}$ consists of the residue classes

$$\overline{0}, \overline{1}, \ldots, \overline{n-1}.$$

We note the following result, which is typical of a class of very useful lemmas that we often use without comment.

1.1.9 The Induced Mapping Theorem for Rings

Let $f : R \to S$ be a ring homomorphism. Then there is a unique injective ring homomorphism $\overline{f} : R/\operatorname{Ker} f \to S$ so that $\overline{f}\pi = f$.

Proof

It must be that $\overline{f}(\overline{r}) = fr$. Note that $\overline{f}$ is often called the *induced* homomorphism. □

1.1.10 The characteristic

As an illustration, we show how to define the characteristic of a ring R and the prime subring of R.

A familiar argument based on the division algorithm shows that any ideal $\mathfrak{a}$ of $\mathbb{Z}$ is principal, of the form $\mathfrak{a} = a\mathbb{Z}$, where a is uniquely defined as the least positive integer belonging to $\mathfrak{a}$ if $\mathfrak{a} \neq 0$, and $a = 0$ if $\mathfrak{a} = 0$. (This argument is given in a much more general context in (3.2.10) below.)

Now, for an arbitrary ring R, let $\chi : \mathbb{Z} \to R$, $\chi(a) = a1_R$, be the unique homomorphism from $\mathbb{Z}$ to R, and write $\operatorname{Ker}\chi = c\mathbb{Z}$ with $c \geq 0$. Then c is called the *characteristic* of R.

The *prime ring* of R is the subring $\chi\mathbb{Z}$ of R. If R has characteristic 0, then χ is an injective ring homomorphism and $\chi\mathbb{Z} \cong \mathbb{Z}$. If R has characteristic $c > 0$, then the Induced Mapping Theorem shows that $\chi\mathbb{Z}$ is isomorphic to the residue ring $\mathbb{Z}/c\mathbb{Z}$.

1.1.11 Units

An element u of the ring R is said to be a *unit* of R or an *invertible element* of R if

$$uv = 1 = vu$$

for some v in R. The element v is then the unique *inverse* of u.

The set of all units in R will be denoted by $U(R)$. Clearly, $U(R)$ is a group under multiplication, called the *unit group*. In particular, $U(0)$ is the trivial group.

If $U(R) = \{r \in R \mid r \neq 0\}$, then R is a *division ring* or *skew field*; if also R is commutative, then R is a *field*. Familiar examples of fields are $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$, and the reader should have no difficulty in verifying that the residue ring $\mathbb{Z}/n\mathbb{Z}$, $n > 0$, is a field precisely when the integer n is a prime number. It is also straightforward to show that if the ring R is a field or division ring, then it must have characteristic either 0 or a prime. An example of a division ring is given in Exercise 2.2.3 below.

One particular type of unit group plays a fundamental role in K-theory. For any ring R and natural number n, the unit group $U(M_n(R))$ of the matrix ring $M_n(R)$ is called the *general linear group* of *degree* n, and is written $\mathrm{GL}_n(R)$.

1.1.12 Constructing the field of fractions

It is a very useful fact that a certain type of commutative ring, namely one that is a domain, can be embedded in a field. First, we make a formal definition. A nontrivial ring $\mathcal{O}$ (not necessarily commutative) is called a *domain* if the following holds.

(Dom) If $r, s \in \mathcal{O}$ and $rs = 0$, then either $r = 0$ or $s = 0$.

The terms *integral domain* and *entire ring* are sometimes used instead.

Suppose now that $\mathcal{O}$ is a commutative domain. We construct a field $\mathcal{K}$ in which every nonzero element r of $\mathcal{O}$ has an inverse $1/r$, and further any element of $\mathcal{K}$ can be written in the form r/s for $r, s \in \mathcal{O}$. Naturally enough, $\mathcal{K}$ is called the *field of fractions* of $\mathcal{O}$. The technique is exactly the same as that used to manufacture the rational numbers $\mathbb{Q}$ from the ring of integers $\mathbb{Z}$.

Let $\Sigma = \mathcal{O}\backslash\{0\}$ be the set of nonzero elements in $\mathcal{O}$. We introduce a relation $\sim$ on the set of pairs $(a, s) \in \mathcal{O} \times \Sigma$ by stipulating that $(a, s) \sim (a', s')$ if and only if there are elements u and u' in Σ with $au = a'u'$ and $su = s'u'$. It is easy to verify that this relation is an equivalence relation.

The fraction a/s is defined to be the equivalence class of (a, s) under this relation and $\mathcal{K}$ is the set of equivalence classes; thus $a/s = a'/s'$ if and only if $ax = a'x'$ and $sx = s'x' \in \Sigma$ for some x and x' in Σ.

We define addition by

$$a/s + b/t = (at + bs)/st,$$

and multiplication by

$$(a/s)(b/t) = ab/st.$$

Another routine check shows that these rules are well-defined and make $\mathcal{K}$ into a ring with zero element $0/1$ and identity $1/1$.

Furthermore, $a/1 = 0$ only if $a = 0$, so that we can identify $\mathcal{O}$ as the subring of $\mathcal{K}$ consisting of all elements of the form $a/1$.

Then the identity $r/r = 1/1$ holds for all nonzero r in $\mathcal{O}$, which confirms that r has an inverse in $\mathcal{K}$, and it is easy to see that $\mathcal{K}$ is a field.

1.1.13 Noncommutative polynomials

Many examples of rings arise from noncommutative rings of polynomials in several variables, so it will be helpful to have a brief outline of their construction. (A more detailed construction of rings of polynomials in one variable is given in section 3.2.)

Let A be a ring, which is referred to in this setting as the *coefficient ring*, and let $\{X_1, \ldots, X_k\}$ be a set of 'variables'. A *monomial* in $X_1, \ldots, X_k$ is any expression of the form

$$X_{h(1)}^{n(1)} \cdots X_{h(\ell)}^{n(\ell)}, \quad \ell \geq 0,$$

where

$$h(1), \ldots, h(\ell) \in \{1, \ldots, k\}$$

and

$$n(1), \ldots, n(\ell) \geq 1.$$

The case $\ell = 0$ is to be interpreted as giving an identity element 1. Note that two monomials are the same only if they have the same factors to the same exponents and in the same order, that is,

$$X_{h(1)}^{n(1)} \cdots X_{h(\ell)}^{n(\ell)} = X_{g(1)}^{p(1)} \cdots X_{g(m)}^{p(m)}$$

if and only if

$$\ell = m, \text{ and } n(i) = p(i) \text{ and } h(i) = g(i) \text{ for all } i.$$

For example, $X_1^2 X_2$, $X_1 X_2 X_1$ and $X_2 X_1^2$ are all different.

The *total degree* of a monomial is the sum

$$n(1) + \cdots + n(\ell).$$