STANDARD DEFINITIONS CONCERNING RINGS

KEITH CONRAD

1. Introduction

The concept of a ring generalizes systems of numbers and of functions that can be added and multiplied.

Definition 1.1. A *ring* is a set R equipped with two operations $+$ (addition) and $\times$ (multiplication) such that R is an abelian group under addition (with identity denoted 0 and the inverse of a denoted $-a$), while multiplication is associative with an identity element 1. Finally, multiplication distributes over addition: $x(y+z) = xy+xz$ and $(x+y)z = xz+yz$ for all x, y, and z in R.

This says a ring is a commutative group under addition, it is a "group without inverses" under multiplication, and multiplication distributes over addition. Examples of rings are $\mathbf{Z}$, $\mathbf{Q}$, all functions $\mathbf{R} \to \mathbf{R}$ with pointwise addition and multiplication, and $\mathrm{M}_2(\mathbf{R})$ – the latter being a noncommutative ring – but $2\mathbf{Z}$ is *not* a ring since it does not have a multiplicative identity.

Some abstract algebra books do not insist rings have a multiplicative identity, leading to the result that $2\mathbf{Z}$ is considered a subring of $\mathbf{Z}$. This is really, really bad. Below we will give the correct definitions of *subring*, *ring homomorphism*, and *ideal*. Our definitions will be the right ones even in the case of noncommutative rings, but little will be lost if you try to get your bearings just on the commutative case. In an appendix we will discuss what "rings without a multiplicative identity" should be called.

2. Subrings

Definition 2.1. A *subring* of a ring R is a subset $R' \subset R$ that is a ring under the same $+$ and $\times$ as R and **shares the same multiplicative identity**.

Example 2.2. The ring $\mathbf{Z}$ is a subring of $\mathbf{Q}$. The ring $\mathbf{Z}/(m)$ for $m > 0$ has no subrings other than itself, since 1 additively generates all of $\mathbf{Z}/(m)$, so a subring contains 1 and then contains everything. The same argument (using $m = 0$) shows $\mathbf{Z}$ contains no subrings other than itself.

It might seem odd to insist in the definition of a subring that it has the same multiplicative identity as the original ring. Shouldn't that follow from the rest of the definition? After all, a subgroup of a group is defined to be a subset that is a group for the same operation, and its identity element can be proved to be the identity for the original group (and inverses for the subgroup are therefore the same as for the original group). But the proof of this uses cancellation in the group law, and in a ring we might not have cancellation for multiplication. This is all made clearer by seeing an actual example.

Example 2.3. In $\mathbf{Z}/(6)$, the subset $\{0, 3\}$ is closed under addition and multiplication since $3^2 = 9 = 3$. So $\{0, 3\}$ is a subset of $\mathbf{Z}/(6)$ "with a ring structure" having multiplicative

identity 3. That is not the multiplicative identity for $\mathbf{Z}/(6)$, so we do not consider $\{0, 3\}$ to be a subring of $\mathbf{Z}/(6)$.

Remark 2.4. If the ring R has cancellation for multiplication (that is, $xz = yz \Rightarrow x = y$ when $z \neq 0$) then a subset of R "with a ring structure" other than $\{0\}$ has to have the same multiplicative identity as R (and thus is a subring) because if x is the multiplicative identity in a subset "with a ring structure" then the equation $x^2 = x$ is satisfied, which is the same as $x \cdot x = x \cdot 1$, forcing $x = 1$ if $x \neq 0$. Thus for rings with cancellation, the constraint on a nonzero subset that it have the same multiplicative identity as the whole ring is automatic from the other properties of a subring.

You might be thinking: what is the big fuss about subrings having the same identity for multiplication? One reason for wanting this has to do with invertible elements. An element $x \in R$ is called a *unit* if it has a 2-sided inverse: $xy = yx = 1$ for some $y \in R$. The set of all units forms a group, denoted $R^\times$. For example, $\mathbf{R}^\times = \mathbf{R} - \{0\}$, $\mathbf{Z}^\times = \{\pm 1\}$, and $\mathrm{M}_n(\mathbf{R})^\times = \mathrm{GL}_n(\mathbf{R}) = \{A \in \mathrm{M}_n(\mathbf{R}) : \det A \neq 0\}$.

Theorem 2.5. *If R is a ring and R' is a subring then $R'^\times$ is a subgroup of $R^\times$.*

Proof. Let 1 be the multiplicative identity in R, so it is also the multiplicative identity in R'. Since R' has the same multiplicative identity as R, if $x \in R'^\times$ then $xy = yx = 1$ for some $y \in R'$, so $x \in R^\times$ and the inverse of x in R' is also its inverse in R. We have shown $R'^\times$ is a subset of $R^\times$. Since the group law (multiplication) and inversion in R' are the same as in R, $R'^\times$ is a subgroup of $R^\times$. □

Example 2.6. We return to the nonexample of $\{0, 3\}$ in $\mathbf{Z}/(6)$. As a subset "with a ring structure," $\{0, 3\}$ has multiplicative identity element 3, which is not a unit in $\mathbf{Z}/(6)$. So the one unit in the "ring that's not a subring" $\{0, 3\}$ is not a unit in $\mathbf{Z}/(6)$.

It would be terrible if the units in a subring were not units in the larger ring, and insisting that subrings have the same multiplicative identity as the whole ring means this weirdness will not happen: units of a subring are units of the larger ring.

3. Ring homomorphisms

Definition 3.1. If R and S are rings, a *ring homomorphism* $f\colon R \to S$ is a function that preserves addition, multiplication, and the multiplicative identity: $f(x + y) = f(x) + f(y)$ and $f(xy) = f(x)f(y)$ for all x and y in R, and $f(1) = 1$.

The last condition, that $f(1) = 1$, is admittedly an awkward part of the definition, since we don't require in the definition that $f(0) = 0$ too. However, it is automatic that $f(0) = 0$ because f is an additive group homomorphism, and group homomorphisms always preserve the identity. But a ring is not a group under multiplication (except for the zero ring), and if we don't insist that $f(1) = 1$ as part of a ring homomorphism then weird things can happen. Consider the next example, which builds on the previous one.

Example 3.2. Let $f\colon \mathbf{Z}/(6) \to \mathbf{Z}/(6)$ by $f(x) = 3x$. Since $3^2 = 3$ in $\mathbf{Z}/(6)$, we have $f(x) + f(y) = 3x + 3y = 3(x + y) = f(x + y)$ and also (the key point) $f(x)f(y) = 3x \cdot 3y = 3^2xy = 3xy = f(xy)$. Thus f is additive and multiplicative, but $f(1) \neq 1$. We do *not* want to consider f to be a ring homomorphism, and the condition that $f(1) = 1$ rules out this example.

The *only* ring homomorphism $\mathbf{Z}/(6) \to \mathbf{Z}/(6)$ is the identity function: once 1 goes to 1 everything else is fixed too by additivity.

Here is a result involving units that would break down if a ring homomorphism did not preserve the multiplicative identities.

Theorem 3.3. *Let $f\colon R \to S$ be a ring homomorphism. Then $f(R^\times) \subset S^\times$ and the function $f\colon R^\times \to S^\times$ is a group homomorphism.*

Proof. If $xy = yx = 1$ in R then applying f gives us $f(x)f(y) = f(y)f(x) = f(1) = 1$, so f sends units in R to units in S. Since f is multiplicative, it is a group homomorphism from $R^\times$ to $S^\times$. □

Theorem 3.4. *If R' is a subring of R then the inclusion mapping $R' \hookrightarrow R$ is a ring homomorphism.*

Proof. Easily the inclusion map sends sums to sums and products to products. The multiplicative identity goes to the multiplicative identity because R' has the same multiplicative identity as R. □

In group theory, the kernel and image of a group homomorphism are subgroups. For a ring homomorphism $f\colon R \to S$, we have the kernel $\ker f = \{x \in R : f(x) = 0\}$ and image $f(R)$. Are these subrings (of R and S respectively)?

Theorem 3.5. *Let $f\colon R \to S$ be a ring homomorphism. The image of f is a subring of S, but the kernel of f is not a subring of R unless S is the zero ring.*

Proof. From the definition of a ring homomorphism, the sum and product of f-values are f-values. The image also contains 1 since $f(1) = 1$. So the image of f is a subring of S.

The kernel of f is closed under addition and multiplication. The kernel of f is not a subring of R unless 1 is in the kernel which requires $f(1) = 0$. Since $f(1) = 1$ for any ring homomorphism, we must have $1 = 0$ in S, so S is the zero ring (the only ring in which $1 = 0$). □

This last theorem is probably why some people do not insist that rings contain 1. Kernels of ring homomorphisms have all the properties of a subring except they almost never contain the multiplicative identity. So if we *want* ring theory to mimic group theory by having kernels of ring homomorphisms be subrings, then we should not insist that subrings contain 1 (and thus perhaps not even insist that rings contain 1). Then kernels of ring homomorphisms could be called subrings. The development of ring theory, particularly for commutative rings, has shown that this is a bad idea. Kernels of group homomorphisms are special kinds of subgroups (normal subgroups), but kernels of ring homomorphisms are something *other than* subrings. What are they? That is the subject of the next section.

4. Ideals

The kernel of a ring homomorphism satisfies a stronger multiplicative condition than being a subset closed under multiplication: if $f\colon R \to S$ is a ring homomorphism and $x \in \ker f$, so $f(x) = 0$, then for *any* $r \in R$ we have $f(rx) = f(r)f(x) = f(r) \cdot 0 = 0$ and $f(xr) = f(x)f(r) = 0 \cdot f(r) = 0$, so rx and xr are in the kernel too. The kernel of f is closed under multiplication by *arbitrary* elements of the ring from either side. Contrast this with $\mathbf{Z}$ as a subring of $\mathbf{Q}$: multiplication of an integer by most elements of $\mathbf{Q}$ will not again be an integer.

Definition 4.1. An *ideal* in a ring R is an additive subgroup $I \subset R$ such that $RI \subset I$ and $IR \subset I$. That is, if $x \in I$ then $Rx \subset I$ and $xR \subset I$: all multiples of x in R lie in I.

Example 4.2. A basic example of an ideal in any *commutative* ring R is the multiples of one element: for $a \in R$, $Ra = \{ra : r \in R\}$ is an ideal in R since a sum and difference of two multiples is again a multiple and (most importantly) any multiple of a multiple is again a multiple. These ideals are called *principal ideals*. For instance, the even numbers $2\mathbf{Z}$ are a principal ideal in the ring $\mathbf{Z}$ but they are not a subring of $\mathbf{Z}$.

If R is noncommutative then this attempt to construct an ideal runs into trouble when you switch the side you multiply on. If an ideal contains a then it contains not only left multiples of a but also right multiples, and in fact multiples from both sides taken together, which is the set $\{ras : r, s \in R\}$. But this set is usually not closed under addition if R is noncommutative. So we have to take finite sums of these two-sided products, getting $r_1as_1 + \cdots + r_nas_n$ for $n \geq 1$ and $r_i, s_i \in R$. Now *that* is an ideal. Very tedious! This is why you should not try to learn about ideals first in noncommutative rings. It's too complicated. Focus on ideals in the commutative setting until you get used to them.

Example 4.3. An ideal in a commutative ring that is not of the special form Ra is the polynomials in $\mathbf{Z}[T]$ that have an even constant term: $I = \{f(T) \in \mathbf{Z}[T] : f(0) \text{ is even}\}$. Examples of elements of I are 2, T, and $T^2 + 3T + 10$. Check yourself that I is an ideal in $\mathbf{Z}[T]$. We will show by contradiction that I is not a principal ideal. Assume I is a principal ideal, so $I = \mathbf{Z}[T]f(T)$ for some $f(T)$. Since $2 \in I$, $2 = g(T)f(T)$ for some $g(T)$, so $f(T)$ has to be a constant polynomial. Write $f(T) = c$. Then $2 = g(T)c$, so $c = \pm 1$ or ± 2. Since c is in the ideal, it must be even, so $c = \pm 2$. Because $T \in I$, $T = h(T)c$ for some $h(T)$, but the left side has leading coefficient 1 and the right side has even leading coefficient. We have a contradiction, so I is not principal.

While the ideal I is not generated by a single element, it is generated by two elements. The general element of I is a linear combination of 2 and T, with coefficients in $\mathbf{Z}[T]$. We can write I symbolically as $2\mathbf{Z}[T] + T\mathbf{Z}[T]$, or as $2\mathbf{Z} + T\mathbf{Z}[T]$.

As David Rohrlich has nicely put it, ideals are "contagious for multiplication." That may help you remember their defining property when you're first working with them. I like to say ideals swallow up multiplication.

Example 4.4. An important way ideals occur in mathematics is as kernels of ring homomorphisms. Any kernel of a ring homomorphism is an ideal. Conversely, one can show any ideal in a ring can be viewed as the kernel of a suitable ring homomorphism using quotient rings (analogous to quotient groups). Thus ideals in a ring are analogous to the normal subgroups of a group: any kernel of a group homomorphism is a normal subgroup and the quotient group construction shows any normal subgroup is the kernel of some group homomorphism.

While $\mathbf{Z}$ is an additive subgroup of $\mathbf{R}$, it is not an ideal in $\mathbf{R}$ since real numbers times integers are usually not integers. Similarly, $\mathbf{Z}$ is a subgroup of $\mathbf{Q}$ but is not an ideal of $\mathbf{Q}$. More generally, a *subring* of a ring R is not an ideal of R unless it's all of R: if R' is a subring of R and also R' is an ideal of R, then since $1 \in R'$ (!) we get for all $r \in R$ that $r = r \cdot 1 \in R'$. Thus $R' = R$. So except for the whole ring, which is both a subring and ideal of itself, subrings and ideals are absolutely separate concepts.

Appendix A. Rings without identity

Having tried to explain why rings should contain a multiplicative identity (and what this implies about the correct definitions of subring and ring homomorphism), we should

admit that "ring-like" systems without a multiplicative identity do occur in mathematics, especially in analysis.

Example A.1. Consider the set of continuous functions $\mathbf{R} \to \mathbf{R}$ whose limit as $x \to \pm\infty$ is 0. Examples include $x/(x^2+1)$ and $(\sin x)e^{-|x|}$. The set of such functions is denoted $C_0(\mathbf{R})$. See Figure 1.

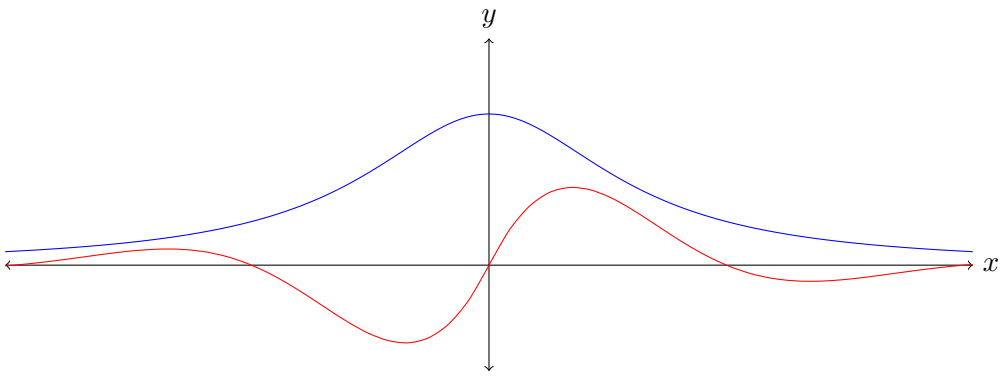

FIGURE 1. Functions in $C_0(\mathbf{R})$.

Under pointwise addition and multiplication, $C_0(\mathbf{R})$ satisfies the definition of a ring, except that it does not have a multiplicative identity. If there were a multiplicative identity it would have to be the constant function 1, which does *not* belong to $C_0(\mathbf{R})$.

Besides pointwise multiplication of functions, another important multiplicative operation on functions in analysis is *convolution* (look up the definition in any analysis book). In many important cases there is no identity for convolution, so under addition and convolution one has another example in analysis of a ring without an identity.

On account of these examples (from analysis), what can we call a "ring without identity" if we can't call it a ring? There is already an available term: a "ring without identity" is a $\mathbf{Z}$-algebra in the sense of the following definition.

Definition A.2. Let R be a commutative ring with identity. An *R-algebra* is an abelian group A that admits a multiplication $A \times A \to A$ and a scalar multiplication by R. Denoting multiplication on A by $(a,b) \mapsto ab$, and scalar multiplication of R and A by $(r,a) \mapsto ra$, the conditions on multiplication are

(1) R-bilinearity of $A \times A \to A$:
   - $a(b+c) = ab + ac$, and $(a+b)c = ac + bc$ for all a, b, and c in A,
   - $r(ab) = (ra)b = a(rb)$ for all $r \in R$ and all a and b in A,

(2) Scalar multiplication $R \times A \to A$ is a ring homomorphism $R \to \mathrm{End}(A)$:
   - $r(a+b) = ra + rb$ for all $r \in R$ and $a, b \in A$,
   - $(r+s)a = ra + sa$ for all r, s in R and all a in A,
   - $(rs)(a) = r(sa)$ for all r and s in R and a in A,
   - $1 \cdot a = a$ for all $a \in A$.

For an R-algebra A there is a distinguished ring R by which we can multiply elements of A, and R may not lie inside A. (Compare with real vector spaces, whose elements can be scaled by real numbers even though real numbers are not themselves vectors.)

Example A.3. The ring $\mathrm{M}_n(\mathbf{R})$ is an $\mathbf{R}$-algebra where we multiply a scalar and a matrix in the usual way.

Example A.4. The set of continuous functions $[0,1] \to \mathbf{R}$ is an $\mathbf{R}$-algebra under pointwise addition and multiplication.

Example A.5. The product ring $\mathbf{Z}/(5) \times \mathbf{Z}/(5)$ is a $\mathbf{Z}$-algebra, where multiplication is given by $n \cdot (x \bmod 5, y \bmod 5) = (nx \bmod 5, ny \bmod 5)$. Similarly, we can treat $\mathbf{Z}/(5) \times \mathbf{Z}/(5)$ not just as a $\mathbf{Z}$-algebra, but as a $\mathbf{Z}/(d)$-algebra when d is any multiple of 5. That a single algebra can be an algebra over more than one ring is like a feature of linear algebra: any complex vector space is also a real vector space.

If S is a ring containing a subring R then S is an R-algebra, where we use the multiplication in S to define how elements of R and S multiply together.

Example A.6. The ring $\mathbf{R}$ is a $\mathbf{Z}$-algebra and a $\mathbf{Q}$-algebra.

A "ring possibly without identity" is the same thing as an associative $\mathbf{Z}$-algebra. (Requiring $1 \cdot a = a$ forces a unique meaning on $n \cdot a$ for all $n \in \mathbf{Z}$, by writing n as a sum or difference of 1's.) Returning to Example A.1, the ring $C_0(\mathbf{R})$ with pointwise addition and multiplication is not just a $\mathbf{Z}$-algebra but also an $\mathbf{R}$-algebra (which is usually how analysts think about it). In analysis there are special types of algebras over $\mathbf{R}$ or $\mathbf{C}$, such as Banach algebras and C^*-algebras, that satisfy additional constraints coming from analysis.

We did not require R-algebras to have an associative multiplication, although many do: often $a(bc) = (ab)c$ for all a, b, and c in the algebra. Many texts in fact use the term R-algebra to mean "associative R-algebra with a multiplicative identity". An example of a nonassociative $\mathbf{R}$-algebra is $\mathbf{R}^3$ with usual addition and the cross product as multiplication: $\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) \neq (\mathbf{x} \times \mathbf{y}) \times \mathbf{z}$ in general and there is no cross product identity vector. Another nonassociative $\mathbf{R}$-algebra is $\mathrm{M}_n(\mathbf{R})$ using usual addition and the bracket operation $[A, B] = AB - BA$ as multiplication. This doesn't have an identity either: for no A is $[A, B] = B$ for all B (for instance, the matrix $[A, B]$ has trace 0, so it is not B if B doesn't have trace 0). These nonassociative algebra structures on $\mathbf{R}^3$ and $\mathrm{M}_n(\mathbf{R})$ are special cases of a *Lie algebra*.