Section 16: Rings

Let's get off of the topic of groups, which is strongly connected to number theory, and move on to another algebraic structure, where there is a closer connection to high school algebra:

Def: A *ring* is a set R with two operations, denoted $+$ and $\cdot$ (or juxtaposition) in which the following properties hold:

- $(R, +)$ is an abelian group;
- $\cdot$ is an associative operation on R; and
- the two "distributive laws" hold: For all a, b, c in R,

$$a(b+c) = ab + ac \qquad \text{and} \qquad (b+c)a = ba + ca \ .$$

If $\cdot$ is commutative, then R is a *commutative ring.* There must be an identity for $+$, denoted as usual by 0 or 0_R, and inverses for this operation are denoted with negative signs. If there is an identity for $\cdot$, it is called a *unity,* usually denoted 1 or 1_R, and R is called a *ring with unity.*

Notice that we are assuming the usual precedence of operations: operations inside parentheses are done first, but otherwise multiplications are done, left to right, before additions, also left to right. This usually isn't a problem, but there are cases where parentheses should be added for clarity. One example is the ring $\mathcal{P}(X)$, consisting of all subsets of a set X, where addition is symmetric difference $\triangle$ and multiplication is intersection $\cap$. One of the distributive laws reads: For all $A, B, C \subseteq X$,

$$A \cap (B \triangle C) = A \cap B \triangle A \cap C \ ;$$

and the right side is nearly incomprehensible without parentheses.

Of course, we know some rings immediately: $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. The additive subgroups $n\mathbb{Z}$ of $\mathbb{Z}$ are commutative rings, but they only have unity if $n = \pm 1$, i.e., if $n\mathbb{Z} = \mathbb{Z}$. The additive groups $\mathbb{Z}_n$, with multiplication mod n, are commutative rings with unity, and their subgroups $d\mathbb{Z}_n$ are commutative rings, usually without unity (though some will surprise by having unities where they are not expected; see the description of "idempotents" below). The basic example of a non-commutative ring is $M_{n\times n}(\mathbb{R})$. Other good examples of commutative rings with unity are $\mathbb{R}[x, y]$, the family of all polynomials with real coefficients in the variables x, y; and $\mathcal{F}(S)$, the family of all functions from the set S into $\mathbb{R}$, with pointwise operations: $f + g$ and fg are defined on S by, for all s in S,

$$(f+g)(s) = f(s) + g(s) \qquad \text{and} \qquad (fg)(s) = f(s)g(s) \ .$$

If S has various structures, they can be used to define subrings of $\mathcal{F}(S)$ — for example, if S is the interval $[a, b]$ on the real line, then we can study the ring of continuous functions on $[a, b]$, or differentiable functions on $[a, b]$, or ...

Terrible Example: Let G be an additive abelian group, and let N be the set of all functions $G \to G$, with the operations of pointwise addition and composition for multiplication. Then one of the distributive laws works but the other doesn't: Let $\alpha, \beta, \gamma : G \to G$; then for all x in G, we do have

$$\begin{aligned} ((\beta+\gamma)\circ\alpha)(x) &= (\beta+\gamma)(\alpha(x)) = \beta(\alpha(x)) + \gamma(\alpha(x)) \\ &= (\beta\circ\alpha)(x) + (\gamma\circ\alpha)(x) = (\beta\circ\alpha + \gamma\circ\alpha)(x) \end{aligned}$$

so $(\beta+\gamma)\circ\alpha = \beta\circ\alpha + \gamma\circ\alpha$, but

$$(\alpha\circ(\beta+\gamma))(x) = \alpha((\beta+\gamma)(x)) = \alpha((\beta(x)+\gamma(x)) \ ,$$

which, because α may not a homomorphism, need not be equal to

$$(\alpha\circ\beta + \alpha\circ\gamma)(x) = \alpha(\beta(x)) + \alpha(\gamma(x)) \ ,$$

so we may not have $\alpha\circ(\beta+\gamma) = \alpha\circ\beta + \alpha\circ\gamma$. This N is an example of a *near-ring,* a concept that I promise I will not mention again.

Even if there is a unity, elements other than the unity may not have multiplicative inverses (if a has one, it is called a *unit,* and the inverse is denoted as usual by a^{-1}), and cancellation from a multiplication may not be possible. For instance, in $\mathbb{Z}_6$, only 1 and 5 have (multiplicative) inverses, and even though $2\cdot 1 = 2\cdot 4$, we don't have $1 = 4$. Because $(R,+)$ is an additive abelian group, the laws of "exponents" hold, but they become laws of coefficients: For $n, m \in \mathbb{Z}$ and $a, b \in R$, we have

$$(n+m)a = na + ma \ , \qquad n(ma) = (nm)a \ , \qquad n(a+b) = na + nb$$

In particular, $0_{\mathbb{Z}}a = 0_R$ and $(-n)a = n(-a) = -(na)$. Positive integer exponents are used for repeated multiplication, as usual, and behave just as in a group; but we can only have 0 as an exponent in a ring with unity, and negative integer exponents only make sense for units. But at least some familiar facts from high school algebra hold in any ring. The first parts of the following proposition just show how addition and multiplication connect in any ring (so the distributive laws, the only connection we have between the operations, are used repeatedly):

Prop and Def: Let R be a ring. Then:

(a) For all a in R, $0_R \cdot a = a \cdot 0_R = 0_R$.

(b) For all a, b in R, $(-a)b = a(-b) = -(ab)$.

(c) For any integer n and a, b in R, $(na)b = a(nb) = n(ab)$.

(d) We define subtraction to be addition of the additive inverse (the negative): $a - b = a + (-b)$. Then multiplication distributes over subtraction: For all a, b, c in R,

$$a(b-c) = ab - ac \qquad \text{and} \qquad (b-c)a = ba - ca \ .$$

(e) The FOIL method works: $(a+b)(c+d) = ac + ad + bc + bd$ (but in a noncommutative ring, the order is important in the middle terms).

Pf: (a) We have $aa + 0_R = aa = (a + 0_R)a = aa + 0_R a$, and we <u>can</u> cancel terms from an addition (because R is a group under addition), so $0_R = 0_R a$. Similarly $a0_R = 0_R$.

(b) We have $ab + (-a)b = (a + (-a))b = 0b = 0$, so $(-a)b$ is the negative of ab. Similarly $a(-b) = -(ab)$.

(c) If n is positive, this is just the distributive law and induction. For the case $n = 0$, we note first that, in (a), 0 means the additive identity in R, not the integer 0. But our convention is that, in a multiplicative group, a^0 means the identity in the group; when the operation is addition, the exponent becomes a coefficient and the identity in the group is denoted 0; so $0_{\mathbb{Z}}a = 0_R$ is that

convention. With that convention, the case $n = 0$ is just (a). For $n < 0$, we use the positive case and (b).

(d)

$$a(b-c) = a(b+(-c)) = ab + a(-c) = ab + (-(ac)) = ab - ac \ ,$$

and similarly $(b-c)a = ba - ca$.

(e) Easy.//

More definitions:

Def: Let R be a ring and a be an element of R.

- a is *nilpotent* if some positive power of a is 0.
- a is *idempotent* if $a^2 = a$.
- a is a *zero-divisor* if there is a nonzero element b of R for which $ab = 0$.
- R is a(n *integral) domain* if it is commutative with unity and the only zero-divisor in R is 0 itself (or equivalently nonzero elements of R can be cancelled from a multiplication).
- R is a *division ring* if it has unity and every nonzero element is a unit (i.e., if $R - \{0\}$ is a group under multiplication).
- R is a *field* if it is a commutative division ring.

Pf of the equivalence in the def of integral domain: Suppose 0 is the only zero-divisor in R and $ab = ac$ where $a \neq 0$. Then $0 = ab - ac = a(b-c)$, and because a is not a zero-divisor, we must have $b - c = 0$, i.e., $b = c$. Conversely, suppose we can cancel nonzero elements of R from a multiplication, and that $ab = 0$ where $a \neq 0$. Then we have $ab = a0$, so by cancellation $b = 0$; i.e., a is not a zero-divisor.//

We can find inside various $\mathbb{Z}_n$'s elements that are nilpotents, idempotents and zero-divisors, provided n is not prime. For example, in $\mathbb{Z}_{24}$, 6 and 12 are nilpotent (because $6^3 = 0$ and $12^3 = 0$); 9 and 16 are idempotents (because $9^2 = 9$ and $16^2 = 16$ — note that $9 = 1 - 16$); and every element that is not relatively prime to 24 is a zero-divisor (for instance, $9 \cdot 8 = 0$).

There are not many interesting examples of division rings; one is the ring of quaternions: Take the elements I, J, K, L of $GL(2, \mathbb{C})$ that gave us Q_8; and form

$$\mathbb{H} = \{aI + bJ + cK + dL : a, b, c, d \in \mathbb{R}\} = \left\{ \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} : z, w \in \mathbb{C} \right\} \ ,$$

a subring of $M_{2\times 2}(\mathbb{C})$ (the overline denotes complex conjugate: $\overline{x + yi} = x - yi$ for x, y in $\mathbb{R}$). The rings $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are fields, and $\mathbb{Z}$ is an integral domain that is not a field. Other integral domains that are not fields are $\mathbb{R}[x, y]$ and the ring of "Gaussian integers" $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$. Others are $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} : a, b \in \mathbb{Z}\}$ and $\mathbb{Z}[\sqrt[3]{2}] = \{a + b\sqrt[3]{2} + c(\sqrt[3]{2})^2 : a, b, c \in \mathbb{Z}\}$.

Just as with groups, we can form the direct product of two rings (or two copies of the same ring); the multiplication on $R \times S$ is coordinatewise, like addition: $(a, c)(b, d) = (ab, cd)$. (The text claims that in the context of rings the direct product is called the direct sum and denoted $R \oplus S$, but that has not been my experience.) If R and S both have unities, then $(1_R, 1_S)$ is a unity for the direct product, and $(1_R, 0_S)$ and $(0_R, 1_S)$ are idempotents in the center of the direct

product. Conversely, if e is a central idempotent in a ring R with unity, then $1-e$ is another central idempotent, the ideals eR and $(1-e)R$ are also rings with unities e and $1-e$ respectively, and R is isomorphic to the direct product $eR \times (1-e)R$ (though we haven't defined isomorphisms of rings yet — you can probably guess what that means). Notice that, even if R and S are integral domains, $R \times S$ has many zero-divisors: $(r,0)(0,s) = (0,0)$.

Aside: In my experience, the difference between direct product and direct sum is as follows: Let $\{A_i : i \in I\}$ be a collection of groups or rings. Here I is the "index set" of the family; it may be just $\{1,2\}$, or it may be $\{1,2,\ldots,n\}$ for any positive integer n, or perhaps $\mathbb{N}$ — the infinite sequences in calculus are indexed by $\mathbb{N}$ — or $\mathbb{R}$, or any other set. The point is that there is an A_i for each i in I. Then the direct product $\prod_{i\in I} A_i$ is the set of "I-tuples", i.e., arrays $(a_i)_{i\in I}$ of elements, where the coordinate a_i in the array comes out of A_i, for each i in I. The operation(s) on this product are defined coordinate-wise:

$$(a_i)_{i\in I} \# (b_i)_{i\in I} = (a_i \#_i b_i)_{i\in I}$$

for any $(a_i)_{i\in I}, (b_i)_{i\in I}$ in the product, where $\#_i$ is the operation in A_i and $\#$ is the operation being defined on the product. And the direct sum $\bigoplus_{i\in I} A_i$ is the subgroup of the product consisting of the "I-tuples" $(a_i)_{i\in I}$ in which all but finitely many of the a_i's are the zero element of their respective A_i. Thus, the direct product and the direct sum are the same if I is finite. (So the text is not wrong in rewriting $R \times S$ as $R \oplus S$; I just haven't seen it done that way.) But if I is infinite, the direct sum is much smaller — for instance, if I and each A_i are countably infinite, then the direct sum is still countably infinite, while the direct product is uncountable.

The text proves the nice fact that a finite integral domain is a field. Here is the proof: Let a be a nonzero element of the finite integral domain R. Then the function $R \to R : r \mapsto ar$ is one-to-one, because a can be cancelled; and because R is finite, the function is also onto R; so there is an element r of R for which $ar = 1$, i.e., r is a multiplicative inverse of a.

The book also talks about Wedderburn's Theorem, which proves that every finite division ring is commutative, i.e., is a field. I've put a proof of this result online — it is not too long (about one-and-a-half pages), but it is very odd, using several unexpected results.