Ring Theory (part 3): Homomorphisms, Ideals, and Quotients (by Evan Dummit, 2018, v. 1.01)

Contents

3 Homomorphisms, Ideals, and Quotients

In this chapter, we will examine some more intricate properties of general rings. We begin with a discussion of isomorphisms, which provide a way of identifying two rings whose structures are identical, and then examine the broader class of ring homomorphisms, which are the "structure-preserving functions" from one ring to another. Next, we study ideals and quotient rings, which provide the most general version of "modular arithmetic" in a ring, and which are fundamentally connected with ring homomorphisms. We close with a detailed study of the structure of ideals and quotients in commutative rings with 1.

3.1 Ring Isomorphisms and Homomorphisms

- We begin our study with a discussion of "structure-preserving maps" between rings.

3.1.1 Ring Isomorphisms

- We have encountered several examples of rings with very similar structures.
- For example, consider the two rings $R = \mathbb{Z}/6\mathbb{Z}$ and $S = (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/3\mathbb{Z})$.
  - ◦ Here are the addition and multiplication tables in R:

+	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

·	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

- Now compare those tables to the tables in S:

$+$	$(0,0)$	$(1,1)$	$(0,2)$	$(1,0)$	$(0,1)$	$(1,2)$
$(0,0)$	$(0,0)$	$(1,1)$	$(0,2)$	$(1,0)$	$(0,1)$	$(1,2)$
$(1,1)$	$(1,1)$	$(0,2)$	$(1,0)$	$(0,1)$	$(1,2)$	$(0,0)$
$(0,2)$	$(0,2)$	$(1,0)$	$(0,1)$	$(1,2)$	$(0,0)$	$(1,1)$
$(1,0)$	$(1,0)$	$(0,1)$	$(1,2)$	$(0,0)$	$(1,1)$	$(0,2)$
$(0,1)$	$(0,1)$	$(1,2)$	$(0,0)$	$(1,1)$	$(0,2)$	$(1,0)$
$(1,2)$	$(1,2)$	$(0,0)$	$(1,1)$	$(0,2)$	$(1,0)$	$(0,1)$

$\cdot$	$(0,0)$	$(1,1)$	$(0,2)$	$(1,0)$	$(0,1)$	$(1,2)$
$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$
$(1,1)$	$(0,0)$	$(1,1)$	$(0,2)$	$(1,0)$	$(0,1)$	$(1,2)$
$(0,2)$	$(0,0)$	$(0,2)$	$(0,1)$	$(0,0)$	$(0,2)$	$(0,1)$
$(1,0)$	$(0,0)$	$(1,0)$	$(0,0)$	$(1,0)$	$(0,0)$	$(1,0)$
$(0,1)$	$(0,0)$	$(0,1)$	$(0,2)$	$(0,0)$	$(0,1)$	$(0,2)$
$(1,2)$	$(0,0)$	$(1,2)$	$(0,1)$	$(1,0)$	$(0,2)$	$(1,1)$

- Notice that these tables look quite similar (especially given the artful reordering of the labels of the elements in S).
- Indeed, if we relabel each entry n in the first set of tables with the ordered pair corresponding to its reduction modulo 2 and 3 (so that 1 becomes $(1,1)$, 2 becomes $(0,2)$, and so forth) we will obtain the second set of tables!

- For another example, consider the rings $R = (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$ and $S = \mathbb{F}_2[x]/(x^2 + x)$.
  - Here are the addition and multiplication tables in R:

$+$	$(0,0)$	$(1,1)$	$(1,0)$	$(0,1)$
$(0,0)$	$(0,0)$	$(1,1)$	$(1,0)$	$(0,1)$
$(1,1)$	$(1,1)$	$(0,0)$	$(0,1)$	$(1,0)$
$(1,0)$	$(1,0)$	$(0,1)$	$(0,0)$	$(1,1)$
$(0,1)$	$(0,1)$	$(1,0)$	$(1,1)$	$(0,0)$

$\cdot$	$(0,0)$	$(1,1)$	$(1,0)$	$(0,1)$
$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$
$(1,1)$	$(0,0)$	$(1,1)$	$(1,0)$	$(0,1)$
$(1,0)$	$(0,0)$	$(1,0)$	$(1,0)$	$(0,0)$
$(0,1)$	$(0,0)$	$(0,1)$	$(0,0)$	$(0,1)$

  - Now compare those tables to the tables in S:

$+$	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

$\cdot$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	x	0
$x+1$	0	$x+1$	0	$x+1$

  - Here, if we relabel $(0,0)$ as 0, $(1,1)$ as 1, $(1,0)$ as x, and $(0,1)$ as $x+1$, the first pair of tables becomes the second set of tables.

- As a third example, consider the rings $R = \mathbb{C} = \{a+bi \,:\, a, b \in \mathbb{R}\}$ and $S = \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \in M_{2\times 2}(\mathbb{R}) \,:\, a, b \in \mathbb{R} \right\}$.
  - Notice that S is a subring of $M_{2\times 2}(\mathbb{R})$: we have $\begin{bmatrix} a & b \\ -b & a \end{bmatrix} - \begin{bmatrix} c & d \\ -d & c \end{bmatrix} = \begin{bmatrix} a-c & b-d \\ -(b-d) & a-c \end{bmatrix}$ and $\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \cdot \begin{bmatrix} c & d \\ -d & c \end{bmatrix} = \begin{bmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{bmatrix}$.
  - Compare these to the addition and multiplication operations in $\mathbb{C}$: $(a+bi) - (c+di) = (a-c) + (b-d)i$ and $(a+bi) \cdot (c+di) = (ac-bd) + (ad+bc)i$.
  - Upon identifying the complex number $a+bi$ with the matrix $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, we see that the ring structure of S is precisely the same as the ring structure of $\mathbb{C}$.

- Let us formalize the central idea in the examples above: in each case, we see that there is a way to "relabel" the elements of R using the elements of S in a way that preserves the ring structure.
  - The desired "relabeling" is a function $\varphi : R \to S$ with the property that φ is a bijection (so that each element of R is "labeled" with a unique element of S) and that φ respects the ring operations.
  - Explicitly, we require $\varphi(r_1 + r_2) = \varphi(r_1) + \varphi(r_2)$ and $\varphi(r_1 \cdot r_2) = \varphi(r_1) \cdot \varphi(r_2)$ for all elements r_1 and r_2 in R.

- Definition: Let R and S be rings. A ring isomorphism φ from R to S is a bijective[1] function $\varphi : R \to S$ such that $\varphi(r_1 + r_2) = \varphi(r_1) + \varphi(r_2)$ and $\varphi(r_1 \cdot r_2) = \varphi(r_1) \cdot \varphi(r_2)$ for all elements r_1 and r_2 in R.

[1] Recall that a function $\varphi : R \to S$ is injective (one-to-one) if $\varphi(x) = \varphi(y)$ implies $x = y$, and φ is surjective (onto) if for every $s \in S$ there exists an $r \in R$ with $\varphi(r) = s$. A bijective function is one that is both injective and surjective. Equivalently, φ is a bijection if it possesses a two-sided inverse function $\varphi^{-1} : S \to R$ with $\varphi(\varphi^{-1}(s)) = s$ and $\varphi^{-1}(\varphi(r)) = r$ for every $r \in R$ and $s \in S$.

- We remark here that in both of the conditions $\varphi(r_1 + r_2) = \varphi(r_1) + \varphi(r_2)$ and $\varphi(r_1 \cdot r_2) = \varphi(r_1) \cdot \varphi(r_2)$, the operations on the left are performed in R while the operations on the right are performed in S.
- Note: Isomorphisms arise in a variety of contexts (e.g., isomorphisms of vector spaces, isomorphisms of groups, etc.), and in some cases the rings we are considering may carry additional structure. We will simply say "isomorphism" rather than explicitly specifying "ring isomorphism" each time, unless there is a particular reason to do otherwise.

- Example: For $R = \mathbb{Z}/6\mathbb{Z}$ and $S = (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/3\mathbb{Z})$, the map $\varphi : R \to S$ defined via $\varphi(n) = (n \bmod 2,\, n \bmod 3)$ is an isomorphism.
  - Note that "reducing" a residue class in $\mathbb{Z}/6\mathbb{Z}$ modulo 2 or modulo 3 is well-defined, since 2 and 3 both divide 6, so φ is well-defined.
  - It is then easy to see that φ is a bijection, and that $\varphi(r_1+r_2) = \varphi(r_1)+\varphi(r_2)$ and $\varphi(r_1 \cdot r_2) = \varphi(r_1) \cdot \varphi(r_2)$ for any residue classes $r_1, r_2 \in \mathbb{Z}/6\mathbb{Z}$. (It is also possible to compare the addition and multiplication tables as we did above.)
  - We therefore conclude that φ is an isomorphism.

- Example: For $S = \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \in M_{2\times 2}(\mathbb{R}) \,:\, a, b \in \mathbb{R} \right\}$, show that the map $\varphi : \mathbb{C} \to S$ defined via $\varphi(a+bi) = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ is an isomorphism.
  - First, we see that φ is a bijection since it has a two-sided inverse; namely, the map $\varphi^{-1} : S \to \mathbb{C}$ defined by $\varphi^{-1}\left(\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \right) = a + bi$.
  - Furthermore, if $z = a + bi$ and $w = c + di$, then
$$\varphi(z + w) = \varphi((a + c) + (b + d)i) = \begin{bmatrix} a+c & b+d \\ -(b+d) & a+c \end{bmatrix} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} + \begin{bmatrix} c & d \\ -d & c \end{bmatrix} = \varphi(z) + \varphi(w)$$
and also
$$\varphi(zw) = \varphi((ac - bd) + (ad + bc)i) = \begin{bmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{bmatrix} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \cdot \begin{bmatrix} c & d \\ -d & c \end{bmatrix} = \varphi(z) \cdot \varphi(w).$$
  - Thus, φ satisfies all the requirements, so it is an isomorphism.

- Definition: If there is an isomorphism $\varphi : R \to S$, we say R and S are isomorphic, and write $R \cong S$.
  - Intuitively, isomorphic rings share the same structure, except that the elements and operations may be labeled differently.

- Proposition (Properties of Isomorphisms): If R, S, T are any rings, the following hold:
  1. The identity map $I : R \to R$ defined by $I(r) = r$ for all $r \in R$ is an isomorphism from R to R.
     - Proof: I is clearly a bijection and respects the ring operations.
  2. If $\varphi : R \to S$ is an isomorphism, then the inverse map $\varphi^{-1} : S \to R$ is also an isomorphism.
     - Proof: Essentially by definition, φ^{-1} is also a bijection.
     - Now suppose $\varphi^{-1}(s_1) = r_1$ and $\varphi^{-1}(s_2) = r_2$, so that $\varphi(r_1) = s_1$ and $\varphi(r_2) = s_2$.
     - Then $\varphi(r_1+r_2) = \varphi(r_1)+\varphi(r_2) = s_1+s_2$, meaning that $\varphi^{-1}(s_1+s_2) = r_1+r_2 = \varphi^{-1}(s_1)+\varphi^{-1}(s_2)$, and likewise for multiplication. Thus, φ^{-1} is also an isomorphism.
  3. If $\varphi : R \to S$ and $\psi : S \to T$ are isomorphisms, then the composition $\psi\varphi : R \to T$ is also an isomorphism.
     - Proof: It is straightforward to see that the composition of two bijections is a bijection.
     - Furthermore, we have $(\psi\varphi)(r_1 + r_2) = \psi(\varphi(r_1 + r_2)) = \psi(\varphi(r_1) + \varphi(r_2)) = \psi\varphi(r_1) + \psi\varphi(r_2)$, and likewise for multiplication. Thus $\psi\varphi$ is an isomorphism.

4. If $\varphi : R \to S$ is an isomorphism, then $\varphi(0_R) = 0_S$, and if R has a 1, then so does S, and $\varphi(1_R) = 1_S$.
   - Proof: For any $r \in R$, we have $\varphi(r) = \varphi(r + 0_R) = \varphi(r) + \varphi(0_R)$: thus, by additive cancellation in S we see $\varphi(0_R) = 0_S$.
   - Likewise, if R has a 1, then let $s \in S$ be arbitrary and $r = \varphi^{-1}(s)$. Then $s \cdot \varphi(1_R) = \varphi(r)\varphi(1_R) = \varphi(r \cdot 1_R) = \varphi(r) = s$, and likewise $\varphi(1_R) \cdot s = s$, so $\varphi(1_R)$ is a multiplicative identity in S.

- From the proposition, we immediately see that "being isomorphic" is an equivalence relation on any collection of rings. In general, it is not easy to determine whether two given rings are isomorphic, and even if two rings are isomorphic, there is no general method for constructing an isomorphism between them.
  - Two isomorphic rings have the same additive and multiplicative structures. Thus, any statement that only depends on the ring operations must be identical in two isomorphic rings.
  - Thus, for example, if $\varphi : R \to S$ is an isomorphism, then R is commutative if and only if S is commutative, and R has a 1 if and only if S has a 1. Likewise, R has zero divisors if and only if S has zero divisors, and the cardinalities of any two isomorphic rings (along with their sets of units) must be equal.
  - So, for example, we see that $M_{2\times 2}(\mathbb{R})$ is not isomorphic to the ring of real quaternions $\mathbb{H}$, since the former has zero divisors and the latter does not.
  - Likewise, we see that none of the rings $\mathbb{Z}/m\mathbb{Z}$ for $m > 1$ are isomorphic to one another, since they all have different cardinalities.
  - In a similar way, the ring $\mathbb{R}$ is not isomorphic to $\mathbb{C}$ since the polynomial equation $x^2 + 1 = 0$ has no solutions in $\mathbb{R}$, but does have solutions in $\mathbb{C}$.
  - As a final example, the rings $\mathbb{Z}/4\mathbb{Z}$ and $(\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$ are not isomorphic: there are two solutions to $x^2 = 0$ in the first ring (namely, 0 and 2) while there is only one solution to $x^2 = 0$ in the second ring (namely, $(0, 0)$). Alternatively, the first ring has 2 units, while the second ring has only 1.

- We also remark that there can exist nontrivial isomorphisms of a ring with itself. Such maps are known as automorphisms.
  - Remark (for those who like group theory): The set of automorphisms of a ring forms a group under function composition.

- Example: Show that the complex conjugation map $\varphi(a + bi) = a - bi$ is an isomorphism from $\mathbb{C}$ to $\mathbb{C}$.
  - It is easy to see that φ is a bijection, since it is its own inverse function.
  - Furthermore, it is a straightforward calculation that $\varphi(z + w) = \varphi(z) + \varphi(w)$ and $\varphi(zw) = \varphi(z)\varphi(w)$ for any complex numbers z and w, so φ is an isomorphism.

3.1.2 Ring Homomorphisms

- We now study maps that respect the structure of ring operations without the requirement that they be bijections.

- Definition: A function $\varphi : R \to S$ is a ring homomorphism if $\varphi(r_1 + r_2) = \varphi(r_1) + \varphi(r_2)$ and $\varphi(r_1 \cdot r_2) = \varphi(r_1) \cdot \varphi(r_2)$ for all elements r_1 and r_2 in R.
  - Note of course that any isomorphism is a homomorphism, but the reverse is not typically true.

- Example: If $m > 1$, show that the map $\varphi : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ defined by $\varphi(a) = \overline{a}$, so that φ maps the integer a to its associated residue class $\overline{a}$ modulo m, is a ring homomorphism.
  - From our results on residue classes, we see $\varphi(a + b) = \overline{a + b} = \overline{a} + \overline{b} = \varphi(a) + \varphi(b)$, and likewise $\varphi(a \cdot b) = \overline{a \cdot b} = \overline{a} \cdot \overline{b} = \varphi(a) \cdot \varphi(b)$. Thus, φ is a homomorphism.
  - Notice that this map is surjective but not injective (since for example $\varphi(0) = \varphi(m)$), so it is not an isomorphism.

- In essentially the same way, we see that the "reduction modulo p" map inside $F[x]$ is also a homomorphism:

- Example: Let F be a field with $R = F[x]$ and let $p(x) \in R$ be nonzero. Then the map $\varphi : R \to R/pR$ given by $\varphi(a) = \overline{a}$, mapping the polynomial a to its associated residue class $\overline{a}$ modulo p, is a ring homomorphism.
  - From our results on residue classes, we see $\varphi(a+b) = \overline{a+b} = \overline{a} + \overline{b} = \varphi(a) + \varphi(b)$, and likewise $\varphi(a \cdot b) = \overline{a \cdot b} = \overline{a} \cdot \overline{b} = \varphi(a) \cdot \varphi(b)$. Thus, φ is a homomorphism.
  - In the next section, we will generalize the ideas in these two examples and describe a general procedure for constructing a "quotient ring".

- Example: Let R be a commutative ring and $a \in R$. Show that the "evaluation at a map" $\varphi_a : R[x] \to R$ defined by $\varphi_a(p) = p(a)$ is a ring homomorphism.
  - We have $\varphi_a(p+q) = (p+q)(a) = p(a) + q(a) = \varphi_a(p) + \varphi_a(q)$ by the definition of polynomial addition.
  - Likewise, we have $\varphi_a(r_b x^b \cdot r_c x^c) = r_b r_c a^{b+c} = (r_b a^b)(r_c a^c) = \varphi_a(r_b x^b)\varphi_a(r_c x^c)$ because R is commutative.
  - Then for any polynomials p and q we see $\varphi_a(pq) = \varphi_a(p)\varphi_a(q)$ by applying distributivity and the fact that φ_a respects multiplication of individual terms and addition.

- Example: Let R and S be any rings. The "zero map" $Z : R \to S$ given by $Z(r) = 0_S$ for every $r \in R$ is a ring homomorphism.

- Example: If S is a subring of R, the map $\iota : S \to R$ given by $\iota(s) = s$ is a ring homomorphism. This map is called the inclusion map (since it simply reflects the set inclusion of S inside R).

- There exist numerous examples of maps that satisfy only one of the two requirements for being a homomorphism.
  - Non-Example: The function $f : \mathbb{Z} \to \mathbb{Z}$ given by $f(n) = 2n$ is not a homomorphism. Explicitly, although it satisfies $f(m+n) = 2(m+n) = f(m) + f(n)$, it is not multiplicative since $f(1 \cdot 1) = 2$ while $f(1) \cdot f(1) = 4$.
  - Non-Example: The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is not a homomorphism. Explicitly, although it satisfies $f(xy) = (xy)^2 = f(x)f(y)$, it is not additive since $f(1+1) = 4$ while $f(1) + f(1) = 2$.

- Here are a few more examples (and non-examples) of homomorphisms:

- Example: Determine whether the map $\varphi : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}$ given by $\varphi\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = b$ is a ring homomorphism.
  - We see that $\varphi\left(\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} + \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}\right) = b_1 + b_2 = \varphi\left(\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}\right) + \varphi\left(\begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}\right)$.
  - However, $\varphi\left(\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \cdot \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}\right) = a_1 b_2 + b_1 d_2$ while $\varphi\left(\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}\right) \cdot \varphi\left(\begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}\right) = b_1 b_2$, and these expressions are not equal in general. Thus, φ is $\boxed{\text{is not a homomorphism}}$.

- Example: Determine whether the map $\varphi : (\mathbb{Z}/15\mathbb{Z}) \to (\mathbb{Z}/15\mathbb{Z})$ given by $\varphi(a) = 10a$ is a ring homomorphism.
  - We have $\varphi(a+b) = 10(a+b) = 10a + 10b = \varphi(a) + \varphi(b)$.
  - Likewise, $\varphi(ab) = 10ab = 100ab = (10a)(10b) = \varphi(a)\varphi(b)$, since $10 \equiv 100 \pmod{15}$.
  - Therefore, φ $\boxed{\text{is a homomorphism}}$.

- Example: Let R be the ring of infinitely differentiable real-valued functions on $\mathbb{R}$. Determine whether the derivative map $D : R \to R$ given by $D(f) = f'$ is a ring homomorphism.
  - We have $D(f+g) = (f+g)' = f' + g' = D(f) + D(g)$, so D is additive.
  - However, D does not respect ring multiplication, since for example $D(x \cdot x^2) = 3x^2$ while $D(x) \cdot D(x^2) = 2x$. Therefore, φ $\boxed{\text{is not a homomorphism}}$.

- Example: Let R be any ring. Determine whether the map $\varphi : R \to R \times R$ given by $\varphi(r) = (r, r)$ is a ring homomorphism.
  - We have $\varphi(r+s) = (r+s, r+s) = (r, r) + (s, s) = \varphi(r) + \varphi(s)$.
  - Likewise, $\varphi(rs) = (rs, rs) = (r, r)(s, s) = \varphi(r)\varphi(s)$, so φ $\boxed{\text{is a homomorphism}}$.
- Like with isomorphisms, homomorphisms have a number of basic properties.
- Proposition (Properties of Homomorphisms): If R, S, T are any rings, the following hold:
  1. If $\varphi : R \to S$ and $\psi : S \to T$ are homomorphisms, then the composition $\psi\varphi : R \to T$ is also a homomorphism.
     - Proof: Follows from the analogous calculation for isomorphisms.
  2. If $\varphi : R \to S$ is a homomorphism, then $\varphi(0_R) = 0_S$, $\varphi(-r) = -\varphi(r)$ for every $r \in R$, and $\varphi(r_1 - r_2) = \varphi(r_1) - \varphi(r_2)$ for every $r_1, r_2 \in R$.
     - Proof: For any $r \in R$, we have $\varphi(r) = \varphi(r + 0_R) = \varphi(r) + \varphi(0_R)$: thus, by additive cancellation in S we see $\varphi(0_R) = 0_S$.
     - Then $0_S = \varphi(0_R) = \varphi(r + (-r)) = \varphi(r) + \varphi(-r)$ so by the uniqueness of additive inverses in S we conclude $\varphi(-r) = -\varphi(r)$.
     - Finally, $\varphi(r_1 - r_2) = \varphi(r_1) + \varphi(-r_2) = \varphi(r_1) - \varphi(r_2)$ by the above calculation.
  3. If $\varphi : R \to S$ is a surjective homomorphism and R has a 1, then S also has a 1 and $\varphi(1_R) = 1_S$. Furthermore, for any unit $u \in R$, the value $\varphi(u)$ is a unit in S whose inverse is $\varphi(u^{-1})$.
     - Proof: Let $s \in S$: then since φ is surjective there exists some $r \in R$ with $\varphi(r) = s$. Then $s\varphi(1_R) = \varphi(r)\varphi(1_R) = \varphi(r1_R) = \varphi(r) = s$, and likewise $\varphi(1_R)s = s$, so $\varphi(1_R)$ is a multiplicative identity in S.
     - For the other part, if u is a unit in R then $1_S = \varphi(1_R) = \varphi(u \cdot u^{-1}) = \varphi(u)\varphi(u^{-1})$, so $\varphi(u)$ is a unit in S with inverse $\varphi(u^{-1})$.
- Associated to a homomorphism are two fundamental objects: the kernel and image.
- Definition: If $\varphi : R \to S$ is a ring homomorphism, the kernel of φ, denoted $\ker \varphi$, is the set of elements in R mapped to 0_S by φ. In other words, $\ker \varphi = \{r \in R : \varphi(r) = 0\}$.
  - Intuitively, the kernel measures how close φ is to being the zero map: if the kernel is large, then φ sends many elements to zero, while if the kernel is small, φ sends fewer elements to zero.
  - Example: The kernel of the reduction homomorphism $\varphi : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ with $\varphi(a) = \overline{a}$ is the subring $m\mathbb{Z}$.
  - Example: The kernel of the evaluation map $\varphi_a : F[x] \to F$ given by $\varphi_a(p) = p(a)$ is the set of polynomials in $F[x]$ with $p(a) = 0$, which is (equivalently) the set of polynomials divisible by $x - a$.
- Definition: If $\varphi : R \to S$ is a ring homomorphism, the image of φ, denoted $\operatorname{im} \varphi$, is the set of elements in S of the form $\varphi(r)$ for some $r \in R$.
  - In the context of general functions, the image is often called the range of φ.
  - Intuitively, the image measures how close φ is to being surjective: indeed (by definition) φ is surjective if and only if $\operatorname{im} \varphi = S$.
- The kernel and image of a homomorphism are subrings of R and S respectively:
- Proposition (Kernel and Image): Let $\varphi : R \to S$ be a ring homomorphism. Then
  1. The image $\operatorname{im} \varphi$ is a subring of S.
     - Proof: Since $\varphi(0_R) = 0_S$, the image contains 0. Furthermore, if s_1 and s_2 are in $\operatorname{im} \varphi$ so that $\varphi(r_1) = s_1$ and $\varphi(r_2) = s_2$ for some $r_1, r_2 \in R$, then $s_1 - s_2 = \varphi(r_1 - r_2)$ and $s_1 s_2 = \varphi(r_1 r_2)$ are also in $\operatorname{im} \varphi$.
     - Thus, $\operatorname{im} \varphi$ contains 0 and is closed under subtraction and multiplication, so it is a subring.

2. The kernel $\ker\varphi$ is a subring of R. In fact, if $x \in \ker\varphi$, then rx and xr are in $\ker\varphi$ for any $r \in R$: in other words, $\ker\varphi$ is closed under multiplication by arbitrary elements of R.
   - Proof: Since $\varphi(0_R) = 0_S$, the kernel contains 0. Furthermore, if r_1 and r_2 are in $\ker\varphi$ then $\varphi(r_1 - r_2) = \varphi(r_1) - \varphi(r_2) = 0$ and $\varphi(r_1 r_2) = \varphi(r_1)\varphi(r_2) = 0 \cdot 0 = 0$
   - Thus, $\ker\varphi$ contains 0 and is closed under subtraction and multiplication, so it is a subring.
   - Moreover, if $x \in \ker\varphi$ then $\varphi(rx) = \varphi(r)\varphi(x) = \varphi(r)0 = 0$ and likewise $\varphi(xr) = \varphi(x)\varphi(r) = 0\varphi(r) = 0$.
3. The kernel is zero (i.e., $\ker\varphi = \{0\}$) if and only if φ is injective. In particular, φ is an isomorphism if and only if $\ker\varphi = \{0\}$ and $\operatorname{im}\varphi = S$.
   - Proof: If $\varphi(a) = \varphi(b)$, then $\varphi(a - b) = \varphi(a) - \varphi(b) = 0$, so $a - b \in \ker\varphi$. Thus, if the only element in $\ker\varphi$ is 0, then we must have $a - b = 0$ so that $a = b$.
   - Conversely, if $x \in \ker\varphi$ and φ is injective, then $\varphi(x) = 0 = \varphi(0)$ implies $x = 0$.
   - The second statement follows from the facts that $\ker\varphi = \{0\}$ is equivalent to φ being injective and $\operatorname{im}\varphi = S$ is equivalent to φ being surjective.

3.2 Ideals and Quotient Rings

- Our next task is to generalize the idea of "modular arithmetic" into general rings.
  - To motivate our discussion, recall the ideas behind the construction of $\mathbb{Z}/m\mathbb{Z}$ and R/pR where $R = F[x]$: we first defined modular modular congruences and studied their properties, and then we constructed residue classes and showed that the collection of all residue classes had a ring structure.
- In both $\mathbb{Z}$ and $F[x]$, we defined modular congruences using divisibility, but let us take a broader approach: if I is a subset of R (whose properties we intend to characterize in a moment) let us say that two elements $a, b \in R$ are "congruent modulo I" if $a - b \in I$.
  - This is a generalization of both types of congruence we have described thus far: for $\mathbb{Z}/m\mathbb{Z}$, the set I consists of the multiples of m, while for R/pR, the set I consists of the multiples of p.
  - We would like "congruence modulo I" to be an equivalence relation: this requires $a \equiv a \pmod{I}$, $a \equiv b \pmod{I}$ implies $b \equiv a \pmod{I}$, and $a \equiv b \pmod{I}$ and $b \equiv c \pmod{I}$ implies $a \equiv c \pmod{I}$.
  - It is easy to see that these three conditions require $0 \in I$, that I be closed under additive inverses, and that I be closed under addition. (Thus, I is in fact closed under subtraction.)
  - Furthermore, we would like the congruences to respect addition and multiplication: if $a \equiv b \pmod{I}$ and $c \equiv d \pmod{I}$, then we want $a + c \equiv b + d \pmod{I}$ and $ac \equiv bd \pmod{I}$.
  - In terms of ring elements, this is equivalent to the following: if $b = a + r$ and $d = c + s$ for some $r, s \in I$, then we want $(b+d) - (a+c) = r + s$ to be in I, and we also want $bd - ac = (a+r)(c+s) - ac = as + rc + rs$ to be in I.
  - The first condition clearly follows from the requirement that I is closed under addition. It is a bit less obvious how to handle the second condition, but one immediate implication follows by setting $a = c = 0$: namely, that $rs \in I$.
  - Thus, I must be closed under multiplication, so it is in fact a subring of R.
  - But the well-definedness of multiplication actually requires more: since $0 \in I$, we can set $r = 0$ to see that $as \in I$, and we can also set $s = 0$ to see that $rc \in I$.
  - So in fact, I must be closed under (left and right) multiplication by *arbitrary* elements of R, in addition to being a subring. It is then easy to see that this condition is also sufficient to ensure that $a \equiv b \pmod{I}$ and $c \equiv d \pmod{I}$ imply $a + c \equiv b + d \pmod{I}$ and $ac \equiv bd \pmod{I}$.
  - Our last task is to define residue classes and then the ring operations: we define the residue class $\overline{a}$ (modulo I) to be the set of ring elements b congruent to a modulo I, which is to say, $\overline{a} = \{a + r : r \in I\}$.
  - Then we take the operations on residue classes to be $\overline{a} + \overline{b} = \overline{a+b}$ and $\overline{a} \cdot \overline{b} = \overline{a \cdot b}$: then from our properties of congruences, we can verify that these operations are well-defined and that the collection of residue classes forms a ring.

3.2.1 Ideals

- Now that we have established the basic properties of the classes of the sets I we can use to construct congruences, we can run through the discussion more formally.

- Definition: A subring I of a ring R is called a left ideal of R if it is closed under arbitrary left multiplication by elements of R, and it is called a right ideal if it is closed under arbitrary right multiplication by elements of R.

    - Explicitly, I is a left ideal if I contains 0 and for any $x, y \in I$ and any $r \in R$, the elements $x - y$ and rx are in I, while I is a right ideal if I contains 0 and for any $x, y \in I$ and any $r \in R$, the elements $x - y$ and xr are in I.

- Definition: A subset I of a ring R that is both a left and a right ideal is called an ideal of R (or, for emphasis, a two-sided ideal).

    - Explicitly, I is an ideal if I contains 0 and for any $x, y \in I$ and any $r \in R$, the elements $x - y$, rx, and xr are all in I.
    - If R is commutative, then left ideals, right ideals, and two-sided ideals are the same. (As we will mention below, when R is not commutative, there may exist left ideals that are not right ideals and vice versa.)

- Here are a few basic examples of ideals:

    - Example: The subrings $n\mathbb{Z}$ are ideals of $\mathbb{Z}$, since they are clearly closed under arbitrary multiplication by elements of $\mathbb{Z}$.
    - Example: If $R = F[x]$ and p is any polynomial, the subring pR of multiples of p is an ideal of $F[x]$, since it is closed under arbitrary multiplication by polynomials in $F[x]$.
    - Non-example: The subring $\mathbb{Z}$ of $\mathbb{Q}$ is not an ideal of $\mathbb{Q}$, since it is not closed under arbitrary multiplication by elements of $\mathbb{Q}$, since for example if we take $r = \dfrac{1}{3} \in \mathbb{Q}$ and $x = 4 \in \mathbb{Z}$, the element $rx = \dfrac{4}{3}$ is not in $\mathbb{Z}$.
    - Example: For any ring R, the subrings $\{0\}$ and R are ideals of R. We refer to $\{0\}$ as the trivial ideal (or the "zero ideal") and refer to any ideal $I \neq R$ as a proper ideal (since it is a proper subset of R).

- Here are a few more examples (and non-examples) of ideals.

- Example: In the polynomial ring $\mathbb{Z}[x]$, determine whether the set S of polynomials with even constant term (i.e., the polynomials of the form $2a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ for integers a_i) forms an ideal.

    - It is easy to see that $0 \in S$ and that S is closed under subtraction.
    - Furthermore, if $q(x)$ is any other polynomial, and $p(x) \in S$, then $p(x)q(x)$ also has even constant term, so it is also in S.
    - Thus, S is closed under multiplication by elements of $\mathbb{Z}[x]$, so it $\boxed{\text{is an ideal}}$.

- Example: Determine whether the set S of upper-triangular 2×2 matrices is a left ideal or a right ideal of $M_{2\times 2}(\mathbb{R})$.

    - The upper-triangular matrices form a subring, so we need only determine whether they are closed under multiplication by arbitrary 2×2 matrices on the left and the right.
    - We can see that if $r = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $x = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ then x is upper-triangular but $rx = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$. Thus, S $\boxed{\text{is not a left ideal}}$.
    - Indeed, with the same choices, we have $xr = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$, so S also $\boxed{\text{is not a right ideal}}$.

- Example: Determine whether the set $S = \{\overline{0}, \overline{2}, \overline{4}, \overline{6}\}$ of "even" residue classes is an ideal of $\mathbb{Z}/8\mathbb{Z}$.

    - We have $0 \in S$, and it is a straightforward calculation to see that S is closed under subtraction, since the sum of two "even" residue classes modulo 8 will still be even.

- Furthermore, the product of any residue class with an even residue class will again be an even residue class (since 8 is even), so S is closed under multiplication by arbitrary elements of R. Thus, S $\boxed{\text{is an ideal}}$.

- Example: Determine whether the set $S = \{(2a, 3a) : a \in \mathbb{Z}\}$ is an ideal of $\mathbb{Z} \times \mathbb{Z}$.
  - We have $0 \in S$, and $(2a, 3a) - (2b, 3b) = (2(a-b), 3(a-b))$ so S is closed under subtraction.
  - But, for example, we can see that $(1, 2) \cdot (2, 3) = (2, 6)$ is not in S, even though $(2, 3)$ is, so S is not closed under arbitrary multiplication by elements of $\mathbb{Z} \times \mathbb{Z}$. Thus, S $\boxed{\text{is not an ideal}}$.

- Example: Determine whether the set S of matrices of the form $\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}$ for $a, b \in \mathbb{R}$ is a left ideal or a right ideal of $M_{2\times 2}(\mathbb{R})$.
  - Clearly $0 \in S$ and S is closed under subtraction. Furthermore, since $\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix} \cdot \begin{bmatrix} c & 0 \\ d & 0 \end{bmatrix} = \begin{bmatrix} ac & 0 \\ bc & 0 \end{bmatrix}$, S is also closed under multiplication, so it is a subring.
  - Also, for $r = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ and $x = \begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}$ we have $rx = \begin{bmatrix} ea + fb & 0 \\ ga + hb & 0 \end{bmatrix}$ and $xr = \begin{bmatrix} ae & af \\ be & bf \end{bmatrix}$.
  - Since $rx \in S$ for every $r \in R$ and $x \in S$, but that xr is not always in S, we see that S $\boxed{\text{is a left ideal}}$ but $\boxed{\text{is not a right ideal}}$ (and hence not a two-sided ideal either).
  - Remark: By taking transposes, we can also see that the set of matrices of the form $\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$ is a right ideal of $M_{2\times 2}$ that is not a left ideal.

- Several of the examples above are particular instances of a general class of ideals:

- Proposition (Principal Ideals): If R is a commutative ring with 1, the set $(a) = \{ra : r \in R\}$ of all R-multiples of a forms a (two-sided) ideal of R, known as the principal ideal generated by a.
  - Proof: Since $0a = 0$ we see $0 \in (a)$. Furthermore, since $ra - sa = (r - s)a$ we see that (a) is closed under subtraction.
  - Furthermore, if $t \in R$ then we have $t(ra) = (tr)a$, so since R is commutative, (a) is closed under multiplication by arbitrary elements of R. Thus, (a) is an ideal.

3.2.2 Quotient Rings

- Now that we have discussed ideals, we can use them to study residue classes, and thereby discuss construct "quotient rings".

- Definition: If I is an ideal of the ring R, then we say a is congruent to b modulo I, written $a \equiv b \pmod{I}$, if $a - b \in I$.
  - As in $\mathbb{Z}$ and $F[x]$, congruence modulo I is an equivalence relation that respects addition and multiplication. The proofs are the same as in $\mathbb{Z}$ and $F[x]$, once we make the appropriate translations from "divisibility" to "containment in I".

- Proposition (Ideal Congruences): Let I be an ideal of R and $a, b, c, d \in R$. Then the following are true:
  1. $a \equiv a \pmod{I}$.
     - Proof: Since $a - a = 0 \in I$, the statement is immediate.
  2. $a \equiv b \pmod{I}$ if and only if $b \equiv a \pmod{I}$.
     - Proof: If $a - b \in I$ then $-(a - b) = b - a \in I$ since I is closed under additive inverses, and conversely if $b - a \in I$ then so is $-(b - a) = a - b$.
  3. If $a \equiv b \pmod{I}$ and $b \equiv c \pmod{I}$, then $a \equiv c \pmod{I}$.

  - Proof: We are given $a - b \in I$ and $b - c \in I$, so since I is closed under addition, we see $(a-b)+(b-c) = a - c \in I$.

4. If $a \equiv b \pmod{I}$ and $c \equiv d \pmod{I}$, then $a + c \equiv b + d \pmod{I}$.
   - Proof: We are given $a - b \in I$ and $c - d \in I$, so since I is closed under addition, we see $(a-b)+(c-d) = (a+c) - (b+d) \in I$.
5. If $a \equiv b \pmod{I}$ and $c \equiv d \pmod{I}$, then $ac \equiv bd \pmod{I}$.
   - Proof: We are given $a - b \in I$ and $c - d \in I$. Then since I is closed under arbitrary left and right multiplication, we see that $(a-b)c$ and $b(c-d)$ are also in I. Hence $ac - bd = (a-b)c + b(c-d)$ is also in I since I is closed under addition.

- Now we can define residue classes:
- Definition: If I is an ideal of the ring R, then for any $a \in R$ we define the residue class of a modulo I to be the set $\overline{a} = a + I = \{a + x : x \in I\}$. This set is also called the coset of I represented by a.
  - We will use the notation $\overline{a}$ and $a+I$ interchangeably. (The latter is intended to evoke the idea of "adding" a to the set I.)
  - We observe, as with our previous examples of residue classes, that any two residue classes are either disjoint or identical and that they partition R: specifically, $\overline{a} = \overline{b}$ if and only if $a \equiv b \pmod{I}$ if and only if $a - b \in I$.
- All that remains is to verify that the residue classes form a ring, in the same way as in $\mathbb{Z}$ and $F[x]$:
- Theorem (Quotient Rings): Let I be an ideal of the ring R. Then the collection of residue classes modulo I forms a ring, denoted R/I (read as "R mod I"), under the operations $\overline{a} + \overline{b} = \overline{a+b}$ and $\overline{a} \cdot \overline{b} = \overline{ab}$. (This ring is called the quotient ring of R by I.) If R is commutative then so is R/I, and likewise if R has a 1 then so does R/I.
  - Remark: The notation R/I is intended to emphasize the idea that I represents a single element (namely, $\overline{0}$) in the quotient ring R/I, and the other elements in R/I are "translates" of I. In this way, R/I is the ring obtained from R by "collapsing" or "dividing out" by I, whence the name "quotient ring".
  - The proof of this fact is exactly the same as in the cases of $\mathbb{Z}$ and $F[x]$, and only requires showing that the operations are well-defined.
  - Proof: First we must show that the addition and multiplication operations are well-defined: that is, if we choose different elements $a' \in \bar{a}$ and $b' \in \overline{b}$, the residue class of $a' + b'$ is the same as that of $a + b$, and similarly for the product.
  - To see this, if $a' \in \bar{a}$ then $a' \equiv a \pmod{I}$, and similarly if $b' \in \overline{b}$ then $b' \equiv b \pmod{I}$.
  - Then $a' + b' \equiv a + b \pmod{I}$, so $\overline{a' + b'} = \overline{a+b}$. Likewise, $a'b' \equiv ab \pmod{I}$, so $\overline{a'b'} = \overline{ab}$.
  - Thus, the operations are well-defined.
  - For the ring axioms [R1]-[R6], we observe that associativity, commutativity, and the distributive laws follow immediately from the corresponding properties in R: the additive identity in R/I is $\bar{0}$ and the additive inverse of $\overline{a}$ is $\overline{-a}$.
  - Finally, if R is commutative then so will be the multiplication of the residue classes, and if R has a 1 then the residue class $\overline{1}$ is easily seen to be a multiplicative identity in R/I.
- This general description of "quotient rings" generalizes the two examples we have previously discussed: $\mathbb{Z}/m\mathbb{Z}$ and R/pR where $R = F[x]$.
  - To be explicit, $\mathbb{Z}/m\mathbb{Z}$ is the quotient of $\mathbb{Z}$ by the ideal $m\mathbb{Z}$, while $F[x]/p$ is the quotient of the polynomial ring $F[x]$ by the principal ideal (p) consisting of all multiples of p.
  - It is not hard to see that the integer congruence $a \equiv b \pmod{m}$, which we originally defined as being equivalent to the statement $m|(b-a)$, is the same as the congruence $a \equiv b \pmod{I}$ where I is the ideal $m\mathbb{Z}$, since $b - a \in m\mathbb{Z}$ precisely when $b - a$ is a multiple of m.

- Here are some additional examples of quotient rings:
- Example: If R is any ring, the quotient ring of R by the zero ideal, namely $R/0$, is (isomorphic to) R itself, while the quotient ring of R by itself, namely R/R, is (isomorphic to) the trivial ring $\{0\}$.
- Example: In $R = \mathbb{Z}[x]$, with I consisting of all multiples of $x^2 + 1$, describe the structure of the quotient ring R/I.
  - It is easy to see that I is an ideal of R, since it is a subring that is closed under arbitrary multiplication by elements of R.
  - From our discussion of polynomial rings, we know that the residue classes in R/I are represented uniquely by residue classes of the form $\overline{a + bx}$ where $a, b \in \mathbb{Z}$. Note that in this quotient ring, we have $\overline{x}^2 + \overline{1} = \overline{0}$, which is to say, $\overline{x}^2 = -\overline{1}$.
  - The addition in this quotient ring is given by $\overline{a + bx} + \overline{c + dx} = \overline{(a + c) + (b + d)x}$ while the multiplication is given by $\overline{a + bx} \cdot \overline{c + dx} = \overline{(ac - bd) + (ad + bc)x}$, which follows from the distributive law and the fact that $\overline{x}^2 = -\overline{1}$.
  - In this case, the quotient ring is isomorphic to the ring of Gaussian integers $\mathbb{Z}[i]$, with the isomorphism $\varphi : R/I \to \mathbb{Z}[i]$ given by $\varphi(\overline{a + bx}) = a + bi$.
- Example: In $R = \mathbb{Z}/8\mathbb{Z}$, with $I = \{0, 4\}$, describe the structure of the quotient ring R/I.
  - It is easy to see that I is an ideal of R, since it is a subring that is closed under arbitrary multiplication by elements of R. (Indeed, it is the principal ideal generated by 4.)
  - Since each residue class contains 2 elements, and R has 8 elements in total, there are four residue classes. With this observation in hand, it is not hard to give a list: $\overline{0} = I = \{0, 4\}$, $\overline{1} = 1 + I = \{1, 5\}$, $\overline{2} = 2 + I = \{2, 6\}$, and $\overline{3} = 3 + I = \{3, 7\}$.
  - Notice, for example, that in the quotient ring R/I, we have $\overline{1} + \overline{3} = \overline{0}$, $\overline{2} \cdot \overline{2} = \overline{0}$, and $\overline{2} \cdot \overline{3} = \overline{2}$: indeed, we can see that the structure of R/I is exactly the same as $\mathbb{Z}/4\mathbb{Z}$ (the labelings of the elements are even the same).
- Example: In the polynomial ring $R = \mathbb{Z}[x]$, with I consisting of the polynomials with even constant term (i.e., the polynomials of the form $2a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ for integers a_i), describe the structure of the quotient ring R/I.
  - We observe that there are only two residue classes, namely $\overline{0}$ and $\overline{1}$: to see this observe that $p(x) \in \overline{0}$ when the constant term of p is even, and $p(x) \in \overline{1}$ when the constant term of p is odd.
  - Then it is fairly easy to see that the structure of this quotient ring is the same as $\mathbb{Z}/2\mathbb{Z}$ (or more formally, it is isomorphic to $\mathbb{Z}/2\mathbb{Z}$), since $\overline{1} + \overline{1} = \overline{0}$.

3.2.3 Homomorphisms and Quotient Rings

- Although homomorphisms and quotient rings may not immediately appear to be connected, in fact they are quite deeply related.
  - To begin, observe that if $\varphi : R \to S$ is a ring homomorphism, then the kernel of φ is an ideal of R.
  - In fact, we proved this fact earlier when we introduced the kernel, but let us remark again: if $x \in \ker\varphi$ and $r \in R$, then $\varphi(rx) = \varphi(r)\varphi(x) = \varphi(r)0 = 0$ and likewise $\varphi(xr) = \varphi(x)\varphi(r) = 0\varphi(r) = 0$.
  - Thus, we can use homomorphisms to construct new ideals.
  - Equally importantly, we can also do the reverse: we can use ideals to construct homomorphisms.
  - The key observation in this direction is that the map $\varphi : R \to R/I$ associating a ring element to its residue class (i.e., with $\varphi(a) = \overline{a}$) is a ring homomorphism.
  - Indeed, the two parts of the definition of homomorphism were precisely the properties we arranged for the residue classes modulo I to possess: $\varphi(a + b) = \overline{a + b} = \overline{a} + \overline{b} = \varphi(a) + \varphi(b)$ and $\varphi(a \cdot b) = \overline{a \cdot b} = \overline{a} \cdot \overline{b} = \varphi(a) \cdot \varphi(b)$.

    - Furthermore, the kernel of this map φ is, by definition, the set of elements in R with $\varphi(r) = \overline{0}$, which is to say, the set of elements $r \in I$.
    - Thus, we see that kernels of homomorphisms and ideals are precisely the same things!

- Let us summarize these observations:

- Proposition (Projection Homomorphisms): If I is an ideal of R, then the map $\varphi : R \to R/I$ defined by $\varphi(a) = \overline{a} = a + I$ is a surjective ring homomorphism called the projection homomorphism from R to R/I.
    - Proof: We have $\varphi(a+b) = \overline{a+b} = \overline{a} + \overline{b} = \varphi(a) + \varphi(b)$ and $\varphi(a \cdot b) = \overline{a \cdot b} = \overline{a} \cdot \overline{b} = \varphi(a) \cdot \varphi(b)$, so φ is a homomorphism.
    - Furthermore, φ is surjective, essentially by definition: any residue class in R/I is of the form $\overline{a}$ for some $a \in R$, and then $\varphi(a) = \overline{a}$.

- The next natural question to ask is: if $\varphi : R \to S$ is a homomorphism with kernel I, what can we say about the structure of R/I?
    - For example, if $R = \mathbb{Q}[x]$ and $\varphi : R \to \mathbb{R}$ is defined by $\varphi(p) = p(0)$, then it is easy to see that φ is a homomorphism.
    - Furthermore, the kernel of φ is the ideal I of $\mathbb{Q}[x]$ consisting of the polynomials divisible by x, while the image of φ is the set of rational numbers.
    - Then it is easy to see (from our description of the kernel) that R/I is precisely the same as R/xR, and from the division algorithm for polynomials we know that the residue classes are represented by the polynomials of degree 0 in $\mathbb{Q}[x]$; namely, the constant polynomials $\overline{c}$ for $c \in \mathbb{Q}$.
    - But now notice that the structure of R/I (namely, of $\mathbb{Q}$) is exactly the same as the structure as the image of φ. More formally, these two rings are isomorphic, with an isomorphism given by identifying a residue class $\overline{c}$ with the rational number c.
    - This relabeling can, equivalently, be thought of as being done via the homomorphism φ: we associate the residue class $\overline{c}$ in R/I with the rational number $\varphi(\overline{c}) = c$.
    - In other words: φ gives an isomorphism between $R/\ker\varphi$ and the image $\operatorname{im}\varphi$.

- Theorem (First Isomorphism Theorem): If $\varphi : R \to S$ is a homomorphism of rings, then $R/\ker\varphi$ is isomorphic to $\operatorname{im}\varphi$.
    - Intuitively, φ is a surjective homomorphism $\varphi : R \to \operatorname{im}\varphi$. To turn it into an isomorphism, we must "collapse" its kernel to a single element: this is precisely what the quotient ring $R/\ker\varphi$ represents.
    - Proof: Let $I = \ker\varphi$. We use φ to construct a map $\psi : R/I \to \operatorname{im}\varphi$, and then show that it is injective and surjective.
    - The map is defined as follows: for any residue class $\overline{r} \in R/I$, we define $\psi(\overline{r}) = \varphi(r)$.
    - We must verify that this map ψ is well-defined, so suppose that r' is some other representative of the residue class $\overline{r}$: then $r' - r \in I$, so $\varphi(r' - r) = 0$ and thus $\varphi(r') = \varphi(r)$.
    - Thus, $\psi(\overline{r'}) = \varphi(r') = \varphi(r) = \psi(\overline{r})$, so the map ψ is well-defined.
    - It is then easy to see ψ is a homomorphism, since $\psi(\overline{r} + \overline{s}) = \varphi(r+s) = \varphi(r) + \varphi(s) = \psi(\overline{r}) + \psi(\overline{s})$ and likewise $\psi(\overline{r} \cdot \overline{s}) = \varphi(r \cdot s) = \varphi(r) \cdot \varphi(s) = \psi(\overline{r}) \cdot \psi(\overline{s})$.
    - Next, we see that $\psi(\overline{r}) = 0$ precisely when $\varphi(r) = 0$, which is to say $r \in \ker(\varphi)$, so that $\overline{r} = \overline{0}$. Thus, the only element in $\ker\psi$ is $\overline{0}$, so ψ is injective.
    - Finally, if s is any element of $\operatorname{im}\varphi$, then by definition there is some $r \in R$ with $\varphi(r) = s$: then $\psi(\overline{r}) = s$, meaning that ψ is surjective.
    - Since ψ is a homomorphism that is both injective and surjective, it is an isomorphism.

- By using the first isomorphism theorem, we can construct isomorphisms of rings.
    - In order to show that R/I is isomorphic to a ring S, we search for a surjective homomorphism $\varphi : R \to S$ whose kernel is I.

- Example: If R is any commutative ring, show that $R[x]/(x)$ is isomorphic to R.
  - Let $\varphi : R[x] \to R$ be the "evaluation at 0" homomorphism $\varphi(p) = p(0)$. This map is clearly surjective since for any $r \in R$ we have $\varphi(r) = r$.
  - Furthermore, the kernel of this homomorphism is precisely the collection of polynomials $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ with $p(0) = 0$, which is easily seen to be the ideal $I = (x)$ consisting of polynomials divisible by x.
  - Thus, by the first isomorphism theorem, for $I = (x)$ we have $R[x]/I \cong R$.
- Example: Show that $\mathbb{Z}/12\mathbb{Z}$ is isomorphic to $(\mathbb{Z}/3\mathbb{Z}) \times (\mathbb{Z}/4\mathbb{Z})$.
  - We seek a surjective homomorphism $\varphi : \mathbb{Z} \to (\mathbb{Z}/3\mathbb{Z}) \times (\mathbb{Z}/4\mathbb{Z})$ whose kernel is $12\mathbb{Z}$.
  - Once this idea is suggested, it is not hard to come up with a candidate, namely, $\varphi(a) = (a \bmod 3,\ a \bmod 4)$.
  - It is easy to verify that map is a homomorphism (since the individual maps of reduction mod 3 and reduction mod 4 are homomorphisms) and it is likewise fairly easy to see that the map is surjective by checking that the images of 0, 1, ... , 11 represent all of the elements in $(\mathbb{Z}/3\mathbb{Z}) \times (\mathbb{Z}/4\mathbb{Z})$.
  - Finally, the kernel of the map consists of all integers a with $\varphi(a) = (0, 0)$, which is equivalent to saying $a \equiv 0 \pmod 3$ and $a \equiv 0 \pmod 4$, so that $3|a$ and $4|a$: thus, the kernel is precisely $12\mathbb{Z}$.
  - Therefore, by the first isomorphism theorem applied to this map φ, we conclude that $\mathbb{Z}/12\mathbb{Z}$ is isomorphic to $(\mathbb{Z}/3\mathbb{Z}) \times (\mathbb{Z}/4\mathbb{Z})$.
  - Remark: In fact, we could have avoided checking surjectivity explicitly by instead observing that the first isomorphism theorem yields an injective homomorphism $\psi : \mathbb{Z}/12\mathbb{Z} \to (\mathbb{Z}/3\mathbb{Z}) \times (\mathbb{Z}/4\mathbb{Z})$, which must therefore also be surjective since there are 12 elements in both sets.

3.3 Properties of Ideals

- Now that we have established basic properties of ideals, homomorphisms, and quotient rings, we embark on a deeper study of these topics.

3.3.1 The Isomorphism Theorems

- We begin by discussing several fundamental theorems about rings, subrings, and ideals that are collectively known as the "isomorphism theorems". We have already proven the first one:
- Theorem (First Isomorphism Theorem): If $\varphi : R \to S$ is a homomorphism of rings, then $R/\ker\varphi$ is isomorphic to $\operatorname{im}\varphi$.
- Theorem (Second Isomorphism Theorem): If A is a subring of R and B is an ideal of R, then $A + B = \{a + b : a \in A, b \in B\}$ is a subring of A, $A \cap B$ is an ideal of A, and $(A + B)/B$ is isomorphic to $A/(A \cap B)$.
  - Proof: Clearly $A + B$ contains 0 and $(a + b) - (a' + b') = (a - a') + (b - b')$ so it is also closed under subtraction. For multiplication, we observe $(a + b)(a' + b') = aa' + ba' + ab' + bb'$: the first term is in A since A is a subring, while the other three terms are in B (hence so is their sum) since B is an ideal.
  - For the last statement, consider the map $\varphi : A \to (A + B)/B$ defined by $\varphi(a) = a + B$. This map is well-defined and a homomorphism by the basic properties of quotient rings, and it is surjective since for any class $r + B$ in $(A + B)/B$ for some $r = a + b \in A + B$, we have $\varphi(a) = a + B = r + B$.
  - The kernel of the map φ consists of all $a \in A$ with $a + B = 0 + B$, which is (by definition) equivalent to saying $a \in B$: thus, $\ker\varphi = A \cap B$. In particular, $A \cap B$ is an ideal since it is a kernel of a homomorphism.
  - Thus, by applying the first isomorphism theorem to φ, we see that the rings $A/(A \cap B)$ and $(A + B)/B$ are isomorphic, as claimed.
- Theorem (Third Isomorphism Theorem): If I and J are ideals of R with $I \subseteq J$, then J/I is an ideal of R/I and $(R/I)/(J/I)$ is isomorphic to R/J.

  - Proof: Define the map $\varphi : R/I \to R/J$ given by setting $\varphi(r + I) = r + J$. This map is well-defined because if $r' + I = r + I$, then since J contains I, we also have $r' + J = r + J$, and it is also surjective since for any class $r + J$ in R/J, we clearly have $\varphi(r + I) = r + J$.
  - Furthermore, φ is a homomorphism by the basic properties of quotient rings, since for example $\varphi((r_1 + r_2) + I) = (r_1 + r_2) + J = (r_1 + J) + (r_2 + J) = \varphi(r_1 + I) + \varphi(r_2 + I)$, which shows that φ is additive because $(r_1 + I) + (r_2 + I) = (r_1 + r_2) + I$.
  - Likewise, since $(r_1 + I)(r_2 + I) = r_1r_2 + I$, we see that $\varphi(r_1r_2 + I) = r_1r_2 + J = (r_1 + J)(r_2 + J) = \varphi(r_1 + I)\varphi(r_2 + I)$ and so φ is multiplicative.
  - The kernel of the map φ consists of all $r + I$ in R/I with the property that $r + J = 0 + J$, which is equivalent to saying $r \in J$: thus, $\ker \varphi$ consists of the classes of the form $r + I$ for $r \in J$; this is simply another way of saying that $\ker \varphi = J/I$.
  - Finally, by applying the first isomorphism theorem to φ, we see that the rings $(R/I)/(J/I)$ and R/J are isomorphic, as claimed.

- Example: Inside $R = \mathbb{Z}[x]$, let I be the ideal of all polynomials with zero constant term and J be the ideal of all polynomials with even constant term.

  - As we have already mentioned, both I and J are ideals of R, and clearly $I \subseteq J$.
  - Furthermore, R/I is isomorphic to $\mathbb{Z}$ (per the division algorithm), and J/I is isomorphic to $2\mathbb{Z}$ (the residue classes are represented by the even integers). Also, R/J is isomorphic to $\mathbb{Z}/2\mathbb{Z}$ (since the residue classes are $\overline{0}$ and $\overline{1}$).
  - Then indeed $(R/I)/(J/I) \cong \mathbb{Z}/2\mathbb{Z} \cong R/J$, as claimed.

- Theorem (Fourth/Lattice Isomorphism Theorem): If I is an ideal of R, then there is an inclusion-preserving bijection between subrings A of R containing I and the subrings $\overline{A} = A/I$ of R/I. Furthermore, a subring A of R containing I is an ideal of R if and only if A/I is an ideal of R/I.

  - Proof: We showed during the proof of the second isomorphism theorem that if A contains I then I is an ideal of A, so the association of A with $\overline{A} = A/I$ is well-defined. Conversely, if S is a subring of R/I, then the set $A = \{r \in R : r + I \in S\}$ is the unique subring of R containing I with the property that $A/I = S$.
  - Furthermore, if B is a subring containing A, then $\overline{B} = B + I$ contains $\overline{A} = A + I$, so the association preserves containment.
  - For the statements about ideals, we showed during the proof of the third isomorphism theorem that if J is an ideal containing I then J/I is an ideal of R/I. Conversely, if J/I is an ideal of R/I, then for any $r \in R$ and $x \in J$ we have $r(x + I) \in J/I$, and this is equivalent to saying that $rx \in J$: thus, J is an ideal of R (since it is already a subring, per the above).

- Example: For $R = \mathbb{Z}$ and $I = 10\mathbb{Z}$, identify the ideals of R containing I and verify that they all yield ideals of R/I.

  - The ideals of R containing I are $\mathbb{Z}$, $2\mathbb{Z}$, $5\mathbb{Z}$, and $10\mathbb{Z}$.
  - The corresponding ideals of $R/I = \mathbb{Z}/10\mathbb{Z}$ are $\mathbb{Z}/10\mathbb{Z}$, $2\mathbb{Z}/10\mathbb{Z} = \{\overline{0}, \overline{2}, \overline{4}, \overline{6}, \overline{8}\}$, $5\mathbb{Z}/10\mathbb{Z} = \{\overline{0}, \overline{5}\}$, and $10\mathbb{Z}/10\mathbb{Z} = \{\overline{0}\}$.
  - As claimed, each of these is indeed an ideal of $\mathbb{Z}/10\mathbb{Z}$.

3.3.2 Generation of Ideals

- In order to study the structure of ideals, we would like a simpler way to describe them. A convenient way is to describe ideals as being "generated" by subsets of a ring:

  - If R is a ring with 1 and A is a subset of R, we would like to define "the ideal generated by A" to be the smallest ideal containing A.

- A priori, it is not obvious that there is such a smallest ideal. However, since the intersection of any nonempty collection of ideals is also an ideal, and since A is contained in at least one ideal (namely the whole ring R), we can equivalently define (A) to be the intersection of all ideals containing A.
- In a similar way, we could define the left ideal generated by A to be intersection of all left ideals containing A, and we could define the right ideal generated by A to be the intersection of all right ideals containing A.
- However, although the above analysis clearly indicates that these definitions are well-posed, we have not actually described what these ideals are.
- If I is the left ideal generated by A, then if $a_1, a_2, \dots, a_n$ are any elements of A, we see that I must contain the elements $r_1 a_1$, $r_2 a_2$, ... , $r_n a_n$ for any $r_i \in R$ and hence also contain their sum.
- On the other hand, if we let S be the set of elements of the form $r_1 a_1 + r_2 a_2 + \cdots + r_n a_n$ for any $a_i \in A$ and $r_i \in R$ (and some $n \geq 0$), then it is easy to see that S is a subring that is closed under left multiplication by elements of R, so S is a left ideal.
- Furthermore, since R contains 1, S contains A, and so by the discussion above, we see that S is a left ideal containing A, hence must actually be the left ideal generated by A.
- In a similar way, the right ideal generated by A consists of the elements of the form $a_1 r_1 + a_2 r_2 + \cdots + a_n r_n$.
- The two-sided ideal generated by A must contain all elements of both forms, but this is not sufficient: indeed, the two-sided ideal must contain elements of the form ras for $r, s \in R$, so the correct definition in this case is the set of elements of the form $r_1 a_1 s_1 + r_2 a_2 s_2 + \cdots + r_n a_n s_n$ for any $a_i \in A$ and $r_i, s_i \in R$ (and some $n \geq 0$).

- Proposition (Generation of Ideals): Let R be a ring with 1 and A be a subset of R. Then the set $RA = \{r_1 a_1 + \cdots + r_n a_n : r_i \in R \text{ and } a_i \in A\}$ is the smallest left ideal containing A, the set $AR = \{a_1 r_1 + a_2 r_2 + \cdots + a_n r_n : r_i \in R \text{ and } a_i \in A\}$ is the smallest right ideal containing A, and the set $(A) = RAR = \{r_1 a_1 s_1 + r_2 a_2 s_2 + \cdots + r_n a_n s_n : r_i, s_i \in R \text{ and } a_i \in A\}$ is the smallest ideal containing A.
  - We will also refer to RA, AR, and (A) as the left ideal generated by A, the right ideal generated by A, and the (two-sided) ideal generated by A, respectively.
  - Note of course that if R is commutative, then $(A) = AR = RA = \{r_1 a_1 + \cdots + r_n a_n : r_i \in R \text{ and } a_i \in A\}$.
  - Proof: As noted above, any left ideal containing A must contain RA, any right ideal containing A must contain AR, and any two-sided ideal containing A must contain (A).
  - Furthermore, since $1 \in R$, each of AR, RA, and (A) contains A.
  - Also, RA, AR, and (A) all contain 0 and are closed under subtraction and multiplication, so they are each subrings.
  - Furthermore, RA is closed under left multiplication, AR is closed under right multiplication, and (A) is closed under both, so they are a left ideal, a right ideal, and a two-sided ideal respectively.
  - Then since each of these is the appropriate type of ideal, by the first observation, we conclude that RA is the smallest left ideal containing A, that AR is the smallest right ideal containing A, and that (A) is the smallest two-sided ideal containing A.

- The simplest class of ideals are those generated by a finite set, and (in particular) those generated by a single element:

- Definition: If R is a ring with 1, we say an ideal I is finitely generated if I is generated by a finite set, and we say I is principal if I is generated by a single element. Thus, a finitely generated ideal has the form $I = (a_1, a_2, \dots, a_n)$, while a principal ideal has the form $I = (a)$.
  - Note that the definition for "principal ideal" extends the one we gave before for commutative rings, since if R is commutative then $(a) = Ra = \{ra : r \in R\}$.
  - If R is not commutative, however, then (a) is the set of elements of the form $r_1 a s_1 + r_2 a s_2 + \cdots + r_n a s_n$ for $r_i, s_i \in R$. (Note in particular that (a) is not just the elements of the form ras for $r, s \in R$, since the sum of two such elements need not also be of that form.)

- Example: If R is any ring with 1, then $R = (1)$ is principal. Likewise, the zero ideal $0 = (0)$ is also principal.

- Example: In $\mathbb{Z}$, for any integer n we have $(n) = n\mathbb{Z}$. Since every ideal of $\mathbb{Z}$ is of the form $n\mathbb{Z}$, we see that every ideal of $\mathbb{Z}$ is principal.
  - Also, we remark that the notation $n\mathbb{Z}$ we have already used is consistent with the definition above. (The same is true for the notation pR for $R = F[x]$.)
  - We also remark that if a and b are integers with greatest common divisor d, then $(a, b) = (d)$: this follows from the pair of observations that a and b are both contained in (d) so that $(a, b) \subseteq (d)$, and that $d = xa + yb$ for some integers x and y by the Euclidean algorithm, so that d is contained in (a, b).
  - Indeed, as a reflection of this fact, many authors write (a, b) to denote the greatest common divisor of a and b.
- Since principal ideals are the easiest to describe, it is often useful to try to determine whether a particular ideal is principal (though this task is not always so easy!):
- Example: Show that the ideal $I = (2, x)$ in $\mathbb{Z}[x]$ is not principal.
  - Note that $I = \{2p(x) + xq(x) \ : \ p, q \in \mathbb{Z}[x]\}$ is the collection of polynomials in $\mathbb{Z}[x]$ with even constant term.
  - If I were principal and generated by some polynomial $r(x)$, then every polynomial in I would be divisible by $r(x)$. Hence, in particular, $r(x)$ would divide 2, so since 2 is a constant polynomial and a prime number, $r(x)$ would have to be one of $\{\pm 1, \pm 2\}$.
  - However, since $r(x)$ must also divide x, the only possibility is that $r(x)$ would be either 1 or -1. But it is easy to see that the ideal generated by 1 (or -1) is all of $\mathbb{Z}[x]$, so $r(x)$ cannot be 1 or -1, since $I \neq \mathbb{Z}[x]$.
  - Thus, there is no possible choice for r, so I is $\boxed{\text{not principal}}$. (Of course, it is still finitely generated!)
- Example: Determine whether or not the ideal $I = (2, 1 + \sqrt{-5})$ in $\mathbb{Z}[\sqrt{-5}]$ is principal.
  - Suppose this ideal were principal with generator $r = a + b\sqrt{-5}$ in $\mathbb{Z}[\sqrt{-5}]$.
  - Then r would necessarily divide 2, meaning that $2 = rs$ for some $s \in \mathbb{Z}[\sqrt{-5}]$. By taking norms, we see that $4 = N(2) = N(r)N(s)$.
  - Likewise, since r divides $1 + \sqrt{-5}$, we would have $1 + \sqrt{-5} = rt$ for some $t \in \mathbb{Z}[\sqrt{-5}]$, so by taking norms we would have $6 = N(1 + \sqrt{-5}) = N(r)N(t)$.
  - Since $N(r) = a^2 + 5b^2$ is a nonnegative integer, we see that $N(r)$ must divide both 4 and 6, hence is either 1 or 2. However, it is easy to see that there are no integer solutions to $a^2 + 5b^2 = 2$, and the only elements of norm 1 are 1 and -1.
  - As in the examples above, the ideal generated by 1 (or -1) is all of $\mathbb{Z}[\sqrt{-5}]$, but $(2, 1 + \sqrt{-5}) \neq \mathbb{Z}[\sqrt{-5}]$ since every element $a + b\sqrt{-5}$ in the ideal has $a + b$ even.
  - Thus, I is $\boxed{\text{not principal}}$.
- Example: Determine whether the ideal $I = (x^3, x + 3)$ in $\mathbb{Q}[x]$ is principal.
  - In the same way as in the example above, if I were principal and generated by a polynomial $r(x)$, then every polynomial in I would be divisible by r.
  - Here, since x^3 and $x + 3$ are relatively prime, we can see that any generator would necessarily divide their gcd, which is 1.
  - In fact, 1 is a generator of I: via the Euclidean algorithm, we can see that $1 = -\frac{1}{27}x^3 + (\frac{1}{3} - \frac{1}{9}x + \frac{1}{27}x^2)(x + 3)$, and so since both x^3 and $x + 3$ are in I, we see that 1 is also in I.
  - Then since 1 is in I, so is $p(x) \cdot 1 = p(x)$ for any $p(x) \in \mathbb{Q}[x]$, meaning that in fact $I = \mathbb{Q}[x]$ and I indeed $\boxed{\text{is principal}}$ (and generated by 1).
- We can in fact generalize the argument from the last example above:

- Proposition (Ideals of $F[x]$): If F is a field, then every ideal in $F[x]$ is principal.
  - Proof: Let I be an ideal of $F[x]$. If I is the zero ideal we are done, so assume I contains a nonzero element.
  - We claim that $I = (d)$, where d is the monic greatest common divisor of all the elements in I. (Equivalently, d is the monic polynomial of largest degree dividing all the elements of I: such a polynomial must exist by the well-ordering axiom.)
  - If d divides every polynomial in I, then clearly $I \subseteq (d)$.
  - Conversely, since d is the gcd, by the Euclidean algorithm and the well-ordering axiom we can write $d = x_1p_1 + x_2p_2 + \cdots + x_np_n$ for some polynomials $x_i \in F[x]$ and $p_i \in I$: then we see that $d \in I$, and hence $(d) \subseteq I$. Thus, $I = (d)$ is principal as claimed.
- As we also saw above, when R is a ring with 1, then 1 is a generator of R. We can likewise generalize this statement:
- Proposition (Ideals and Units): If I is an ideal of the ring R with 1, then $I = R$ if and only if I contains a unit.
  - Proof: If $I = R$ then certainly I contains a unit (namely, 1).
  - Conversely, if $u \in I$ is a unit with $ur = 1$, then since I is an ideal we have $1 = ur \in I$, and then for any $s \in R$, the element $s = 1s$ is also in I, and so $I = R$.
- Since every nonzero element in a field is a unit, we immediately see that the only nonzero ideal of a field is the full ring. The converse is also true:
- Corollary (Ideals of Fields): A commutative ring R with 1 is a field if and only if the only ideals of R are 0 and R.
  - Proof: If F is a field and I is any nonzero ideal, then I contains some nonzero element r. Since F is a field, r is a unit, and so by the proposition above, $I = R$.
  - Conversely, if the only ideals of R are 0 and R, let $r \in R$ be any nonzero element. Then (r) contains $r \neq 0$ so it cannot be the zero ideal, so we must have $(r) = R$.
  - By the previous proposition, this means (r) contains 1: then $rs = 1$ for some $s \in R$, so r is a unit. Hence every nonzero element of R is a unit, so R is a field as claimed.
  - Remark: In fact, the proof above shows that the only ideals of a division ring R are 0 and R. However, the converse direction does not hold: there exist noncommutative rings R with zero divisors whose only ideals are 0 and R. (One such ring is $M_{2\times 2}(\mathbb{R})$, although this is not completely trivial to prove.)

3.3.3 Maximal and Prime Ideals

- An important class of ideals are those that are "maximal" under inclusion (i.e., which are not contained in any other ideal except the full ring):
- Definition: If R is a ring, a maximal ideal of R is an ideal $M \neq R$ with the property that the only ideals of R containing M are M and R.
  - Example: If F is a field, then since the only ideals of F are 0 and F, the zero ideal is a maximal ideal of F.
  - Example: In $\mathbb{Z}$, the ideal $m\mathbb{Z}$ is contained in $n\mathbb{Z}$ precisely when n divides m. Accordingly, the maximal ideals of $\mathbb{Z}$ are precisely the ideals of the form $p\mathbb{Z}$, where p is a prime.
  - Non-example: The ideal (x) is not a maximal ideal of $\mathbb{Z}[x]$ because it is contained in the proper ideal $(2, x)$.
- A general ring need not possess any maximal ideals.
  - A trivial example is the zero ring, since its only ideal is itself.

- For a less trivial example, let R be the ring of rational numbers with trivial multiplication (i.e., so that $ab = 0$ for any a and b). Since multiplication is trivial, the ideals of R are precisely the sets containing 0 that are closed under subtraction.
- Now suppose that I is any proper ideal of R, and let $S = \{n \in \mathbb{N} : \frac{1}{n} \notin I\}$. If it were true that $\frac{1}{n} \in I$ for every positive integer n, then since I is closed under addition and additive inverses, we would necessarily have $I = R$. Thus, S is a nonempty set of positive integers, so it contains some minimal element d.
- Then define $J = I + (\frac{1}{d})$: it is not hard to verify that J is an ideal properly containing I. If $\frac{1}{d^2}$ were in J, then we could write $\frac{1}{d^2} = x + \frac{a}{d}$ for some $a \in \mathbb{Z}$ and $x \in I$: multiplying through by d yields $\frac{1}{d} = dx + a$, but since dx and a are both in I, this would imply $1/d$ is in I, which is impossible. Thus, J is a proper ideal, and so I is not maximal.

- However, it is true that a ring with 1 must have maximal ideals:
- Theorem (Existence of Maximal Ideals): If R is a ring with 1, then any proper ideal of R is contained in a maximal ideal.
  - Like a number of other general existence theorems (e.g., the proof that every vector space has a basis), this proof requires the (in)famous “axiom of choice” from set theory. The version of the axiom of choice typically used in algebra is known as Zorn’s lemma: if S is a nonempty partially ordered set with the property that every chain in S has an upper bound, then S contains a maximal element[2].
  - Proof: Suppose R is a ring with 1 and I is a proper ideal of R.
  - Let S be the set of all proper ideals of R containing I, partially ordered under inclusion. Since $I \in S$, S is nonempty.
  - If C is any nonempty chain in S, let J be the union of all ideals in C: then $0 \in J$ since 0 is contained in any ideal in C.
  - Furthermore, if $x, y \in J$ and $r \in R$, then by definition $x \in I_i$ and $y \in I_j$ for some I_i and I_j in C. Since $I_i \subseteq I_j$ or $I_j \subseteq I_i$ since C is a chain, it follows that $x - y$, rx, and xr are all in one of I_i or I_j, hence in J. Thus, J is an ideal.
  - Also, if it were true that $J = R$, then the element 1 would be in J. But this is impossible, since by definition J is the union of a collection of proper ideals of R, none of which therefore contains 1.
  - Therefore, J is an upper bound for S. Hence, by Zorn’s lemma, J contains a maximal element, which is therefore a maximal ideal of R that contains I.
- It might initially appear to be difficult to detect whether a particular ideal is maximal. However, by using the isomorphism theorems, it is actually quite easy to detect maximal ideals in commutative rings:
- Proposition (Maximal Ideals and Quotients): If R is a commutative ring with 1, then the ideal M is maximal if and only if R/M is a field.
  - We will remark that this result is *not* true if we drop either of the assumptions on R (i.e., that it is commutative and has a 1).
  - Proof: By the lattice isomorphism theorem, the ideals of R/M are in bijection with the ideals of R containing M: therefore, M is maximal precisely when the only ideals of R/M are 0 and R/M.
  - Furthermore, if R is commutative with 1, then R/M is also a commutative ring with 1, so R/M is a field if and only if the only ideals of R/M are 0 and R/M. Putting these two statements together yields the proposition.
- Corollary: If F is a field, the maximal ideals of $F[x]$ are precisely the principal ideals (p) where p is irreducible.

[2] A partial ordering on a set S a relation $\leq$ such that for any $x, y, z \in S$, (i) $x \leq x$ (ii) $x \leq y$ and $y \leq x$ implies $x = y$, and (iii) $x \leq y$ and $y \leq z$ implies $x \leq z$. If S is a partially-ordered set, a subset C is a chain if for any $x, y \in C$, either $x \leq y$ or $y \leq x$, an upper bound for a subset B is an element $w \in B$ such that $b \leq w$ for all $b \in B$, and a maximal element of a subset B is an element $m \in B$ such that if $x \in B$ has $m \leq x$ then $m = x$.

- Proof: Every ideal of $F[x]$ is principal, and the quotient ring $F[x]/(p)$ is a field if and only if p is irreducible.

- Example: Determine whether the ideal $I = (2, x)$ is a maximal ideal of $R = \mathbb{Z}[x]$.
  - As we have already shown, the quotient ring $R/(2, x)$ is isomorphic to $\mathbb{Z}/2\mathbb{Z}$, which is a field. Thus, I is a maximal ideal of R.

- Example: Determine whether the ideal $I = (2)$ is maximal in $R = \mathbb{Z}[\sqrt{2}]$.
  - In the quotient ring R/I, the residue class $\sqrt{2} + I$ is nonzero, but has the property that $(\sqrt{2} + I)^2 = 2 + I = 0 + I$ is equal to zero.
  - Thus, the quotient ring R/I has zero divisors hence is not a field, meaning that I is not a maximal of R.

- In addition to maximal ideals, we have another important class of ideals in commutative rings:

- Definition: If R is a commutative ring, a prime ideal of R is an ideal $P \neq R$ with the property that for any $a, b \in R$ with ab in P, at least one of a and b is in P.
  - Remark: There is also a definition of "prime ideal" in a noncommutative ring, but it is more complicated (ultimately because the definition above involves products of elements).
  - As naturally suggested by the name, prime ideals are a generalization of the idea of a prime number in $\mathbb{Z}$: for $n > 1$, the ideal $n\mathbb{Z}$ is a prime ideal of $\mathbb{Z}$ precisely when $ab \in n\mathbb{Z}$ implies $a \in n\mathbb{Z}$ or $b \in n\mathbb{Z}$. Equivalently (in the language of divisibility) this means $n|ab$ implies $n|a$ or $n|b$, and this is precisely the condition that n is either a prime number (or zero).
  - Example: The prime ideals of $\mathbb{Z}$ are (0) and the ideals $p\mathbb{Z}$ where p is a prime number.
  - A similar statement holds in $R = F[x]$: the ideal (p) is prime precisely when p is not a unit and $p|ab$ implies $p|a$ or $p|b$, and the latter condition is equivalent to saying that p is either irreducible or zero.
  - Example: The prime ideals of $F[x]$ are (0) and the ideals (p) where p is an irreducible polynomial of positive degree.

- Like with maximal ideals, there is an easy way to test whether an ideal is prime using quotient rings:

- Proposition (Prime Ideals and Quotients): If R is a commutative ring with 1, then the ideal P is prime if and only if R/P is an integral domain.
  - This proof is essentially just a restatement of the definition of a prime ideal using residue classes in the quotient ring using the observation that $r \in P$ if and only if $\overline{r} = \overline{0}$ in R/P.
  - Proof: If R is commutative with 1 and $P \neq R$, then R/P is also commutative with 1, so we need only test for zero divisors.
  - If P is a prime ideal, then $ab \in P$ implies $a \in P$ or $b \in P$. In the quotient ring, this says that $\overline{ab} = \overline{0}$ implies $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$, which is precisely the statement that R/P has no zero divisors.
  - Conversely, if R/P has no zero divisors, then $\overline{ab} = \overline{0}$ implies $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$, which is to say, $ab \in P$ implies $a \in P$ or $b \in P$. Furthermore, since R/P is not the zero ring (since this possibility is excluded by the definition of integral domain), we see $P \neq R$, and therefore P is a prime ideal of R.

- Corollary: A commutative ring with 1 is an integral domain if and only if 0 is a prime ideal.
  - Proof: 0 is prime if and only if the quotient $R/0 \cong R$ is an integral domain.

- Corollary: In a commutative ring with 1, every maximal ideal is prime.
  - Proof: If M is a maximal ideal, then R/M is a field. Every field is an integral domain, so M is a prime ideal.

- Example: Determine whether the ideals (x) and (x^2) in $\mathbb{Z}[x]$ are prime ideals.

- ○ Note that (x) is the kernel of the evaluation homomorphism $\varphi : \mathbb{Z}[x] \to \mathbb{Z}$ given by $\varphi(p) = p(0)$, and this homomorphism is surjective.
- ○ Thus, by the first isomorphism theorem, we see that $\mathbb{Z}[x]/(x)$ is isomorphic to $\mathbb{Z}$. Since $\mathbb{Z}$ is an integral domain, we conclude that (x) is a prime ideal. (Note that it is not maximal, however, since $\mathbb{Z}$ is not a field.)
- ○ On the other hand, by the division algorithm, we see that the residue classes in $\mathbb{Z}[x]/(x^2)$ are of the form $\overline{a+bx}$ where $a, b \in \mathbb{Z}$. Since $\overline{x} \cdot \overline{x} = \overline{0}$ but $\overline{x} \neq \overline{0}$, we see that $\mathbb{Z}[x]/(x^2)$ has zero divisors, and so (x^2) is not a prime ideal.

3.3.4 The Chinese Remainder Theorem

- We now state an important theorem regarding quotient rings by products of ideals. We first require a few preliminary definitions:
- Definition: If R is commutative with 1 and I and J are ideals of R, then the sum $I+J = \{a+b : a \in I, b \in J\}$ is defined to be the set of all sums of elements of I and J, and the product $IJ = \{a_1b_1 + \cdots + a_nb_n, : a_i \in I, b_i \in J\}$ is the set of finite sums of products of an element of I with an element of J.
  - ○ It is not difficult to verify that $I + J$ and IJ are both ideals of R, and that IJ contains the intersection $I \cap J$.
  - ○ We can also speak of the product $I_1 I_2 \cdots I_n$ of more than two ideals, defined as the set of finite sums of products of an element from each of $I_1, I_2, \ldots, I_n$.
- Definition: If R is commutative with 1, the ideals I and J are comaximal if $I + J = R$.
  - ○ Note that $a\mathbb{Z} + b\mathbb{Z} = \mathbb{Z}$ precisely when a and b are relatively prime. (The appropriate notion in general rings is not "primality" but "maximality", so we use the term comaximal rather than coprime.)
- We can now state the theorem:
- Theorem (Chinese Remainder Theorem): Let R be commutative with 1 and $I_1, I_2, \ldots, I_n$ be ideals of R. Then the map $\varphi : R \to (R/I_1) \times (R/I_2) \times \cdots \times (R/I_n)$ defined by $\varphi(r) = (r + I_1, r + I_2, \ldots, r + I_n)$ is a ring homomorphism with kernel $I_1 \cap I_2 \cap \cdots \cap I_n$. If all of the ideals $I_1, I_2, \ldots, I_n$ are pairwise comaximal, then φ is surjective and $I_1 \cap I_2 \cap \cdots \cap I_n = I_1 I_2 \cdots I_n$, and thus $R/(I_1 I_2 \cdots I_n) \cong (R/I_1) \times (R/I_2) \times \cdots \times (R/I_n)$.
  - ○ Proof: First, φ is a homomorphism since $\varphi(a+b) = (a+b+I_1, \ldots, a+b+I_n) = (a+I_1, \ldots, a+I_n) + (b+I_1, \ldots, b+I_n) = \varphi(a) + \varphi(b)$ and similarly $\varphi(ab) = (ab+I_1, \ldots, ab+I_n) = (a+I_1, \ldots, a+I_n) \cdot (b+I_1, \ldots, b+I_n) = \varphi(a)\varphi(b)$.
  - ○ The kernel of φ is the set of elements $r \in R$ such that $\varphi(r) = (0 + I_1, \ldots, 0 + I_n)$, which is equivalent to requiring $r \in I_1$, $r \in I_2$, ... , and $r \in I_n$: thus, $\ker \varphi = I_1 \cap I_2 \cap \cdots \cap I_n$.
  - ○ For the second statement, we will prove the results for two ideals and then deduce the general statement via induction.
  - ○ So suppose I and J are ideals of R and $\varphi : R \to (R/I) \times (R/J)$ has $\varphi(r) = (r + I, r + J)$. We must show that if $I + J = R$, then $I \cap J = IJ$ and φ is surjective.
  - ○ If $I + J = R$ then by definition there exist elements $x \in I$ and $y \in J$ with $x + y = 1$.
  - ○ Then for any $r \in I \cap J$, we can write $r = r(x + y) = rx + yr$, and both rx and yr are in IJ: hence $I \cap J \subseteq IJ$, and since $IJ \subseteq I \cap J$ we conclude $IJ = I \cap J$.
  - ○ Furthermore, for any $a, b \in R$ we can write $ay + bx = a(1-x) + bx = a + (b-a)x$ so $ay + bx \in a + I$, and likewise $ay + bx = ay + b(1-y) = b + (a-b)y \in b + J$.
  - ○ Then $\varphi(ay + bx) = (ay + bx + I, ay + bx + J) = (a + I, b + J)$, and therefore φ is surjective as claimed.
  - ○ Finally, the statement that $R/IJ \cong (R/I) \times (R/J)$ then follows immediately by the first isomorphism theorem. This establishes all of the results for two ideals.
  - ○ For the general statement, we use induction on n: the base case $n = 2$ was done above, and for the inductive step, it is enough to show that the ideals I_1 and $I_2 \cdots I_n$ are comaximal, since then we may write $R/(I_1 I_2 \cdots I_n) \cong (R/I_1) \times (R/I_2 \cdots I_n)$ and apply the induction hypothesis to $R/I_2 \cdots I_n$.

- If I_1 and I_i are comaximal for $2 \le i \le n$, then there exist elements $x_i \in I_1$ and $y_i \in I_i$ such that $x_i + y_i = 1$. Then $1 = (x_2 + y_2)(x_3 + y_3) \cdots (x_n + y_n) \equiv y_2 y_3 \cdots y_n$ modulo I_1. But since $y_2 y_3 \cdots y_n$ is in $I_2 I_3 \cdots I_n$, this means that $I_1 + I_2 I_3 \cdots I_n$ contains 1 and is therefore all of R, as required.

- The name of this theorem comes from its application inside $\mathbb{Z}$ to solving simultaneous modular congruences.
  - Explicitly, if $m_1, m_2, \dots m_n$ are relatively prime positive integers, then $\varphi : \mathbb{Z} \to (\mathbb{Z}/m_1\mathbb{Z}) \times (\mathbb{Z}/m_2\mathbb{Z}) \times \cdots \times (\mathbb{Z}/m_n\mathbb{Z})$ given by $\varphi(a) = (a \bmod m_1, a \bmod m_2, \dots, a \bmod m_n)$ is a surjective homomorphism with kernel $m_1 m_2 \cdots m_n \mathbb{Z}$.
  - The fact that this map is surjective says that the system of simultaneous congruences $x \equiv a_1 \bmod m_1$, $x \equiv a_2 \bmod m_2$, ... , $x \equiv a_n \bmod m_n$ always has a solution in $\mathbb{Z}$. Furthermore, the characterization of the kernel says that the solution is unique modulo $m_1 m_2 \cdots m_n$.
  - Systems of congruences of this form were studied by the ancient Chinese, whence the theorem's name.

- A useful application of the Chinese remainder theorem is to decompose $\mathbb{Z}/m\mathbb{Z}$ as the direct product of other rings when m is composite (some examples of which we have already seen):

- Corollary: If m is a positive integer with prime factorization $m = p_1^{a_1} p_2^{a_2} \cdots p_n^{a_n}$, then $\mathbb{Z}/m\mathbb{Z} \cong (\mathbb{Z}/p_1^{a_1}\mathbb{Z}) \times \cdots \times (\mathbb{Z}/p_n^{a_n}\mathbb{Z})$. In particular, the number of units in $\mathbb{Z}/m\mathbb{Z}$ is $m(1 - 1/p_1)(1 - 1/p_2) \cdots (1 - 1/p_n)$.
  - Proof: The first statement follows from the Chinese remainder theorem along with the observation that if p and q are distinct primes, then the ideals $p^a\mathbb{Z}$ and $q^b\mathbb{Z}$ are comaximal in $\mathbb{Z}$.
  - The second statement follows from the observation that the number of units in $\mathbb{Z}/m\mathbb{Z}$ is the same as the number of units in $(\mathbb{Z}/p_1^{a_1}\mathbb{Z}) \times \cdots \times (\mathbb{Z}/p_n^{a_n}\mathbb{Z})$.
  - Finally, an element of a direct product is a unit if and only if each of its components is a unit, and for any prime p, the units in $\mathbb{Z}/p^a\mathbb{Z}$ are precisely the $p^a - p^{a-1} = p^a(1 - 1/p)$ residue classes that are not divisible by p.
  - Remark: The function $\varphi(m) = m(1 - 1/p_1)(1 - 1/p_2) \cdots (1 - 1/p_n)$ giving the the number of units in $\mathbb{Z}/m\mathbb{Z}$ is called the Euler φ-function.

3.4 Rings of Fractions

- Let R be a commutative ring. Our goal in this section is to discuss a construction for creating "rings of fractions", which (as both a motivating example and a special case) includes the construction of the rational numbers $\mathbb{Q}$ from the integers $\mathbb{Z}$.
  - A natural first attempt is simply to construct symbols of the form $\frac{a}{b}$ where $a, b \in R$, and then define addition and multiplication operations on these symbols.
  - However, even in the case of constructing $\mathbb{Q}$ from $\mathbb{Z}$, complications already arise since it is not possible to divide by zero, and all rational numbers can be written in multiple forms (e.g., $1/2 = 3/6$).
  - Indeed, we say that a/b and c/d are equal (as rational numbers) precisely when $ad = bc$.
  - To make this more precise, we can think of fractions as ordered pairs (a, b) of integers, with the rational numbers then being equivalence classes of these ordered pairs under the relation $(a, b) \sim (c, d)$ when $ad = bc$, and then define $(a, b) + (c, d) = (ad + bc, bd)$ and $(a, b) \cdot (c, d) = (ac, bd)$, per the usual arithmetic rules $\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$ and $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$.
  - This approach also explains one reason why we should not allow 0 in denominators: if we did, then we would have $(0, 0) \sim (a, b)$ for any integers a and b, and then $\sim$ would not be an equivalence relation.
  - More generally, if we want $\sim$ to be an equivalence relation, then we would need $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$ to imply $(a, b) \sim (e, f)$: thus we want $ad - bc = 0$ and $cf - de = 0$ to imply $af - be = 0$.
  - Since $f(ad - bc) + b(cf - de) = (adf - bcf) + (bcf - bde) = d(af - be)$, if $d \neq 0$ we would be able to conclude that $af - be = 0$.

- This calculation suggests that, if we want to extend the construction of fractions to general rings, we should avoid having zero divisors (and zero) in our denominators. Indeed, if $bd = 0$, then we could write $(b, 1) \sim (bd, d) \sim (0, d) \sim (0, 1)$, and so the associated fraction $b/1$ would be equal to $0/1 = 0$.
- Furthermore, the arithmetic rules clearly require the collection of denominators to be closed under multiplication, since we must be able to add and multiply fractions.
- It turns out that these two restrictions are sufficient to allow us to construct a "ring of fractions" with a given set of denominators.

- Theorem (Rings of Fractions): Let R be a commutative ring and D be a nonempty subset of R that is closed under multiplication and does not contain 0 or any zero divisors. Then there is a commutative ring $D^{-1}R$ with 1 consisting of elements of the form $r \cdot d^{-1}$ for $r \in R$ and $d \in D$. This ring $D^{-1}R$ contains (an isomorphic copy of) R as a subring, and every element of D is a unit in $D^{-1}R$.
  - Proof: Let $S = \{(r, d) : r \in R \text{ and } d \in D\}$ and define a relation on S via $(r, d) \sim (s, e)$ precisely when $re = sd$.
  - First we observe that $\sim$ is an equivalence relation on S: clearly $(r, d) \sim (r, d)$ since $rd - rd = 0$, and $(r, d) \sim (s, e)$ implies $(s, e) \sim (r, d)$ since $re - sd = 0$ implies $sd - re = 0$.
  - Furthermore, if $(r, d) \sim (s, e)$ and $(s, e) \sim (t, f)$ then $re - sd = sf - et = 0$, and therefore we see that $e(fr - dt) = f(re - sd) + d(sf - et) = 0$. But since $e \in D$, e is not a zero divisor, and therefore $e(fr - dt) = 0$ implies $fr = dt$, so $(r, d) \sim (t, f)$.
  - Now let $D^{-1}R$ be the set of equivalence classes of S under the relation $\sim$, and write $\frac{r}{d}$ to represent the equivalence class of (r, d). We define the addition and multiplication operations in $D^{-1}R$ to be $\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$ and $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$.
  - In order to show that $D^{-1}R$ is a commutative ring with 1, we must verify that these operations are well-defined and that they satisfy the ring axioms [R1]-[R8].
  - To see $+$ is well-defined: if $\frac{a}{b} = \frac{a'}{b'}$ and $\frac{c}{d} = \frac{c'}{d'}$, then $\frac{a'}{b'} + \frac{c'}{d'} = \frac{a'd' + b'c'}{b'd'}$ and we must show that this equals $\frac{ad + bc}{bd}$. But $(a'd' + b'c')(bd) - (ad + bc)(b'd') = (a'b - ab')dd' - (c'd - cd')bb' = 0$ since $a'b - ab' = c'd - cd' = 0$.
  - To see $\cdot$ is well-defined: if $\frac{a}{b} = \frac{a'}{b'}$ and $\frac{c}{d} = \frac{c'}{d'}$, then $\frac{a'}{b'} \cdot \frac{c'}{d'} = \frac{a'c'}{b'd'}$ and we must show that this equals $\frac{ac}{bd}$. But $(a'c')(bd) - (ac)(b'd') = c'd(a'b - ab') + ab'(c'd - cd') = 0$ since $a'b - ab' = c'd - cd' = 0$.
  - The ring axioms [R1]-[R8] are straightforward calculations (not even requiring the equivalence relation): the additive identity is $\frac{0}{d}$ for any $d \in D$, the additive inverse of $\frac{a}{b}$ is $\frac{-a}{b}$, and the multiplicative identity is $\frac{d}{d}$ for any $d \in D$.
  - Furthermore, we can embed R in $D^{-1}R$ via the map $\iota : R \to D^{-1}R$ with $\iota(r) = \frac{dr}{d}$ for any fixed $d \in D$ (note that this embedding does not actually depend on d, since for any other $d' \in D$ we have $\frac{dr}{d} = \frac{d'r}{d'}$): we have $\iota(a + b) = \frac{d(a + b)}{d} = \frac{d^2(a + b)}{d^2} = \frac{da}{d} + \frac{db}{d} = \iota(a) + \iota(b)$ and likewise $\iota(ab) = \frac{d(ab)}{d} = \frac{d^2(ab)}{d^2} = \frac{da}{d} \cdot \frac{db}{d} = \iota(a)\iota(b)$, so ι is a homomorphism.
  - Furthermore, if $\iota(a) = \frac{0}{d}$ then this means $\frac{ad}{d} = \frac{0}{d}$ whence $ad^2 = 0$ so that $a = 0$ (since d^2 is not a zero divisor). Hence ι is injective, and so $\iota(R)$ is isomorphic to R.
  - Finally, for any $e \in D$, we have $\frac{de}{d} \cdot \frac{d}{de} = \frac{d^2e}{d^2e} = \frac{d}{d}$, so every element of D inside $D^{-1}R$ is a unit, and then any element $\frac{r}{d} \in D^{-1}R$ can be written as $r \cdot d^{-1}$ for $r \in R$ and $d \in D$.

- Using this result, we can show that every integral domain can be viewed naturally as a subset of its "field of fractions":
- Corollary (Fields of Fractions): If R is an integral domain, then R is a subring of its field of fractions $D^{-1}R$, where $D = R\backslash\{0\}$.
  - Proof: If R is an integral domain, then $D = R\backslash\{0\}$ is a multiplicatively closed subset not containing zero or any zero divisors.
  - Then $D^{-1}R$ is a commutative ring with 1 in which every element of D is a unit, which is to say, in which every nonzero element of R is a unit.
  - But since the elements of $D^{-1}R$ are all of the form r/s for $r, s \in R$ and $s \neq 0$, this means that every nonzero element of $D^{-1}R$ is a unit in $D^{-1}R$, so it is a field.
- Here are a few examples of rings and fields of fractions:
  - Example: The field of fractions of $\mathbb{Z}$ is $\mathbb{Q}$.
  - Example: The field of fractions of $\mathbb{Z}[\sqrt{D}]$, or more generally the quadratic integer ring $\mathcal{O}_{\mathbb{Q}(\sqrt{D})}$ is $\mathbb{Q}(\sqrt{D})$.
  - Example: The field of fractions of $2\mathbb{Z}$ is the set of rational numbers of the form $\dfrac{2m}{2n}$, hence is also $\mathbb{Q}$. Notice in particular that although $2\mathbb{Z}$ does not have a multiplicative identity, the fraction $\dfrac{2}{2}$ is a multiplicative identity in its field of fractions.
  - Example: If F is any field, the field of fractions of F is simply F itself.
  - Example: If F is any field, the field of fractions of $F[x]$ consists of elements $\dfrac{p(x)}{q(x)}$ with $q(x) \neq 0$: this is simply the field of rational functions with coefficients in F.
  - Example: If $R = \mathbb{Z}$ and $D = \{1, p, p^2, p^3, \dots\}$ where p is a prime number, the ring of fractions $D^{-1}R$ consists of the rational numbers whose denominator is a power of p. This ring is often denoted $\mathbb{Z}[1/p]$, since it is obtained by "adjoining" the number $1/p$ to $\mathbb{Z}$ (and indeed, it is not hard to see that it is the smallest subring of $\mathbb{Q}$ that contains $1/p$).
- Inside $D^{-1}R$, every element of D is a unit. In fact, $D^{-1}R$ is (in a fairly strong sense) the smallest ring in which this property holds:
- Proposition (Minimality of $D^{-1}R$): Let D be a multiplicatively closed subset of the ring R not containing 0 or any zero divisors. Suppose S is also a commutative ring with 1 and there is an injection $\varphi : R \to S$ with the property that $\varphi(d)$ is a unit for all $d \in D$. Then there is an injection $\Phi : D^{-1}R \to S$ such that $\Phi(r) = \varphi(r)$ for all $r \in R$.
  - Equivalently, this result says that any ring that contains (an isomorphic copy of) R in which every element of D is a unit must actually contain (an isomorphic copy of) all of $D^{-1}R$.
  - In the specific case for fields of fractions, we obtain the following statement: if F is any field that contains (an isomorphic copy of) an integral domain R and every element of R is a unit in F, then F contains (an isomorphic copy of) the field of fractions of R.
  - Proof: Suppose $\varphi : R \to S$ is an injective ring homomorphism such that $\varphi(d)$ is a unit for all $d \in D$.
  - We then define $\Phi : D^{-1}R \to S$ by setting $\Phi(r/d) = \varphi(r)\varphi(d)^{-1}$.
  - To see that Φ is well-defined, suppose that $r'/d' = r/d$, so that $r'd = rd'$. Then $\varphi(r')\varphi(d) = \varphi(r'd) = \varphi(rd') = \varphi(r)\varphi(d')$, so multiplying by $\varphi(d)^{-1}\varphi(d')^{-1}$ yields $\varphi(r')\varphi(d')^{-1} = \varphi(r)\varphi(d)^{-1}$. Finally, we see $\Phi(r'/d') = \varphi(r')\varphi(d')^{-1} = \varphi(r)\varphi(d)^{-1} = \Phi(r/d)$.
  - Finally, Φ is a ring homomorphism, since $\Phi(r/d + s/e) = \varphi(re + ds)\varphi(de)^{-1} = [\varphi(r)\varphi(e) + \varphi(d)\varphi(s)] \cdot \varphi(d)^{-1}\varphi(e)^{-1} = \varphi(r)\varphi(d)^{-1} + \varphi(s)\varphi(e)^{-1} = \Phi(r/d) + \Phi(s/e)$ and $\Phi(r/d \cdot s/e) = \varphi(rs)\varphi(de)^{-1} = \varphi(r)\varphi(s)\varphi(d)^{-1}\varphi(e)^{-1} = [\varphi(r)\varphi(d)^{-1}] \cdot [\varphi(s)\varphi(e)^{-1}] = \Phi(r/d) \cdot \Phi(s/e)$, and Φ is injective since $\Phi(r/d) = 0$ implies $\varphi(r)\varphi(d)^{-1} = 0$ so $\varphi(r) = 0$, and thus $r = 0$ since φ is injective.

Well, you're at the end of my handout. Hope it was helpful.