Homomorphisms, Ideals and Quotients

Revised November 14, 2013

1. Homomorphisms

Let R and S be rings. A function $\theta : R \to S$ is called a *homomorphism* if $\theta(x+y) = \theta(x) + \theta(y)$ and $\theta(xy) = \theta(x)\theta(x)$ for all $x, y \in R$. The *image* of $\theta : R \to S$ is

$$\theta(R) = \{\theta(r) : r \in R\}$$

and the *kernel* of θ is

$$\ker(\theta) = \big\{r \in R : \theta(r) = 0\}.$$

Homomorphisms are the most natural functions to consider in ring theory, just as linear transformations are the most natural functions to consider in linear algebra. This analogy between ring theory and linear algebra will be very helpful to keep in mind. Under this analogy, the following concepts correspond:

Linear Algebra	$\leftrightarrow$	Ring Theory
vector space		ring
subspace		subring
linear transformation		homomorphism
isomorphism		isomorphism
image (range)		image
null space (kernel)		kernel

Examples

The map $\theta : \mathbb{Z} \to \mathbb{Z}_2$ given by

$$\theta(n) = \begin{cases} 0, & \text{if } n \text{ is even;} \\ 1, & \text{if } n \text{ is odd} \end{cases}$$

is a homomorphism. More generally for any $n \in \mathbb{N}$, there is an obvious homomorphism $\mathbb{Z} \to \mathbb{Z}_n$ given by 'reduction modulo n': each integer goes to its congruence class mod n.

The map $\theta : \mathbb{Z}[t] \to \mathbb{R}$ defined by $\theta\big(f(t)\big) = f(\sqrt{2})$ is a homomorphism since for all polynomials $f(t), g(t) \in \mathbb{Z}[t]$ we have

$$\theta\big(f(t)+g(t)\big) = f(\sqrt{2}) + g(\sqrt{2}) = \theta\big(f(t)\big) + \theta\big(g(t)\big); \quad \text{and}$$
$$\theta\big(f(t)g(t)\big) = f(\sqrt{2})g(\sqrt{2}) = \theta\big(f(t)\big)\theta\big(g(t)\big).$$

Its image is $\mathbb{Z}[\sqrt{2}]$ and its kernel consists of all polynomials in $\mathbb{Z}[t]$ divisible by t^2-2. Observe that the map θ may be described as an evaluation map: it evaluates each polynomial at a particular point.

More generally, evaluation at *any* point is a homomorphism: Let $a \in \mathbb{C}$; then evaluation at a gives a homomorphism $\theta_a : \mathbb{Z}[t] \to \mathbb{C}$ given by $\theta_a\big(f(t)\big) = f(a)$. The image of θ_a is

$$\theta_a(\mathbb{Z}[t]) = \big\{\theta_a\big(f(t)\big) \ : \ f(t) \in \mathbb{Z}[t]\big\} = \big\{f(a) \ : \ f(t) \in \mathbb{Z}[t]\big\} = \mathbb{Z}[a].$$

Thus for example, the image of $\theta_{\sqrt{2}} : \mathbb{Z}[t] \to \mathbb{R}$ is $\mathbb{Z}[\sqrt{2}] \subset \mathbb{R}$; the image of $\theta_i : \mathbb{Z}[t] \to \mathbb{C}$ is $\mathbb{Z}[i]$. It takes some thought to see why this agrees with our earlier definition of $\mathbb{Z}[\sqrt{2}]$ and $\mathbb{Z}[i]$; but the point is that evaluating any polynomial $f(t) \in \mathbb{Z}[t]$ at $\sqrt{2}$ always gives a number of the form $a + b\sqrt{2}$ where $a, b \in \mathbb{Z}$: terms of even degree in $f(t)$ contribute to the integer term $a \in \mathbb{Z}$, while terms of odd degree in $f(t)$ contribute to the other term $b\sqrt{2}$ (where $b \in \mathbb{Z}$). Similar remarks show that for every integer d, the previously studied ring $\mathbb{Z}[\sqrt{d}]$ is nothing other than the image of $\mathbb{Z}[t]$ under the map of evaluation at $\sqrt{d}$.

A related example is the following: let X be any set, and let R be the ring of all real-valued functions defined on X, i.e. functions $X \to \mathbb{R}$ under pointwise addition and multiplication: $(f+g)(x) = f(x) + g(x)$ and $(fg)(x) = f(x)g(x)$ for all $f, g \in R,\ x \in X$. Evaluation at a particular point $a \in X$ gives a map $R \to \mathbb{R}$, $f \mapsto f(a)$. This is a homomorphism.

An isomorphism is the same thing as a bijective homomorphism. This provides an abundant supply of further examples; some examples of isomorphisms are

$$\begin{aligned} \mathbb{C} \to \mathbb{C}, &\quad z \mapsto \overline{z}; \\ \mathbb{R}[t] \to \mathbb{R}[t], &\quad f(t) \mapsto f(2t); \\ M_2(\mathbb{R}) \to M_2(\mathbb{R}), &\quad \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right). \end{aligned}$$

(Here we use the standard arrow notation for functions, in which '$\to$' is used to express the domain and range of the function; and '$\mapsto$' is used to express the image of individual elements. This distinction between two types of arrows is required to avoid ambiguity, in cases where the elements of the domain are themselves sets.)

Every ring homomorphism $\theta : R \to S$ must map $0 \mapsto 0$ (i.e. $0_R \mapsto 0_S$ where 0_R and 0_S are the zero elements of R and S respectively). To see this, note that $\theta(0_R) = \theta(0_R + 0_R) = \theta(0_R) + \theta(0_R)$; now subtracting $\theta(0_R)$ from both sides gives $\theta(0_R) = 0_S$ as desired. Also note that for all $n \in \mathbb{N}$ and $r \in R$,

$$\theta(nr) = \theta\big(\underbrace{r + r + \cdots + r}_{n \text{ times}}\big) = \underbrace{\theta(r) + \theta(r) + \cdots + \theta(r)}_{n \text{ times}} = n\theta(r).$$

Now by taking additive inverses, we obtain $\theta(nr) = n\theta(r)$ for all $n \in \mathbb{Z}$ and $r \in R$.

Given any two rings R and S, the map $R \to S$ mapping $r \mapsto 0_S$ for all $r \in R$, is clearly a homomorphism; it is called the *trivial homomorphism*. When R and S are rings with identity, we typically want to exclude such degenerate homomorphisms; to do so, we may speak of a *homomorphism of rings with an identity* as a homomorphism $\theta : R \to S$ such that $\theta(1_R) = 1_S$. (Here 1_R and 1_S refer to the identity elements of R and S respectively.)

As further examples, for every $k \in \mathbb{Z}$ there is a homomorphism $\mu_k : \mathbb{Z} \to \mathbb{Z}$ given by $x \mapsto kx$. These are *all* the homomorphisms from $\mathbb{Z}$ to $\mathbb{Z}$. To see this, let $\theta : \mathbb{Z} \to \mathbb{Z}$ be a homomorphism, and take $m = \theta(1) \in \mathbb{Z}$. Then for all $n \in \mathbb{Z}$, we have

$$\theta(n) = \theta(n1) = n\theta(1) = nm = mn = \mu_m(n),$$

so $\theta = \mu_m$.

Here is a short glossary of related terminology:

epimorphism = surjective ("onto") homomorphism
monomorphism = injective ("one-to-one") homomorphism
isomorphism = bijective ("one-to-one and onto") homomorphism

Proposition 1. Let $\theta : R \to S$ be a homomorphism. Then θ is one-to-one (i.e. a monomorphism) iff its kernel is $\{0\}$.

Proof. We have seen that $\theta(0) = 0$ (i.e. $\theta(0_R) = 0_S$) so $\ker\theta$ contains 0. If θ is one-to-one then $\theta(r) = 0 = \theta(0)$ implies $r = 0$, so $\ker\theta = \{0\}$.

Conversely, suppose that $\ker\theta = \{0\}$. If $\theta(x) = \theta(y)$ for some , $y \in R$, then $\theta(x-y) = \theta(x) - \theta(y) = 0$ so $x - y \in \ker\theta = \{0\}$. This forces $x - y = 0$, so $x = y$. This means that θ is one-to-one. □

An *automorphism of a ring* R is an isomorphism $R \to R$. Every ring has at least one automorphism, the identity automorphism mapping every element to itself. Some rings have many automorphisms; but some (such as $\mathbb{R}$) have only the identity automorphism. The field $\mathbb{C}$ has at least two automorphisms: the identity, and complex conjugation. In fact, the field of complex numbers has *many* automorphisms, although we cannot explicitly describe them all here.

2. Motivation

Our goal is to provide a very general way to construct new rings from old. Given a ring R with an ideal $J \subseteq R$, the new ring we construct is called the *quotient ring* R/J. This process generalizes the construction of $\mathbb{Z}_m$ from $\mathbb{Z}$. Rather than precisely defining ideals and quotient rings now, let us first motivate the general process using a series of examples.

Example: $\mathbb{Z}/\langle 3\rangle$

In this case we start with the ring of integers $\mathbb{Z}$. Define the *principal ideal generated by 3* as the set of all multiples of 3:

$$\langle 3\rangle = 3\mathbb{Z} = \{3k : k \in \mathbb{Z}\} = \{\ldots, -6, -3, 0, 3, 6, 9, \ldots\}.$$

A *coset* of $\langle 3\rangle$ is a set of the form

$$r+\langle 3\rangle = \{r+x : x \in \langle 3\rangle\}$$

where $r \in \mathbb{Z}$, otherwise known as the congruence class of the number r. There are only three cosets in this case:

$$\begin{aligned} 0+\langle 3\rangle &= \{\ldots, -9, -6, -3, 0, 3, 6, 9, \ldots\} = \langle 3\rangle, \\ 1+\langle 3\rangle &= \{\ldots, -8, -5, -2, 1, 4, 7, 10, \ldots\}, \\ 2+\langle 3\rangle &= \{\ldots, -7, -4, -1, 2, 5, 8, 11, \ldots\}. \end{aligned}$$

Each coset can be represented by any of its elements. For example the coset

$$5+\langle 3\rangle = \{\ldots, -4, -1, 2, 5, 8, \ldots\} = 2+\langle 3\rangle.$$

This is why we have only three distinct cosets: once we have listed the three cosets $\langle 3\rangle$, $1+\langle 3\rangle$ and $2+\langle 3\rangle$, every integer is already covered and every coset $r+\langle 3\rangle$ coincides with one of the three cosets listed above.

We add two cosets by adding elementwise, which gives a new coset. For example

$$\begin{aligned} (1+\langle 3\rangle) + (1+\langle 3\rangle) &= \{\ldots, -5, -2, 1, 4, 7, \ldots\} + \{\ldots, -5, -2, 1, 4, 7, \ldots\} \\ &= \{\ldots, -10, -7, -4, -1, 2, 5, 8, 11, 14, \ldots\} = 2+\langle 3\rangle; \\ (1+\langle 3\rangle) + (2+\langle 3\rangle) &= \{\ldots, -5, -2, 1, 4, 7, \ldots\} + \{\ldots, -4, -1, 2, 5, 8, \ldots\} \\ &= \{\ldots, -9, -6, -3, 0, 3, 6, 9, 12, 15, \ldots\} = \langle 3\rangle. \end{aligned}$$

Similarly, we multiply two cosets by multiplying their elements, which again gives a new coset; thus for example

$$\begin{aligned}(2+\langle 3\rangle)(2+\langle 3\rangle) &= \{\ldots,-4,-1,2,5,8,\ldots\}\{\ldots,-4,-1,2,5,8,\ldots\}\\ &= \{\ldots,-8,-5,-2,1,4,7,10,13,\ldots\} = 1+\langle 3\rangle.\end{aligned}$$

The simple rule is to add two cosets by adding their representatives, and to multiply two cosets by multiplying their representatives, thus:

$$\begin{aligned}(r+\langle 3\rangle)+(s+\langle 3\rangle) &= (r+s)+\langle 3\rangle;\\ (r+\langle 3\rangle)(s+\langle 3\rangle) &= rs+\langle 3\rangle.\end{aligned}$$

The resulting collection of cosets forms a ring

$$\mathbb{Z}/\langle 3\rangle = \{\langle 3\rangle,\ 1+\langle 3\rangle,\ 2+\langle 3\rangle\}$$

with addition and multiplication defined as above. In general, the principal ideal generated by an integer m is $\langle m\rangle = m\mathbb{Z}$, the set of multiples of m. The quotient ring

$$\mathbb{Z}/\langle m\rangle = \{r+\langle m\rangle : r\in\mathbb{Z}\} = \{\langle m\rangle,\ 1+\langle m\rangle,\ 2+\langle m\rangle,\ m-1+\langle m\rangle\}$$

is a ring isomorphic to $\mathbb{Z}_m$; it is a field iff m is irreducible (i.e. of the form $\pm p$ where p is prime.

Example: $\mathbb{Z}[t]/\langle t^2-2\rangle$

For our next example, consider $R=\mathbb{Z}[t]$ and the irreducible polynomial $t^2-2\in\mathbb{Z}[t]$. The principal ideal generated by t^2-2 is the set of all multiples of t^2-2:

$$\langle t^2-2\rangle = (t^2-2)\mathbb{Z}[t] = \{(t^2-2)q(t) : q(t)\in\mathbb{Z}[t]\}.$$

The *coset of* $\langle t^2-2\rangle$ *containing* $f(t)$ is

$$f(t)+\langle t^2-2\rangle = \{f(t)+m(t) : m(t)\in\langle t^2-2\rangle\} = \{f(t)+q(t)(t^2-2) : q(t)\in\mathbb{Z}[t]\}.$$

For example, the coset containing $2t^3+5t^2-1$ is

$$\begin{aligned}2t^3+5t^2-1+\langle t^2-2\rangle &= (2t+5)(t^2-2)+4t+9+\langle t^2-2\rangle\\ &= 4t+9+\langle t^2-2\rangle.\end{aligned}$$

Note that the polynomial $(2t+5)(t^2-2) \in \langle t^2-2\rangle$ is absorbed into $\langle t^2-2\rangle$ since the ideal is closed under addition; so all we are left with is the remainder term $4t+9$. In general, to simplify the expression $f(t)+\langle t^2-2\rangle$, we first use the Division Algorithm to obtain

$$f(t) = q(t)(t^2-2) + at+b, \qquad q(t) \in \mathbb{Z}[t],\ a, b \in \mathbb{Z}.$$

Although the Division Algorithm does not always work in $\mathbb{Z}[t]$ (since $\mathbb{Z}$ is not a field), in this case it does work because we are dividing by a monic polynomial t^2-2, and division by 1 works in $\mathbb{Z}$. For a general coset we obtain

$$f(t)+\langle t^2-2\rangle = q(t)(t^2-2) + at+b + \langle t^2-2\rangle = at+b + \langle t^2-2\rangle.$$

The quotient ring $\mathbb{Z}[t]/\langle t^2-2\rangle$ consists of all such cosets, with addition and multiplication defined using representatives in $\mathbb{Z}[t]$. For example, consider the elements $\alpha, \beta \in \mathbb{Z}[t]/\langle t^2-2\rangle$ given by

$$\begin{aligned}\alpha &= 2t^3+5t^2-1 + \langle t^2-2\rangle = 4t+9 + \langle t^2-2\rangle;\\ \beta &= 3t-11 + \langle t^2-2\rangle.\end{aligned}$$

Adding and multiplying, we have

$$\begin{aligned}\alpha+\beta &= (4t+9 + \langle t^2-2\rangle) + (3t-11 + \langle t^2-2\rangle)\\ &= 7t-2 + \langle t^2-2\rangle;\\ \alpha\beta &= (4t+9 + \langle t^2-2\rangle)(3t-11 + \langle t^2-2\rangle)\\ &= 12t^2-17t-99 + \langle t^2-2\rangle\\ &= 12(t^2-2)\ -17t-75 + \langle t^2-2\rangle\\ &= -17t-75 + \langle t^2-2\rangle.\end{aligned}$$

With these operations, $\mathbb{Z}[t]/\langle t^2-2\rangle$ becomes a ring. In fact this is a ring you have seen before; it is isomorphic to $\mathbb{Z}[\sqrt{2}]$. To illustrate this, observe that

$$\begin{aligned}(9+4\sqrt{2}) + (-11+3\sqrt{2}) &= -2+7\sqrt{2}\,;\\ (9+4\sqrt{2})(-11+3\sqrt{2}) &= -75-17\sqrt{2}\,.\end{aligned}$$

The isomorphism $\mathbb{Z}[t]/\langle t^2-2\rangle \to \mathbb{Z}[\sqrt{2}]$ is given explicitly by $at+b+\langle t^2-2\rangle \mapsto a\sqrt{2}+b$.

Example: $\mathbb{R}[t]/\langle t^2+1\rangle$

Next consider the ring $R = \mathbb{R}[t]$, with irreducible polynomial $t^2+1 \in \mathbb{R}[t]$. The principal ideal generated by t^2+1 is

$$\langle t^2+1\rangle = (t^2+1)\mathbb{R}[t] = \{(t^2+1)q(t) : q(t) \in \mathbb{R}[t]\}.$$

Every coset has the form

$$f(t)+\langle t^2+1\rangle = at+b + \langle t^2+1\rangle.$$

Here we have used the Division Algorithm to subtract off a multiple of t^2+1 from $f(t)$, leaving a remainder $at+b$ of degree ≤ 1. The set of all cosets gives the quotient ring

$$\mathbb{R}[t]/\langle t^2+1\rangle = \{f(t)+\langle t^2+1\rangle : f(t) \in \mathbb{R}[t]\} = \{at+b+\langle t^2+1\rangle : a, b \in \mathbb{R}\}.$$

Consider two elements in this quotient ring given by

$$\alpha = 3t+8 + \langle t^2+1\rangle; \qquad \beta = 2t+5 + \langle t^2+1\rangle.$$

Adding and multiplying, we have

$$\begin{aligned}
\alpha + \beta &= (3t+8 + \langle t^2+1\rangle) + (2t+5 + \langle t^2+1\rangle) \\
&= 5t+13 + \langle t^2+1\rangle; \\
\alpha\beta &= (3t+8 + \langle t^2+1\rangle)(2t+5 + \langle t^2+1\rangle) \\
&= 6t^2+31t+40 + \langle t^2+1\rangle \\
&= 6(t^2+1) +31t+34 + \langle t^2+1\rangle \\
&= 31t+34 + \langle t^2+1\rangle.
\end{aligned}$$

The resulting ring $\mathbb{R}[t]/\langle t^2+1\rangle$ is in fact isomorphic to $\mathbb{C}$. To illustrate this, observe that

$$\begin{aligned}
(8+3i) + (5+2i) &= 13+5i\,; \\
(8+3i)(5+2i) &= 34+31i\,.
\end{aligned}$$

The isomorphism $\mathbb{R}[t]/\langle t^2+1\rangle \to \mathbb{C}$ is given explicitly by $at+b + \langle t^2+1\rangle \mapsto ai + b$.

Example: $\mathbb{Q}[t]/\langle t^3+t+1\rangle$

Next consider the ring $R = \mathbb{Q}[t]$, with irreducible polynomial $m(t) = t^3 + t + 1 \in \mathbb{Q}[t]$. (Although $m(t)$ is reducible in $\mathbb{R}[t]$, it is irreducible in $\mathbb{Q}[t]$.) As usual, the principal ideal $\langle m(t)\rangle$ consists of all multiples of $m(t)$. Every coset $f(t)+\langle m(t)\rangle$ is represented by the remainder of $f(t)$ after dividing by $m(t)$, so every coset has the form $at^2+bt+c + \langle m(t)\rangle$. We obtain the quotient ring

$$\mathbb{Q}[t]/\langle m(t)\rangle = \{at^2+bt+c + \langle m(t)\rangle : a, b, c \in \mathbb{Q}\}.$$

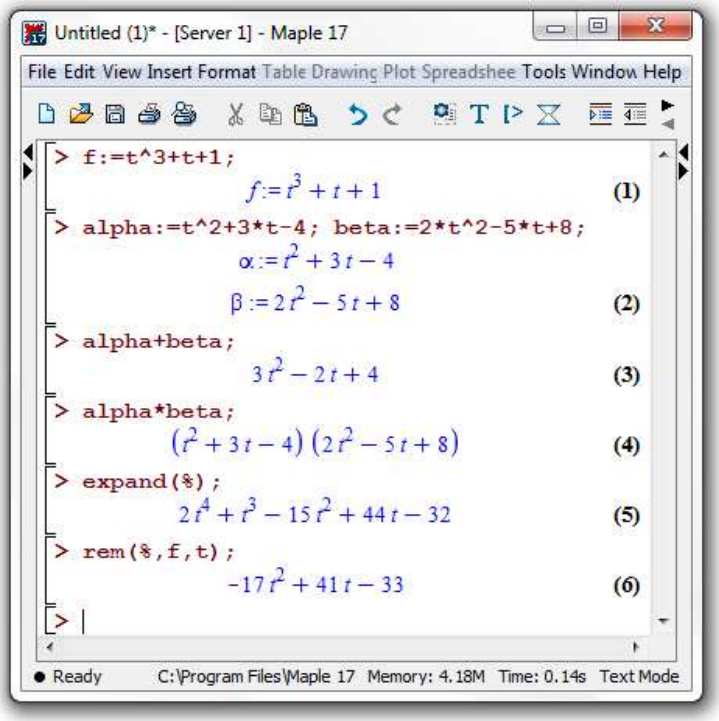
Consider two typical elements of this ring:

$$\alpha = t^2+3t-4 + \langle m(t)\rangle; \qquad \beta = 2t^2-5t+8 + \langle m(t)\rangle.$$

Their sum and product are

$$\begin{aligned}\alpha+\beta &= (t^2{+}3t{-}4+\langle m(t)\rangle)+(2t^2{-}5t{+}8+\langle m(t)\rangle)\\ &= 5t{+}13+\langle m(t)\rangle;\\ \alpha\beta &= (t^2{+}3t{-}4+\langle m(t)\rangle)(2t^2{-}5t{+}8+\langle m(t)\rangle)\\ &= 2t^4{+}t^3{-}15t^2{+}44t{-}32+\langle m(t)\rangle\\ &= -17t^2{+}41t{-}33+\langle m(t)\rangle.\end{aligned}$$

The latter remainder (after dividing by $m(t)$) was found using MAPLE®:



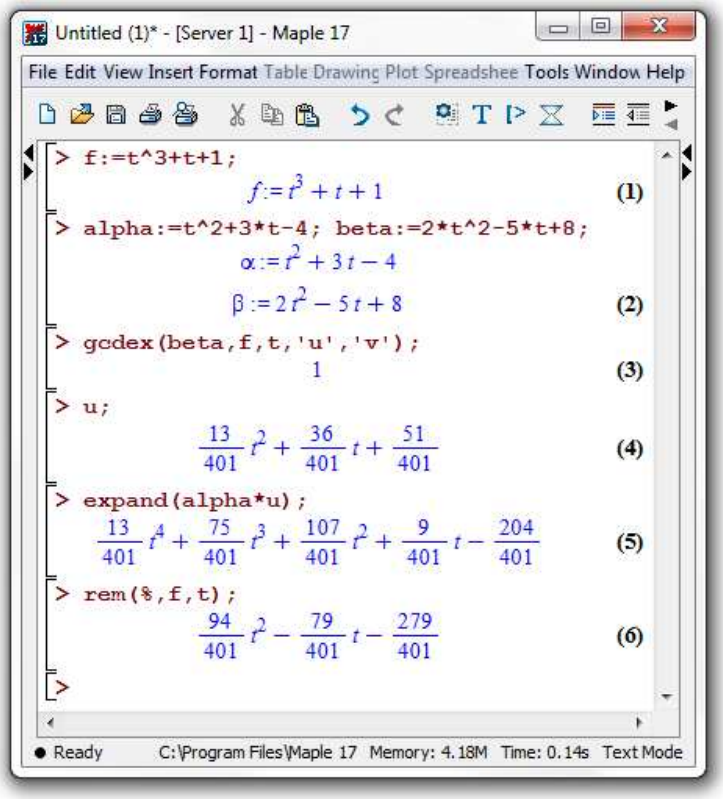
The ring $\mathbb{Q}[t]/\langle m(t)\rangle$ is actually a field. (Rings of the form $\mathbb{Z}/\langle m\rangle$ are fields iff m is irreducible, i.e. iff $m=\pm p$ where p is prime. In the same way, for any field F, the quotient ring $F[t]/\langle m(t)\rangle$ is a field iff $m(t)$ is irreducible.) Let us illustrate by dividing α/β. First we need the inverse of β. Since $m(t)$ is irreducible, it doesn't divide $2t^2{-}5t{+}8$, so

$$\gcd(2t^2{-}5t{+}8,\ m(t)) = 1 = u(t)(2t^2{-}5t{+}8)+v(t)m(t)$$

for some $u(t), v(t)\in\mathbb{Q}[t]$. Ignoring multiples of $m(t)$, which are absorbed into $\langle m(t)\rangle$, we see that

$$(u(t)+\langle m(t)\rangle)(2t^2{-}5t{+}8+\langle m(t)\rangle) = 1{+}\langle m(t)\rangle$$

so $\beta^{-1}=u(t){+}\langle m(t)\rangle$. Finally, we compute $\alpha/\beta=\alpha\beta^{-1}$ by multiplying as before. For convenience, let us use MAPLE®:



Thus

$$\beta^{-1} = \tfrac{13}{401}t^2 + \tfrac{36}{401}t + \tfrac{51}{401} + \langle m(t)\rangle;$$
$$\frac{\alpha}{\beta} = \alpha\beta^{-1} = \tfrac{94}{401}t^2 - \tfrac{79}{401}t - \tfrac{279}{401} + \langle m(t)\rangle.$$

The general idea for constructing a quotient ring R/J is to throw away all elements of J, as they are absorbed into J. This much works very much like working in $\mathbb{Z}_m$, where we throw away all multiples of m. The nice thing about the rings we have considered so far, is that the Division Algorithm allows us to reduce every coset $a+\langle m\rangle$ to $r+\langle m\rangle$ where r is the remainder of a mod m. In other rings we are not always so lucky to have a nicest representative for every coset. However, we can always form a ring from cosets of $J \subseteq R$ as long as J is an ideal.

3. Ideals

To make things as simple as possible, suppose for the time being that R is a commutative ring with identity. An *ideal of* R is a subring $J \subseteq R$ such that $rx \in J$ for all $r \in R$ and $x \in J$ (we write this condition as $RJ \subseteq J$). So in order to be an ideal, J must contain 0; J must be closed under addition and subtraction; and J must be closed under multiplication by all elements of R (not just closed under multiplying by other elements of J, as required for a subring). [The situation is a little more complicated for general rings, in which case

a *left ideal of* R is a subring $J \subseteq R$ satisfying $rx \in J$ for all $r \in R$ and $x \in J$, i.e. $RJ \subseteq J$; and a *right ideal of* R is a subring $J \subseteq R$ satisfying $xr \in J$ for all $x \in J$ and $r \in R$, i.e. $JR \subseteq J$. An *ideal of* R is then a two-sided ideal, i.e. a subring $J \subseteq R$ satisfying $RJ \subseteq J$ and $JR \subseteq J$.] The condition $RJ \subseteq J$ actually becomes $RJ = J$ in our case since R has an identity.

For example, let $m \in \mathbb{Z}$. We obtain an ideal in $\mathbb{Z}$ consisting of all integer multiples of m, namely

$$m\mathbb{Z} = \{mk : k \in \mathbb{Z}\} \subseteq \mathbb{Z}.$$

We will show (in Theorem 2 below) that every ideal of $\mathbb{Z}$ has this form.

Let $a \in R$ where R is a commutative ring with identity. Then $Ra = \{ra : r \in R\}$ is an ideal of R containing a. To see that $a \in Ra$, use the fact that $a = 1a \in Ra$. To verify that Ra is an ideal of R, we check that Ra is closed under addition and subtraction; and that Ra is closed under multiplication—by elements of R, not just closed under multiplication by elements of Ra itself. For all $ra, sa \in Ra$ and $t \in R$ (where $r, s, t \in R$), we have

$$ra + sa = (r + s)a \in Ra;$$
$$ra - sa = (r - s)a \in Ra;$$
$$t(ra) = (tr)a \in Ra.$$

An ideal of the form $Ra \subset R$ is called *principal*, and a is called a *generator* for the principal ideal Ra. The principal ideal $Ra \subseteq R$ is often denoted (a) or $\langle a \rangle$. We will typically write $Ra = \langle a \rangle$ since the textbook uses this notation. We now show that *every* ideal of $\mathbb{Z}$ is principal:

Theorem 2. Every ideal of $\mathbb{Z}$ is of the form $\langle m \rangle = m\mathbb{Z}$ for some $m \in \mathbb{Z}$.

Proof. Let $J \subseteq \mathbb{Z}$ be an ideal. Since J is a subring, it contains 0. If $J = \langle 0 \rangle$ then we are done. Otherwise J contains a nonzero integer a. Since J contains both a and $-a$, J contains at least one positive integer. Let S be the set of all positive integers contained in J; so from what we have seen, S is a nonempty subset of $\mathbb{N}$. Since $\mathbb{N}$ is well-ordered, we may take m to be the least element of S. Since J is an ideal containing m, J contains every multiple of m, i.e. $J \supseteq \langle m \rangle$. Conversely, let $x \in J$; we must show that $x \in m\mathbb{Z}$. By the Division Algorithm, $x = qm + r$ for some integers q, r satisfying $0 \leq r < m$. However, $r = x - qm \in J$; and since m is the smallest *positive* element of J, we must in fact have $r = 0$; thus $x = qm \in m\mathbb{Z}$ as required, giving $J = \langle m \rangle$. □

The same argument works for other Euclidean domains, including $F[t]$ where F is any field:

Theorem 2′. Let F be a field. Then every ideal of the polynomial ring $F[t]$ is principal. Every ideal $J \subseteq F[t]$ satisfies either $J = \langle 0 \rangle$, or J has a unique monic generator $g(t) \in F[t]$ so that $J = \langle g(t) \rangle = \{f(t)g(t) : f(t) \in F[t]\}$. □

A ring in which *every* ideal is principal is called a *principal ideal ring*. Thus $\mathbb{Z}$ and $F[t]$ are principal ideal rings, where F is any field. An example of a ring which is *not* a principal ideal ring is the ring $\mathbb{R}[x, y]$ consisting of all polynomials in two variables x and y, with real coefficients. Let $J \subset \mathbb{R}[x, y]$ be the set of all polynomials $f(x, y) \in \mathbb{R}[x, y]$ with no constant term, i.e. satisfying $f(0, 0) = 0$. It is easy to see that J is an ideal of $\mathbb{R}[x, y]$; and we will show by contradiction that J is not principal. Suppose that $g(x, y)$ is a generator for J, so that every element of J (such as x and y) is divisible by $g(x, y)$. Then $x = g(x, y)q_1(x, y)$ and $y = g(x, y)q_2(x, y)$ for some $q_1(x, y), q_2(x, y) \in \mathbb{R}[x, y]$. However any polynomial dividing both x and y must be a nonzero constant, i.e. $\gcd(x, y) = 1$. Thus $g(x, y)$ is a nonzero constant. However this cannot be in J since it has nonzero constant term, a contradiction.

Theorem 3. Let $\theta : R \to S$ be a homomorphism of rings. Then
(i) the kernel $K = \ker \theta$ is an ideal of R; and
(ii) the image $\theta(R)$ is a subring of S.

Proof. We have seen that $\theta(0) = 0$, so $0 \in K$. If $x, y \in K$ then $\theta(x \pm y) = \theta(x) \pm \theta(y) = 0 \pm 0 = 0$, so K is closed under addition and subtraction. Also if $x \in K$ and $r \in R$ then $\theta(rx) = \theta(r)\theta(x) = \theta(r)0 = 0$ so $rx \in K$. This proves (i).

The image $\theta(R)$ contains $\theta(0) = 0$. It is closed under addition, subtraction and multiplication since if $r, s \in R$ then $\theta(r) \pm \theta(s) = \theta(r \pm s) \in \theta(R)$ and $\theta(r)\theta(s) = \theta(rs) \in \theta(R)$. This proves (ii). □

From Theorem 3, we note that the kernel of a homomorphism is always an ideal. In Section 3, we prove a converse of this: every ideal is the kernel of a homomorphism. Thus ideals are the same things as kernels of homomorphisms.

For example, the ideals of $\mathbb{Z}$ have the form $\langle m \rangle \subseteq \mathbb{Z}$ as we have claimed. When $m \neq 0$, note that $\langle m \rangle$ is in fact the kernel of the ring homomorphism $\mathbb{Z} \to \mathbb{Z}_m$ which reduces every

integer modulo m. And in the exceptional case when $m = 0$, the ideal $\langle 0 \rangle$ is the kernel of the homomorphism $\mathbb{Z} \to \mathbb{Z}$ which maps $x \mapsto x$. [*Aside:* Although the notation $\mathbb{Z}_m$ has become popular, especially by authors of undergraduate textbooks, this notation is unfortunate as it conflicts with an established notation for rings of p-adic numbers.]

As another example, for each $a \in \mathbb{C}$ let $K_a = \{f(x) \in \mathbb{Q}[x] : f(a) = 0\}$. Then by definition, K_a is the kernel of the evaluation map $\theta_a : \mathbb{Q}[t] \to \mathbb{C}$ given by $f(t) \mapsto f(a)$. Thus K_a is an ideal (which is also easy to see directly). Now by Theorem $2'$, K_a is a principal ideal: it consists of all multiples of some generator polynomial $m_a(t) \in \mathbb{Q}[t]$. If a is *transcendental*, then $K_a = \{0\}$ and $m_a(t) = 0$. Otherwise a is *algebraic*, K_a is a nonzero ideal and if we take $m_a(t) \in \mathbb{Q}[t]$ to be monic, then it is unique. We call $m_a(t)$ the *minimal polynomial of a.* For example

- π and e are transcendental; $K_\pi = K_e = \{0\}$. There are no nonzero polynomials $f(t) \in \mathbb{Q}[t]$ having either π or e as a root.
- $\sqrt{2}$ is algebraic with minimal polynomial $m_{\sqrt{2}}(t) = t^2 - 2$.
- $i = \sqrt{-1}$ is algebraic with minimal polynomial $m_i(t) = t^2 + 1$.
- $2^{1/3}$ is algebraic with minimal polynomial $m_{2^{1/3}}(t) = t^3 - 2$.
- $\sqrt{1+\sqrt{2}}$ is algebraic with minimal polynomial $m_{\sqrt{1+\sqrt{2}}}(t) = t^4 - 2t^2 - 1$.

4. Quotient Rings

Let J be an ideal of a ring R. For each $r \in R$, define the *coset of J containing r* by

$$r + J = \{r + x : x \in J\}.$$

Proposition 4. Let J be an ideal of a ring R; and let $r, r' \in R$.

(i) If $r - r' \notin J$, then the two cosets $r + J$ and $r' + J$ are disjoint, i.e. $(r + J) \cap (r' + J) = \varnothing$.

(ii) If $r - r' \in J$, then the two corresponding cosets coincide, i.e. $r + J = r' + J$.

Proof. If $r + J = r' + J$ then $f = f + 0 \in r + J = r' + J$, so $r = r' + x$ for some $x \in J$. This proves (ii). Now suppose $(r + J) \cap (r' + J)$ contains an element $r + x = r' + x'$ where $x, x' \in J$; then $r - r' = x' - x \in J$. This is the contrapositive of (i). □

The cosets of J form a new ring, called the *quotient ring R/J*. Addition and multiplication of cosets are defined by

$$(r + J) + (s + J) = (r + s) + J; \qquad (r + J)(s + J) = rs + J.$$

The zero element of R/J is $0 + J = J$. The problem is to show that these operations are *well-defined.* As an example, the quotient of $\mathbb{Z}$ by the ideal $\langle m \rangle \subseteq \mathbb{Z}$ gives $\mathbb{Z}/\langle m \rangle \cong \mathbb{Z}_m$.

> What is this concern about? An example of a binary operation that is not well-defined is exponentiation on $\mathbb{Z}_3$: for $x, y \in \mathbb{Z}_3$ we would like to define $x^y \in \mathbb{Z}_3$ in the most natural way possible. Since $2 = 5$ and $1 = 4$, we should have $2^1 = 5^4$. In $\mathbb{Z}_3$, this says that $2 = 625$, which is *not* correct in $\mathbb{Z}_3$. By contrast, the binary operations of addition and multiplication on $\mathbb{Z}_3$ are well-defined. Note for example that $2+1 = 5+4$ (both sides give $0 \in \mathbb{Z}_3$) and $2 \cdot 1 = 5 \cdot 4$ (both sides give $2 \in \mathbb{Z}_3$). The fact that addition and multiplication are well-defined binary operations on $\mathbb{Z}_m$, follows from:

Theorem 5. The structure R/J, as defined above, is a ring.

Proof. As pointed out above, our first and main job is to check that the operations of addition and multiplication of cosets, are well-defined. The concern is that if $r + J = r' + J$ and $s + J = s' + J$, then the sum and product should work out the same way no matter which choice of representatives we choose. By Proposition 3, the hypotheses imply that $r - r' \in J$ and $s - s' \in J$. Now

$$(r + s) + J = (r' + s') + (r - r') + (s - s') + J \subseteq (r' + s') + J$$

since J is closed under addition. The reverse inclusion

$$(r' + s') + J \subseteq (r + s) + J$$

follows by the same argument, so actually we must have equality:

$$(r + s) + J = (r' + s') + J.$$

This means that the sum of $r + J$ and $s + J$ is well-defined: we get the same answer whether we add them using the old coset representatives r, s or using the new coset representatives r', s'. Next we check that multiplication is also well-defined:

$$rs + J = r's' + (r - r')s + r'(s - s') + J \subseteq r's' + J$$

since $(r - r')s \in J$; $r'(s - s') \in J$; and J is closed under addition. The reverse inclusion

$$(r' + s') + J \subseteq (r + s) + J$$

follows by the same argument, so again we must have equality:

$$rs + J = r's' + J.$$

This means that the product of $r + J$ and $s + J$ is well-defined: we get the same answer whether we add them using the old coset representatives r, s or using the new coset representatives r', s'.

It is easy to see that R satisfies all the other requirements for a ring. For example, the left-distributive law for R/J follows from the left-distributive law for R, as follows:

$$\begin{aligned}(r + J)\big((s + J) + (t + J)\big) &= (r + J)\big((s + t) + J)\big) \\ &= r(s + t) + J \\ &= (rs + rt) + J \\ &= (rs + J) + (rt + J) \\ &= (r + J)(s + J) + (r + J)(t + J)\end{aligned}$$

for all $r, s, t \in R$. Verifying the other ring axioms for R/J in this way is left as a straightforward exercise. □

Theorem 6. Let J be an ideal of a ring R. The map $\pi : R \to R/J$, $r \mapsto r + J$ is a homomorphism with kernel J. In particular, every ideal is the kernel of some homomorphism.

The homomorphism $\pi : R \to R/J$ of Theorem 6 is called the *canonical homomorphism*. Since π is clearly *onto* R/J (i.e. its image is $\pi(R) = R/J$ by definition), π is in fact an epimorphism.

Proof. We have $r \in \ker \pi$ iff the coset $\pi(r) = r + J \in R/J$ coincides with the zero element $0 + J = J \in R/J$. By Proposition 1, this is equivalent to the condition $r = r - 0 \in J$. This proves that $\ker \pi = J$. The fact that $\pi(r + s) = \pi(r) + \pi(s)$ and $\pi(rs) = \pi(r)\pi(s)$ follows from the way we defined sums and products of cosets. □

The following is called the *First Isomorphism Theorem for Rings*.

Theorem 7. Let $\theta : R \to S$ be any homomorphism of rings. Recall (from Theorem 2) that the kernel $K = \ker\theta$ is an ideal of R, and $\theta(R)$ is a subring of S. Then

$$R/K = R/\ker\theta \cong \theta(R) \subseteq S.$$

Proof. We have shown that K is an ideal of R. Define $\overline{\theta} : R/K \to \theta(R)$ by $\overline{\theta}(r+K) = \theta(r)$. We must first check that $\overline{\theta}$ is well-defined: if $r + K = r' + K$ then $r - r' \in K$ so $\theta(r) - \theta(r') = \theta(r - r') = 0$, i.e. $\theta(r) = \theta(r')$. Whether we compute $\overline{\theta}(r+K) = \overline{\theta}(r'+K)$ using the representative r or the representative r', we get the same answer; so $\overline{\theta}$ is well-defined. It is easy to verify that $\overline{\theta}$ is a homomorphism: if $r, s \in R$ then

$$\overline{\theta}\big((r+K)+(s+K)\big) = \overline{\theta}\big((r+s)+K\big) = \theta(r+s) = \theta(r)+\theta(s) = \overline{\theta}(r+K)+\overline{\theta}(s+K);$$
$$\overline{\theta}\big((r+K)(s+K)\big) = \overline{\theta}\big(rs+K\big) = \theta(rs) = \theta(r)\theta(s) = \overline{\theta}(r+K)\overline{\theta}(s+K).$$

The fact that $\overline{\theta}$ is onto $\theta(R)$ is really by definition: every element of $\theta(R)$ has the form $\theta(r) = \overline{\theta}(r+K)$ for some $r \in R$. To show that $\overline{\theta}$ is one-to-one, suppose $\overline{\theta}(r+K) = 0$ for some $r \in R$. This says that $\theta(r) = 0$, so $r \in \ker\theta = K$; but then $r + K = K$. This shows that the only coset in the kernel of $\overline{\theta}$ is the zero coset $0 + K = K$, showing (by Proposition 2) that $\overline{\theta}$ is one-to-one. Thus $\overline{\theta}$ is an isomorphism.

Examples

Consider the evaluation map $\theta_i : \mathbb{R}[t] \to \mathbb{C}$, $f(t) \mapsto f(i)$. As before, like all evaluation-at-a-point maps, this is a homomorphism. Clearly θ_i is onto $\mathbb{C}$. The kernel of θ_i is a nonzero ideal generated by $m_i(t)$. We clearly see that $m_i(t) = t^2 + 1$. (Let $K = \ker\theta_i = \{f(t) \in \mathbb{R}[t] : f(i) = 0\}$. If $f(t) \in K$ then $f(i) = 0$ and $f(-i) = \overline{f(i)} = \overline{0} = 0$ so $f(t)$ is divisible by both $t - i$ and $t + i$; thus $f(t)$ is divisible by $(t+i)(t-i) = t^2 + 1$. It follows that $K = \langle m_i(t)\rangle = \langle t^2+1\rangle$.) We have

$$\mathbb{R}[t]/\langle t^2+1\rangle \cong \mathbb{C}.$$