ALGEBRA II: RINGS AND MODULES OVER LITTLE RINGS.

KEVIN MCGERTY.

1. RINGS

The central characters of this course are algebraic objects known as rings. A ring is any mathematical structure where you can add and multiply, and could be thought of generalising $\mathbb{Z}$ the integers. Formally speaking we have:

Definition 1.1. A ring is a datum $(R, +, \times, 0, 1)$ where R is a set, $1, 0 \in R$ and $+, \times$ are binary operations on R such that

(1) R is an abelian group under $+$ with identity element 0.
(2) The binary operation $\times$ is associative and $1 \times x = x \times 1 = x$ for all $x \in R$.[1]
(3) Multplication distributes over addition:

$$x \times (y + z) = (x \times y) + (x \times z), \quad \forall x, y, z \in R.$$

Just as for multiplication of real numbers or integers, we will tend to suppress the symbol for the operation $\times$, and write "." or omit any notation at all. If the operation $\times$ is commutative (*i.e.* if $x.y = y.x$ for all $x, y \in R$) then we say R is a commutative ring[2]. Sometimes people consider rings which do not have a multiplicative indentity[3]. We won't. It is also worth noting that some texts require an additional axioms asserting that $1 \neq 0$. In fact it's easy to see from the other axioms that if $1 = 0$ then the ring has only one element. We will refer to this ring as the "zero ring", which is a somewhat degenerate object, but it seems unnecessary to me to exclude it.

Example 1.2.
i) The integer $\mathbb{Z}$ form the fundamental example of a ring. In some sense much of the course will be about finding an interesting class of rings which behave a lot like $\mathbb{Z}$. Similarly if $n \in \mathbb{Z}$ then $\mathbb{Z}/n\mathbb{Z}$, the integers modulo n, form a ring with the usual addition and multiplication.
ii) The subset $\mathbb{Z}[i] = \{a + ib \in \mathbb{C} : a, b \in \mathbb{Z}\}$ is easily checked to be a ring under the normal operations of addition and multiplication of complex numbers. It is known as the *Gaussian integers*. We shall see later that it shares many of the properties with the ring $\mathbb{Z}$ of ordinary integers.
iii) Any field, *e.g.* $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, is a ring – the only difference between the axioms for a field and for a ring is that in the case of a ring we do not require the existence of multiplicative inverses (and that, for fields one insists that $1 \neq 0$, so that the smallest field has two elements).
iv) If k is a field, and $n \in \mathbb{N}$, then the set $M_n(\mathsf{k})$ of $n \times n$ matrices with entries in k is a ring, with the usual addition and multiplication of matrices.
v) Saying the previous example in a slightly more abstract way, if V is a vector space over a field k then $\mathrm{End}(V)$ the space of linear maps from V to V, is a ring. In this case the multiplication is given by composition of linear maps, and hence is not commutative. We will mostly focus on commutative rings in this course.
vi) Example *iii*) also lets us construct new rings from old, in that there is no need to start with a field k. Given any ring R, the set $M_n(R)$ of $n \times n$ matrices with entries in R is again a ring.
vii) Polynomials in any number of indeterminates form a ring: if we have n variables $t_1, t_2, \ldots, t_n$ and k is a field then we write $\mathsf{k}[t_1, \ldots, t_n]$ for the ring of polynomials in the variables $t_1, \ldots, t_n$ with coefficients in k.

Date: October, 2011.

[1] That is, R is a monoid under $\times$ with identity element 1 if you like collecting terminology.

[2] We will try and use the letter R as our default symbol for a ring, in some books the default letter is A. This is the fault of the French, as you probably guess.

[3] In some texts they use (rather hideous) term *rng* for such an object.

viii) Just as in *v*), there is no reason the coefficients of our polynomials have to be a field – if R is a ring, we can build a new ring $R[t]$ of polynomials in t with coefficients in R in the obvious way. It is important to note in both this and the previous example is that polynomials are no longer function: given a polynomial $f \in R[t]$ we may evaluate it at and $r \in R$, thus we can associate it to a function from R to R, but this function may not determine f. For example if $R = \mathbb{Z}/2\mathbb{Z}$ then clearly there are only finitely many functions from R to itself, but $R[t]$ still contains infinitely many polynomials.

ix) If we have two rings R and S, then we can form the *direct sum* of the rings $R \oplus S$: this is the ring whose elements are pairs (r, s) where $r \in R$ and $s \in S$ with addition and multiplication given componentwise.

x) Another way to construct new rings from old is to consider, for a ring R, functions taking values in R. The simplest example of this is $R^n = (a_1, \dots, a_n)$, where we add and multiply coordinatewise. This is just[4] the ring of R-valued functions on the set $\{1, 2, \dots, n\}$. We can generalise this and consider, for any set X, the set $R^X = \{f : X \to R\}$ of functions from X to R, and make it a ring by adding values (exactly as we define the sum of two $\mathbb{R}$ or $\mathbb{C}$-valued functions).

xi) To make the previous example more concrete, the set of all functions $f : \mathbb{R} \to \mathbb{R}$ is a ring. Moreover, the set of all continuous (or differentiable, infinitely differentiable,...) functions also forms a ring by standard algebra of limits results.

Definition 1.3. If R is a ring, a subset $S \subseteq R$ is said to be a *subring* if it inherits the structure of a ring from R, thus we must have $0, 1 \in S$ and moreover S is closed under the addition and multiplication operations in R so that $(S, +)$ is a subgroup of $(R, +)$.

For example, the integers $\mathbb{Z}$ are a subring of $\mathbb{Q}$, the ring of differentiable functions from $\mathbb{R}$ to itself is a subring of the ring of all functions from $\mathbb{R}$ to itself. The ring of Gaussian integers is a subring of $\mathbb{C}$, as are $\mathbb{Q}$,$\mathbb{R}$ (the latter two being fields of course). Recall that for a group G containing a subset H, the *subgroup criterion* says that H is a subgroup if and only if it is nonempty and whenever $h_1, h_2 \in H$ we have $h_1 h_2^{-1} \in H$ (here I'm writing the group operation on G multiplicatively). We can use this to give a similar criterion for a subset of a ring to be a subring.

Lemma 1.4 (Subring criterion)**.** *Let R be a ring and S a subset of R, then S is a subring if and only if $1 \in S$ and for all $s_1, s_2 \in S$ we have $s_1 s_2, s_1 - s_2 \in S$.*

Proof. The condition that $s_1 - s_2 \in S$ for all $s_1, s_2 \in S$ implies that S is an additive subgroup by the subgroup test (note that as $1 \in S$ we know that S is nonempty). The other conditions for a subring hold directly. □

When studying any kind of algebraic object[5] it is natural to consider maps between those kind of objects which respect their structure. For example, for vector spaces the natural class of maps are linear maps, and for groups the natural class are the group homomorphisms. The natural class of maps to consider for rings are defined similarly:

Definition 1.5. A map $f : R \to S$ between rings R and S is said to be a (*ring*) *homomorphism* if

(1) $f(1_R) = 1_S$,
(2) $f(r_1 + r_2) = f(r_1) + f(r_2)$,
(3) $f(r_1.r_2) = f(r_1).f(r_2)$,

where strictly speaking we might have written $+_R$ and $+_S$ for the addition operation in the two different rings R and S, and similarly for the multiplication operation[6]. Apart from the fact that things then become hard to read, because the required syntax is clear from context, hopefully this (conventional) sloppiness in notation will not bother anyone. Note that it follows from (2) that $f(0) = 0$.

[4]Recall, for example, that sequences of real numbers are defined to be functions $a : \mathbb{N} \to \mathbb{R}$, we just tend to write a_n for the value of a at n (and refer to it as the n-th term) rather than $a(n)$.

[5]Or more generally any mathematical structure: if you're taking Topology this term then continuous maps are the natural maps to consider between topological spaces, similarly in Integration you consider measurable functions: loosely speaking, you want to consider maps which play nicely with the structures your objects have, be that a topology, a vector space structure, a ring structure or a measure.

[6]though since I've already decided to supress the notation for it, it's hard to distinguish the two when you supress both...

It is easy to see that the image of a ring homomorphism $f\colon R \to S$, that is $\{s \in S : \exists r \in R, f(r) = s\}$ is a subring of S. If it is all of S we say f is surjective, and $f\colon R \to S$ is an isomorphism if there is a homomorphism $g\colon S \to R$ such that $f \circ g = \mathrm{id}_S$ and $g \circ f = \mathrm{id}_R$. It is easy to check that f is an isomorphism if and only if it is a bijection (that is, to check that the set-theoretic inverse of f is automatically a ring homomorphism – you probably did a similar check for linear maps between vector spaces before.)

Example 1.6.
i) For each positive integer n, there is a natural map from $\mathbb{Z}$ to $\mathbb{Z}/n\mathbb{Z}$ which just takes an integer to its equivalence class modulo n. The standard calculations that check modular arithmetic makes sense exactly show that this map is a ring homomorphism.

ii) Let V be a k-vector space and let $\alpha \in \mathrm{End}_{\mathsf{k}}(V)$. Then $\phi\colon \mathsf{k}[t] \to \mathrm{End}_{\mathsf{k}}(V)$ given by $\phi(\sum_{i=0}^n a_i t^i) = \sum_{i=0}^n a_i \alpha^i$ is a ring homomorphism. Ring homomorphisms of this type will connnection the study of the ring $\mathsf{k}[t]$ to linear algebra. (In a sense you saw this last term when defining things like the minimal polynomial of a linear map, but we will explore this more fully in this course.)

iii) Obviously the inclusion map $i\colon S \to R$ of a subring S into a ring R is a ring homomorphism.

iv) Let $A = \{\begin{pmatrix} a & -b \\ b & a \end{pmatrix} : a, b \in \mathbb{R}\}$. It is easy to check this A is a subring of $\mathrm{Mat}_2(\mathbb{R})$. The map $\phi\colon \mathbb{C} \to A$ given by $a + ib \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ is a ring isomorphism. (This homomorphism arises by sending a complex number z to the map of the plane to itself given by multiplication by z.)

The first of the above examples has an important generalisation which shows that any ring R in fact has a smallest subring: For $n \in \mathbb{Z}_{\geq 0}$ set $n_R = 1 + 1 + \ldots + 1$ (that is, 1, added to itself n times), and for n a negative integer $n_R = -(-n)_R$. You can check that $\{n_R : n \in \mathbb{Z}\}$ is a subring of R, and indeed that the map $n \mapsto n_R$ gives a ring homomorphism from $\phi\colon \mathbb{Z} \to R$. Since a ring homomorphism is in particular a homomorphism of the underlying abelian groups under addition, using the first isomorphism theorem for abelian groups we see that $\{n_R : n \in \mathbb{Z}\}$, as an abelian group, is isomorphic to $\mathbb{Z}/d\mathbb{Z}$ for some $d \in \mathbb{Z}_{\geq 0}$. Since any subring S of R contains 1, and hence, since it is closed under addition, n_R for all $n \in \mathbb{Z}$, we see that S contains the image of ϕ, so that the image is indeed the smallest subring of R.

Definition 1.7. The integer d defined above is called the *characteristic* of the ring R.

Remark 1.8. The remark above that in general polynomials with coefficients in a ring cannot always be viewed as functions might have left you wondering what such a polynomial actually is. In other words, what do we mean when we say $\mathsf{k}[t]$ is a ring where "t is an indeterminate." The answer is a bit like the definition of complex numbers: to construct them you just define an addition and multiplication on $\mathbb{R}^2$ and then check the definitions do what you want them to, so that $(0, 1)$ becomes "i". For polynomials, we just start with the sequence of coefficients, and do the same thing. More precise, if R is a ring, consider the set $R[t]$ of sequences[7] $(a_n)_{n\in\mathbb{N}}$ where all but finitely many of the a_ns are zero, that is, that there is some $N \in \mathbb{N}$ such that $a_n = 0$ for all $n \geq N$. Then if $(a_n), (b_n)$ are two such sequences, define

$$(a_n) + (b_n) = (a_n + b_n);$$

$$(a_n).(b_n) = (\sum_{k=0}^{n} a_k b_{n-k}).$$

It is easy to see that if $a_n = 0$ for all $n \geq N$ and $b_n = 0$ for all $n \geq M$, then $a_n + b_n = 0$ if $n \geq \max\{N, M\}$ which $\sum_{k=0}^n a_k b_{n-k} = 0$ if $n \geq N + M$, since then at least one of a_k or b_{n-k} must be zero (otherwise $k < N$ and $n - k < M$ so that $n = k + (n - k) < N + M$). It is then routine to check that $R[t]$ forms a ring, which contains R viewed as the sequences (a_n) where $a_0 = r \in R$ and $a_n = 0$ for $n > 0$. The "indeterminate" t is then just the sequence $(0, 1, 0, \ldots)$. In fact it is easy to check that $t^n = (0, \ldots, 0, 1, 0, \ldots)$ where the 1 is in position n, and thus if (a_n) is a sequence as above with $a_n = 0$ for all $n \geq N$ then $(a_n) = \sum_{n=0}^N a_n t^n$.

In fact, the set of *all* sequences $(a_n)_{n\in\mathbb{N}}$ forms a ring with the same definitions for addition and multiplication. This is known as the ring of formal power series in t, and is denoted $R[[t]]$. (The name comes from the fact that, we view elements of $R[[t]]$ as "infinite sums" $\sum_{n\geq 0} a_n t^n$.) Perhaps surprisingly, it turns out that that, say, $\mathbb{C}[[t]]$ has a simpler structure in many ways that $\mathbb{C}[t]$.

[7] In this course, $\mathbb{N}$ will denote the non-negative integers unless it's obviously supposed to denote the positive integers.

2. BASIC PROPERTIES

From now on, unless we explicitly state otherwise, all rings will be assumed to be commutative.

Now that we have seen some examples of rings, we will discuss some basic properties of rings and their elements. Note that it is a routine exercise[8] in axiom grubbing to check that, for any ring R, we have $a.0 = 0$ for all $a \in R$. The next definition records the class of rings for which this is the only case in which the product of two elements is zero.

Definition 2.1. If R is a ring, then an element $a \in R\backslash\{0\}$ is said to be a *zero-divisor* if there is some $b \in R\backslash\{0\}$ such that $a.b = 0$. A ring which is not the zero ring and has no zero-divisors is called an *integral domain*. Thus if a ring is an integral domain and $a.b = 0$ then one of a or b is equal to zero.

Another way to express the fact that a ring is an integral domain is observe that it is exactly the condition which permits cancellation[9], that is, if $x.y = x.z$ then in an integral domain you can conclude that either $y = z$ or $x = 0$. This follows immediately from the definition of an integral domain and the fact that $x.y = x.z \iff x.(y - z) = 0$, which follows from the distributive axiom.

Example 2.2. If R is a ring, then R^2 is again a ring, and $(a, 0).(0, b) = (0, 0)$ so that $(a, 0)$ and $(0, b)$ are zero-divisors. The (noncommutative) ring of $n \times n$ matrices $M_n(\mathsf{k})$ for a field k also has lots of zero divisors, even though a field k does not. The integers modulo n have zero-divisors whenever n is not prime.

On the other hand, it is easy to see that a field has no zero-divisors. The integers $\mathbb{Z}$ are an integral domain (and *not* a field). Slightly more interestingly, if R is an integral domain, then $R[t]$ is again an integral domain. Moreover, the same is true of $R[[t]]$.

Exercise 2.3. Show that if R is an integral domain then $R[t]$ is also.

Recall the characteristic of a ring defined in the last lecture.

Lemma 2.4. *Suppose that R is an integral domain. Then any subring S of R is also an integral domain. Moreover, $\mathrm{char}(R)$, the characteristic of R, is either zero or a prime $p \in \mathbb{Z}$.*

Proof. It is clear from the definition that a subring of an integral domain must again be an integral domain. Now from the definition of the characteristic of a ring, if $\mathrm{char}(R) = n > 0$ then $\mathbb{Z}/n\mathbb{Z}$ is a subring of R. Clearly if $n = a.b$ where $a, b \in \mathbb{Z}$ are both greater than 1, then $a_R.b_R = 0$ in R with neither a_R nor b_R zero, thus both are zero divisors. It follows that if R is an integral domain then $\mathrm{char}(R)$ is zero or a prime. □

Recall that in a ring we do not require that nonzero elements have a multiplicative inverse[10]. Nevertheless, because the multiplication operation is associative and there is a multiplicative identity, the elements which happen to have multiplicative inverses form a group:

Definition 2.5. Let R be a ring. The subset

$$R^\times = \{r \in R : \exists s \in R, r.s = 1\},$$

is called the group of *units* in R – it is a group under the multiplication operation $\times$ with identity element 1.

Example 2.6. The units in $\mathbb{Z}$ form the group $\{\pm 1\}$. On the other hand, if k is a field, then the units $\mathsf{k}^\times = \mathsf{k}\backslash\{0\}$. If $R = M_n(\mathsf{k})$ then the group of units is $\mathrm{GL}_n(\mathsf{k})$.

Note that in particular, the characteristic of a field is always zero or a prime.

In our example of $\mathbb{Z}/n\mathbb{Z}$ notice that this ring either has zero-divisors (when n is composite) or is a field (when n is prime). In fact this is dichotomy holds more generally.

Lemma 2.7. *Let R be an integral domain which has finitely many elements. Then R is a field.*

---

[8] Until you feel you might die of boredom, it's a good idea to try and check that the axioms for a ring mean do indeed imply that you can indeed perform standard algebraic manipulations you are used to, so things like $0.x = 0$ hold in any ring.

[9] Except for the assertion the ring is not the zero ring, the zero ring having cancellation vacuously.

[10] As noted above, the axioms for a ring imply that $0.x = 0$ for all $x \in R$, thus the additive identity cannot have a multiplicative identity, hence the most we can ask for is that every element of $R\backslash\{0\}$ does – this is exactly what you demand in the axioms for a field.

Proof. We need to show that if $a \in R\backslash\{0\}$ then a has a multiplicative inverse, that is, we need to show there is a $b \in R$ with $a.b = 1$. But consider the map $m_a \colon R \to R$ given by left multiplication by a, so that $m_a(x) = a.x$. We claim that m_a is injective: indeed if $m_a(x) = m_a(y)$ then we have

$$a.x = a.y \implies a.(x - y) = 0,$$

and since R is an integral domain and $a \neq 0$ it follows that $x - y = 0$, that is, $x = y$. But now since R is finite, an injective map must be surjective, and hence there is some $b \in R$ with $m_a(b) = 1$, that is, $a.b = 1$ as required. □

Remark 2.8. Note that the argument in the proof which shows that multiplicative inverses exist does not use the assumption that the ring R was commutative (we only need it in order to conclude R is a field). A noncommutative ring where any nonzero element has an inverse is called a *division ring* (or sometimes a *skew field*). Perhaps the most famous example of a division ring is the ring of quaternions.

2.1. **The field of fractions.** If R is an integral domain which is infinite, it does not have to be a field (*e.g.* consider the integers $\mathbb{Z}$). However, generalising the construction of the rational numbers from the integers, we may build a field $F(R)$ from R by, loosely speaking, "taking ratios": the elements of $F(R)$ are "fractions" a/b where $a, b \in R$ and $b \neq 0$, where we multiply in the obvious way and add by taking common denominators. The field $F(R)$ will have the property that it is, in a sense we will shortly make precise, the smallest field into which you can embed the integral domain R.

To do this a bit more formally, define a relation on $R \times R\backslash\{0\}$ by setting $(a, b) \sim (c, d)$ if $a.d = b.c$ (to see where this comes from note that it expresses the equation $a/b = c/d$ without using division).

Lemma 2.9. *The relation $\sim$ is an equivalence relation.*

Proof. The only thing which requires work to check is that the relation is transitive. Indeed suppose that $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$. Then we have $ad = bc$ and $cf = de$ and need to check that $(a, b) \sim (e, f)$, that is, $af = be$. But this holds if

$$af - be = 0 \iff d.(af - be) = 0 \iff (ad).f - b.(de) = 0 \iff (bc).f - b.(cf) = 0,$$

as required (note in the first "if and only if" we used the fact that R is an integral domain and that $d \neq 0$. □

Write $\frac{a}{b}$ for the equivalence class of a pair (a, b) and denote the set of equivalence classes as $F(R)$.

Lemma 2.10. *The binary operations $(R \times R\backslash\{0\}) \times (R \times R\backslash\{0\}) \to R \times R\backslash\{0\}$ given by:*

$$((a, b), (c, d)) \mapsto (ad + bc, bd)$$
$$((a, b), (c, d)) \mapsto (ac, bd)$$

induce binary operations on $F(R)$.

Proof. Note first that since R is an integral domain and b, d are nonzero, bd is also nonzero, hence the above formulas do indeed define binary operations on $R \times R\backslash\{0\}$. To check that they induce binary operations on $F(R)$ we need to check that the equivalence class of the pairs on the right-hand side depends only on the equivalence classes of the two pairs on the left-hand side. We check this for the first operation (the second one being similar but easier).

Suppose that $(a_1, b_1) \sim (a_2, b_2)$ and $(c_1, d_1) \sim (c_2, d_2)$, so that $a_1b_2 = a_2b_1$ and $c_1d_2 = c_2d_1$. Then we need to show that $(a_1d_1 + b_1c_1, b_1d_1) \sim (a_2d_2 + b_2c_2, b_2d_2)$, which holds if and only if

$$(a_1d_1 + b_1c_1)(b_2d_2) = (a_2d_2 + b_2c_2)(b_1d_1) \iff (a_1b_2d_1d_2) + (b_1b_2c_1d_2) = (a_2b_1d_1d_2 + b_1b_2c_2d_1),$$

but $a_1b_2 = a_2b_1$ so $a_1b_2d_1d_2 = a_2b_1d_1d_2$ and $b_2c_1 = b_1c_2$ so that $b_1b_2c_1d_2 = b_1b_2c_2d_1$ and we are done. □

Let $+$ and $\times$ denote the binary operations the first and second operations above induce on $F(R)$. Thus we have

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \quad \frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}.$$

Theorem 2.11. *The above formulas give well-defined addition and multiplication operations on $F = F(R)$ the set of equivalence classes $\{\frac{a}{b} : a, b \in R, b \neq 0\}$, and F is a field with respect to these operations with additive identity $\frac{0}{1}$ and multiplicative identity $\frac{1}{1}$. Moreover there is a unique injective homomorphism $\iota \colon R \to F(R)$ sending $a \mapsto \frac{a}{1}$.*

Proof. (*Non-examinable*). One just has to check that the axioms for a field are satisfied. The ring axioms are routine to check by calculating in $R \times R\backslash\{0\}$. To see that $F(R)$ is a field, note that $(a, b) \sim (0, 1)$ if and only if $a = 0$. Thus if $\frac{a}{b} \neq 0$ then $a \neq 0$ and so $\frac{b}{a} \in F(R)$, and by definition $\frac{a}{b}.\frac{b}{a} = \frac{a.b}{a.b} = \frac{1}{1}$. Thus the multiplicative inverse of $\frac{a}{b}$ is $\frac{b}{a}$.

The map $a \mapsto \frac{a}{1}$ is certainly injective, since $(a, 1) \sim (b, 1)$ if and only if $a.1 = b.1$, that is, $a = b$. It is then immediate from the definitions that this map is a homomorphism as required.

□

Definition 2.12. The field $F(R)$ is known as the *field of fractions* of R.

Remark 2.13. All of this may look a little formal, but it is really no more than you have to do to construct the rational numbers from the integers. You should think of it as no more or less difficult (or to be fair, interesting) than that construction: essentially it just notices that all you needed to construct the rationals was the cancellation property which is the defining property of integral domains.

Finally we make precise the sense in which $F(R)$ is the smallest field containing R.

Proposition 2.14. *Let* k *be a field and let* $\theta\colon R \to \mathsf{k}$ *be an embedding (that is, an injective homomorphism). Then there is a unique injective homomorphism* $\tilde{\theta}\colon F(R) \to \mathsf{k}$ *extending* θ *(in the sense that* $\tilde{\theta}_{|\mathsf{k}} = \theta$*).*

Proof. (*non-examinable*): Suppose that $f\colon F(R) \to \mathsf{k}$ was such a homomorphism. Then by assumption $f(\frac{a}{1}) = \theta(a)$, and since homomorphism of rings respect multiplicative inverses this forces $f(\frac{1}{a}) = \theta(a)^{-1}$. But then, again because f is supposed to be a homomorphism, we must have $f(\frac{a}{b}) = f(\frac{a}{1}.\frac{1}{b}) = f(\frac{a}{1}).f(\frac{1}{b}) = \theta(a).\theta(b)^{-1}$. Thus if f exists, it has to be given by this formula.

The rest of the proof consists of checking that this recipe indeed works: Given $(a, b) \in R\times R\backslash\{0\}$ first define $\Theta(a, v) = \theta(a).\theta(b)^{-1}$. Then it is easy to check that Θ is constant on the equivalence classes of $\sim$ the relation defining $F(R)$, so that it induces a map $\tilde{\theta}\colon F(R) \to \mathsf{k}$. Finally it is straight-forward to see that this map is a homomorphism extending θ as required. □

Remark 2.15. Notice that this theorem implies that any field k of characteristic zero contains a (unique) copy of the rationals. Indeed by definition of characteristic, the unique homomorphism from $\mathbb{Z}$ to k is an embedding, and the above theorem shows that it therefore extends uniquely to an embedding of $\mathbb{Q}$ into k as claimed.

3. Homomorphisms and ideals

From now on we will assume all our rings are commutative. In this section we study the basic properties of ring homomorphisms, and establish an analogue of the "first isomorphism theorem" which you have seen already for groups. Just as for homomorphisms of groups, homomorphisms of rings have kernels and images.

Definition 3.1. Let $f\colon R \to S$ be a ring homomorphism. The *kernel* of f is

$$\ker(f) = \{r \in R : f(r) = 0\},$$

and the *image* of f is

$$\mathrm{im}(f) = \{s \in S : \exists r \in R, f(r) = s\}.$$

Just as for groups, the image of a homomorphism is a subring of the target ring. For kernels the situation is a little different. In the case of groups, kernels of homomorphisms are subgroups, but not any subgroup is a kernel – the kernels are characterised intrinsically by the property of being normal (*i.e.* perserved by the conjugation action of the group). We will show that the kernels of ring homomorphisms can similarly be characterised intrinsically, but the situation, because we have two binary operations, is slightly different: a kernel is both more and less than a subring. Indeed since homomorphisms are required to send 1 to 1, the kernel never contains 1 unless it is the entire ring, thus a kernel is *not* a subring. However, it is closed under addition and mulitplication (as is straight-forward to check) and because $0.x = 0$ for any x, it in fact obeys a stronger kind of closure with respect to multiplication:[11] If $x \in \ker(f)$ and $r \in R$ is any element of R, then $f(x.r) = f(x).f(r) = 0.f(r) = 0$ so that $x.r \in \ker(f)$. This motivates the following definition:

[11]This is analogous to the fact that kernels of group homomorphisms, are "more closed" than arbitrary subgroups.

Definition 3.2. Let R be a ring. A subset $I \subseteq R$ is called an *ideal* if it is a subgroup of $(R, +)$ and moreover for any $a \in I$ and $r \in R$ we have $a.r \in I$.

Lemma 3.3. *If $f : R \to S$ is a homomorphism, then $\ker(f)$ is an ideal.*

Proof. This is immediate from the definitions. □

3.1. **Basic properties of ideals.** Note that if I is an ideal of R which contains 1 then $I = R$. We will shortly see that in fact any ideal is the kernel of a homomorphism. First let us note a few basic properties of ideals:

Lemma 3.4. *Let R be a ring, and I, J ideals in R. Then $I + J$ $I \cap J$ and IJ are ideals, where*

$$I + J = \{i + j : i \in I, j \in J\}; \quad IJ = \{\sum_{k=1}^{n} i_k j_k : i_k \in I, j_k \in I, n \in \mathbb{N}\}.$$

Moreover we have $IJ \subseteq I \cap J$ and $I, J \subseteq I + J$.

Proof. For $I + J$ it is clear that this is an abelian subgroup of R, while if $i \in I$, $j \in J$ and $r \in R$, then $r(i + j) = (r.i) + (r.j) \in I + J$ as both I and J are ideals, hence $I + J$ is an ideal. Checking $I \cap J$ is an ideal is even less interesting – in fact it is easy to see that an arbitrary (not necessarily finite) intersection of ideals is an ideal. To see that IJ is an ideal, note that it is clear that the sum of two elements of IJ is clearly of the same form, and if $\sum_{k=1}^{n} x_k y_k \in IJ$ then

$$-\sum_{k=1}^{n} x_k y_k = \sum_{k=1}^{n} (-x_k).y_k \in IJ,$$

since if $x_k \in I$ then $-x_k \in I$. Thus IJ is an abelian subgroup. It is also straight-forward to check the multiplicative condition. The containments are all clear once you note that if $i \in I$ and $j \in J$ then ij in in $I \cap J$ because both I and J are ideals. □

In fact given a collection of ideals $\{I_\alpha : \alpha \in A\}$ in a ring R, their intersection $\bigcap_{\alpha \in A} I_\alpha$ is easily seen to again be an ideal. This easy fact is very useful for the following reason:

Definition 3.5. Given *any* subset T of R, one can define

$$\langle T \rangle = \bigcap_{T \subseteq I} I$$

(where I is an ideal) the ideal *generated* by T. We can also give a more explicit "from the ground up" description of the ideal generated by a subset X:

Lemma 3.6. *Let $T \subseteq R$. Then we have*

$$\langle T \rangle = \{\sum_{i=1}^{n} r_k t_k : r_k \in R, t_k \in T, n \in \mathbb{N}\}.$$

Proof. Let I denote the right-hand side. It is enough to check that I is an ideal and that I is contained in any ideal which contains T. We first check that I is an ideal – the proof is very similar to the proof that IJ is an ideal when I and J are: The multiplicative property is immediate: if $r \in R$ and $\sum_{k=1}^{n} r_k x_k \in I$ then $r(\sum_{k=1}^{n} r_k x_k) = \sum_{k=1}^{n} (r.r_k) x_k \in I$. Moreover the sum of two such elements is certainly of the same form, and I is closed under additive inverses since $-\sum_{k=1}^{n} r_k x_k = \sum_{k=1}^{n} (-r_k).x_k$,so that it is an additive subgroup of R.

It remains to show that if J is an ideal containing X then J contains I. But if $\{x_1, \ldots, x_k\} \subseteq T \subseteq J$ and $r_1, \ldots, r_k \in R$, then since J is an ideal certainly $r_k x_k \in J$ and hence $\sum_{k=1}^{n} r_k x_k \in J$. Since the x_k, r_k and $n \in \mathbb{N}$ were arbitrary it follows that $I \subseteq J$ as required. □

This is completely analogous the notion of the "span" of a subset in a vector space. If I and J are ideals, it is easy to see that $I + J = \langle I \cup J \rangle$. In the case where $T = \{a\}$ consists of a single element, we often write aR for $\langle a \rangle$.

Remark 3.7. Note that in the above, just as for span in a vector space, there is no need for the set X to be finite.

MATHEMATICAL INSTITUTE, OXFORD.