MODULES OVER PRINCIPAL IDEAL DOMAINS

N. P. STRICKLAND

1. Rings

A *commutative ring* is a set R of things that can be added, negated and multiplied in a sensible way to get new elements of R. More precisely, we require that the following axioms be satisfied:

(a) If $a, b \in R$ then $a + b \in R$. [closure under addition]
(b) There is an element $0 \in R$ such that $a + 0 = a$ for all $a \in R$. [additive identity]
(c) For each element $a \in R$ there is an element $-a \in R$ such that $a + (-a) = 0$. [additive inverses]
(d) $a + (b + c) = (a + b) + c$ for all $a, b, c \in R$. [associativity of addition]
(e) $a + b = b + a$ for all $a, b \in R$. [commutativity of addition]
(f) If $a, b \in R$ then $ab \in R$. [closure under multiplication]
(g) There is an element $1 \in R$ such that $1a = a$ for all $a \in R$. [multiplicative identity]
(h) $a(bc) = (ab)c$ for all $a, b, c \in R$. [associativity of multiplication]
(i) $ab = ba$ for all $a, b \in R$. [commutativity of multiplication]
(j) $a(b + c) = ab + ac$ for all $a, b, c \in R$. [distributivity]

In Example 8, we will discuss an example of a noncommutative ring. Everywhere else in these notes the word “ring” will mean “commutative ring”. We will use without comment various standard consequences of the axioms, such as the facts that $-(-a) = a$, $0.a = 0$ and $(-1).a = -a$.

Example 1. The set $\mathbb{Z}$ of integers is a commutative ring. If we add, subtract or multiply any two integers, we get another integer, so axioms (a), (c) and (f) hold. The numbers 0 and 1 are integers so axioms (b) and (g) hold. The remaining axioms are familiar properties of addition, subtraction and multiplication.

Example 2. The set $\mathbb{N}$ of natural numbers is not a commutative ring, because $0 \notin \mathbb{N}$, so there is no additive identity, in other words axiom (b) does not hold. Moreover, if $n \in \mathbb{N}$ then $-n \notin \mathbb{N}$, so $\mathbb{N}$ does not have additive inverses.

Example 3. The set $2\mathbb{Z}$ of even integers is not a commutative ring, because it does not contain the multiplicative identity element 1.

Example 4. The set $\mathbb{Q}$ of rational numbers, the set $\mathbb{R}$ of real numbers, and the set $\mathbb{C}$ of complex numbers, are all commutative rings.

Example 5. The set $X = \mathbb{R} \setminus \mathbb{Q}$ of irrational numbers is not a commutative ring. Indeed, we have $\pi \in X$ and $-\pi \in X$ but $0 = \pi + (-\pi) \notin X$, so X is not closed under addition. Moreover, $\sqrt{2} \in X$ but $\sqrt{2}.\sqrt{2} = 2 \notin X$, so X is not closed under multiplication. Even more obviously, X contains neither 0 nor 1, so it does not have an additive identity or a multiplicative identity.

Example 6. For any natural number n, the set $\mathbb{Z}_n = \{\overline{0}, \overline{1}, \dots, \overline{n-1}\}$ of integers modulo n is a commutative ring. For example, we have

$$\begin{aligned} \mathbb{Z}_4 &= \{\overline{0}, \overline{1}, \overline{2}, \overline{3}\} \\ \overline{2} + \overline{3} &= \overline{5} = \overline{1} \\ \overline{1} - \overline{2} &= \overline{-1} = \overline{3} \\ \overline{2} \,.\, \overline{3} &= \overline{6} = \overline{2}. \end{aligned}$$

Date: December 11, 2013.

Strictly speaking, we should say that a ring is a set R *together with a definition of addition, subtraction and multiplication* such that the axioms hold. In the examples discussed above, there is an obvious way to define these operations. We next discuss an example where the definition is not so obvious, and another example where there are two different possible definitions.

Example 7. Let R be the set of all subsets of $\mathbb{N}$. Define

$$\begin{aligned} 0 &= \emptyset \\ 1 &= \mathbb{N} \\ A+B &= \{n \in \mathbb{N} \mid n \in A \text{ or } n \in B \text{ but not both. }\} \\ -A &= A \\ AB &= A \cap B. \end{aligned}$$

For example, we have

$$\{1,2,3,4,5\} + \{4,5,6,7\} = \{1,2,3,6,7\}.$$

One can check that these operations make R into a commutative ring. Axioms (a), (b), (e), (f), (g), (h) and (j) are all easy. For axiom (c), note that $A + (-A) = A + A$. In general, a number n lies in $B + C$ if it lies in B but not in C, or if it lies in C but not in B. It is of course impossible for a number to lie in A but not in A, so $A + A = \emptyset$. For axiom (d), one checks that $A + (B + C)$ consists of the numbers that either

(i) lie in A but not in B or in C; or
(ii) lie in B but not in A or in C; or
(iii) lie in C but not in A or in B; or
(iv) lie in $A \cap B \cap C$.

One then checks that $(A + B) + C$ has the same description, so $A + (B + C) = (A + B) + C$. Similarly, one checks that $A(B + C)$ consists of the numbers that either

(i) lie in A and in B but not in C; or
(ii) lie in A and in C but not in B.

One then checks that $AB + AC$ has the same description, so $A(B + C) = AB + AC$.

Example 8. Let R be the set of all 2×2 matrices of integers. If $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and $A' = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$ are two elements of R we define their matrix product $A \circ A'$ and their pointwise product $A * A'$ by

$$\begin{aligned} A \circ A' &= \left(\begin{smallmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{smallmatrix}\right) \\ A * A' &= \left(\begin{smallmatrix} aa' & bb' \\ cc' & dd' \end{smallmatrix}\right). \end{aligned}$$

If we use $A \circ A'$ as our definition of multiplication, then all axioms are satisfied except for axiom (k), because $A \circ A' \neq A' \circ A$ in general. The element 0 in axiom (b) is the matrix $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, and the element 1 in axiom (h) is the matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. The set R with this definition of multiplication is a noncommutative ring.

If we instead use $A * A'$ as our definition of multiplication, then all the axioms hold. The element 0 in axiom (b) is again the matrix $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, but the element 1 in axiom (h) is the now matrix $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$.

Example 9. Let p be a prime number. We write $\mathbb{Z}[\frac{1}{p}]$ for the set of rational numbers x that can be written in the form $x = a/p^k$ for some integer a and some nonnegative integer k. For example, $5/2$, $-1/8 = -1/2^3$ and $9/48 = 3/2^3$ are elements of $\mathbb{Z}[\frac{1}{2}]$, but $3/5$ and $1/13$ are not. Note that any integer n can be written as n/p^0 and thus lies in $\mathbb{Z}[\frac{1}{p}]$, in other words $\mathbb{Z} \subseteq \mathbb{Z}[\frac{1}{p}]$, and in particular $0, 1 \in \mathbb{Z}[\frac{1}{p}]$. Now suppose that $x, y \in \mathbb{Z}[\frac{1}{p}]$, so we can write $x = a/p^j$ and $y = b/p^k$ for some integers a, b, j, k with $j, k \geq 0$. Note that

$$\begin{aligned} x + y &= (ap^k + bp^j)/p^{j+k} \\ xy &= (ab)/p^{j+k} \\ -x &= (-a)/p^j. \end{aligned}$$

As $ap^k + bp^j$, ab and $-a$ are integers and $j + k$ and j are nonnegative integers, this shows that $x + y$, xy and $-x$ lie in $\mathbb{Z}[\frac{1}{p}]$. Thus $\mathbb{Z}[\frac{1}{p}]$ is closed under addition, multiplication and negation, so it is a subring of $\mathbb{Q}$.

Example 10. Now let $\mathbb{Z}_{(p)}$ be the set of rational numbers x that can be written in the form a/b, where a and b are integers and b is not divisible by p. For example, $123/101$ and $-4/12 = (-1)/3$ and $8 = 8/1$ are elements of $\mathbb{Z}_{(2)}$, but $5/6$ and $1/8$ are not. As any integer n can be written as $n/1$ and 1 is not divisible by p, we see that $\mathbb{Z} \subseteq \mathbb{Z}_{(p)}$. Now suppose that $x, y \in \mathbb{Z}_{(p)}$, so we can write $x = a/b$ and $y = c/d$ for some integers a, b, c and d, where b and d are not divisible by p. As p is prime this means that bd is also not divisible by p. We have

$$\begin{aligned} x + y &= (ad + bc)/(bd) \\ xy &= (ac)/(bd) \\ -x &= -a/b. \end{aligned}$$

As $ad + bc$, ac, $-a$, b and bd are integers, and bd and b are not divisible by p, this means that $x + y$, xy and $-x$ lie in $\mathbb{Z}_{(p)}$. Thus $\mathbb{Z}_{(p)}$ is a subring of $\mathbb{Q}$, called the ring of integers localised at p. (There is a long story coming from algebraic geometry that explains why the word "localised" is appropriate.)

Example 11. We write $\mathbb{Z}[i]$ for the set of complex numbers of the form $a + bi$, where a and b are integers (possibly zero). Thus 7, $6 - 4i$ and $12i$ are elements of $\mathbb{Z}[i]$, but $2/3$ and $1 - i/5$ are not. Note that

$$\begin{aligned} (a + bi) + (c + di) &= (a + c) + (b + d)i \\ (a + bi)(c + di) &= (ac - bd) + (ad + bc)i \\ -(a + bi) &= (-a) + (-b)i. \end{aligned}$$

It follows easily that $\mathbb{Z}[i]$ is closed under addition, multiplication and negation, so it is a subring of $\mathbb{C}$. The elements of $\mathbb{Z}[i]$ are called *Gaussian integers*. More generally, for any integer m we can consider the set $\mathbb{Z}[\sqrt{m}]$ of complex numbers of the form $a + b\sqrt{m}$. We have

$$\begin{aligned} (a + b\sqrt{m}) + (c + d\sqrt{m}) &= (a + c) + (b + d)\sqrt{m} \\ (a + b\sqrt{m})(c + d\sqrt{m}) &= (ac + bdm) + (ad + bc)\sqrt{m} \\ -(a + b\sqrt{m}) &= (-a) + (-b)\sqrt{m}. \end{aligned}$$

Using this we see that $\mathbb{Z}[\sqrt{d}]$ is a subring of $\mathbb{C}$. Of course, if $d > 0$ then it is actually a subring of $\mathbb{R}$.

Example 12. We write $\mathbb{Z}[x]$ for the set of all polynomials with integer coefficients. For example, $7x^3 - 22x + 3$ and x^{1001} are elements of $\mathbb{Z}[x]$ but $(x + 1)/(x - 1)$ and $x^2 - 1/2$ and $x - x^{-1}$ are not. The general form of an element of $\mathbb{Z}[x]$ is $f(x) = a_0 + a_1x + \ldots + a_nx^n$ for some integer $n \geq 0$ and integers $a_0, \ldots, a_n$. Integers are polynomials of degree zero, so $\mathbb{Z} \subseteq \mathbb{Z}[x]$. The usual operations of addition, multiplication and negation of polynomials make $\mathbb{Z}[x]$ into a ring.

More generally, given any ring R we can consider the ring $R[x]$ of polynomials with coefficients in R. For example, we can consider $f(x) = \overline{2}x^2 + \overline{3} \in \mathbb{Z}_6[x]$ and $g(x) = \overline{3}x + \overline{2} \in \mathbb{Z}_6[x]$ and we find that

$$\begin{aligned} f(x)g(x) &= \overline{6}x^3 + \overline{4}x^2 + \overline{9}x + \overline{6} \\ &= \overline{4}x^2 + \overline{3}x. \end{aligned}$$

This gives us rings $\mathbb{Z}[x] \subseteq \mathbb{Q}[x] \subseteq \mathbb{R}[x] \subseteq \mathbb{C}[x]$. We can also use more than one variable; for example, we have a ring $\mathbb{Q}[x, y, z]$ containing elements like $(x^2 + y^2 + z^2)/4$ or $1 + xyz$ (but not x/y or $\sqrt{2}x$ or e^{x+y}).

Remark 13. Here is a slightly subtle point about polynomials. Consider the element $f(x) = x^3 - x \in \mathbb{Z}_3[x]$. You might like to argue that $f = 0$, for the following reason. The elements of $\mathbb{Z}_3$ are $\overline{0}$, $\overline{1}$ and $\overline{2}$. We have

$$\begin{aligned} f(\overline{0}) &= \overline{0}^3 - \overline{0} = \overline{0} \\ f(\overline{1}) &= \overline{1}^3 - \overline{1} = \overline{1} - \overline{1} = \overline{0} \\ f(\overline{2}) &= \overline{2}^3 - \overline{2} = \overline{8} - \overline{2} = \overline{0}. \end{aligned}$$

Thus, *if we think of f as a function*, we have $f = 0$. However, in this context we will *not* think of f as a function, but just as a formal expression. We only regard two polynomials as the same if all their coefficients are the same, not if all their values are the same.

In the more usual context of polynomials with coefficients in $\mathbb{C}$ or subrings of $\mathbb{C}$, it turns out that we do not need to worry about this distinction: two polynomials have the same values if and only if they have the same coefficients. You will doubtless have used this fact many times in the past, probably without noticing it.

2. Domains and fields

We next make a few remarks about the zero ring. The set $\{0\}$ (with the obvious definitions of addition, multiplication and negation) is a ring. The only subtle point is the existence of a multiplicative identity. The axiom says that there should be an element $1 \in R$ such that $1.a = a$ for all $a \in R$. As 0 is the only element of R, this reduces to the requirement that $1.0 = 0$. As $0.0 = 0$, we may take $1 = 0$.

Conversely, let R be an arbitrary ring, so we have elements $0, 1 \in R$. If $1 = 0$ then for all $a \in R$ we have $a = 1.a = 0.a = 0$, so $R = \{0\}$. Thus, the zero ring is the only ring in which $1 = 0$.

Definition 14. Let R be a commutative ring. We say that R is an *integral domain* (or sometimes just a *domain*) if $1 \neq 0$ and for all $a, b \in R$ with $a, b \neq 0$ we have $ab \neq 0$.

Example 15. The ring $\mathbb{Z}_6$ is not a domain. Indeed, $\overline{2}$ and $\overline{3}$ are nonzero elements of $\mathbb{Z}_6$, but $\overline{2}\overline{3} = \overline{6} = \overline{0}$. More generally, if $n > 1$ and n is not prime then $\mathbb{Z}_n$ is not a domain, because we can write $n = ab$ for some a, b with $1 < a, b < n$, and then we have $\overline{a}, \overline{b} \neq 0$ but $\overline{a}\overline{b} = 0$.

Example 16. Now suppose instead that p is prime; I claim that $\mathbb{Z}_p$ is a domain. Indeed, if $\overline{a}$ and $\overline{b}$ are nonzero elements of $\mathbb{Z}_p$ then a and b are not divisible by p and p is prime so ab is not divisible by p, so $\overline{ab}$ is nonzero. It is also clear that $1 \neq 0$.

Example 17. Consider the ring R of subsets of $\mathbb{N}$, as in example 7. If A and B are any two nonempty, disjoint subsets of $\mathbb{N}$ then A and B are nonzero elements of R but $AB = A \cap B = 0$. (For example, we could take $A = \{1, 2, 3\}$ and $B = \{4, 5, 6, 7\}$, or just take $A = \{1\}$ and $B = \{2\}$.) This shows that R is not a domain.

Example 18. If a and b are nonzero real numbers then of course $ab \neq 0$, so $\mathbb{R}$ is a domain. It follows that any subring of $\mathbb{R}$ is a domain, so $\mathbb{Z}$ and $\mathbb{Q}$ are domains, as are $\mathbb{Z}[\frac{1}{p}]$ and $\mathbb{Z}_{(p)}$ for any prime p. If a and b are nonzero complex numbers then $|a|, |b| > 0$ so $|ab| = |a||b| > 0$ so $ab \neq 0$, so $\mathbb{C}$ is a domain, as is any subring of $\mathbb{C}$. For example $\mathbb{Z}[i]$ is a domain.

Proposition 19. *If R is an integral domain, then so is $R[x]$.*

Proof. Suppose that $f(x)$ and $g(x)$ are nonzero elements of $R[x]$. Let n be the degree of f, in other words the highest power of x that occurs in f with a nonzero coefficient. Then $f(x) = ax^n + \text{ lower terms }$ for some $a \in R$ with $a \neq 0$. Similarly, if m is the degree of g then $g(x) = bx^m + \text{ lower terms }$ for some $b \in R$ with $b \neq 0$. When we multiply out $f(x)g(x)$, we get abx^{n+m} and then lots of other terms involving powers x^k with $k < n + m$, which therefore cannot cancel out the term abx^{n+m}. As R is a domain we have $ab \neq 0$ so the term abx^{n+m} is nonzero, so we have $f(x)g(x) \neq 0$, as required. □

Definition 20. Let a be an element of a ring R. We say that R is *invertible* (or is a *unit* in R) if there is an element b such that $ab = 1$. We write $R^\times$ for the set of invertible elements of R. We say that R is a *field* if $1 \neq 0$ and every nonzero element is invertible.

Remark 21. If $ab = 1$ and $ac = 1$ then $b = b.1 = b(ac) = (ba)c = 1c = c$, so $b = c$. Thus the element b is unique if it exists. If so, it is called the *inverse* of a, and written a^{-1} or $1/a$.

Remark 22. It is not hard to see that the set $R^\times$ forms an Abelian group under multiplication.

Example 23.
(1) The only units in $\mathbb{Z}$ are 1 and -1.
(2) The only units in $\mathbb{Z}[i]$ are 1, -1, i and $-i$.
(3) Note that $(2 + \sqrt{3})(2 - \sqrt{3}) = 2^2 - (\sqrt{3})^2 = 1$. This shows that the number $u = 2 + \sqrt{3}$ is a unit in $\mathbb{Z}[\sqrt{3}]$, with inverse $u^{-1} = 2 - \sqrt{3}$. In fact, the units in this ring are precisely the numbers of the form $\pm u^n$ with $n \in \mathbb{Z}$.
(4) If K is a field, then the only units in $K[x]$ are the nonzero constant polynomials.

(5) The units in $\mathbb{Z}_{(2)}$ are the numbers of the form a/b where a and b are odd integers and $b \neq 0$.

Example 24. The rings $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are well-known to be fields.

Example 25. The ring $\mathbb{Z}$ is not a field, because 2 is a nonzero element of $\mathbb{Z}$ and there is no integer b such that $2b = 1$. Similarly, for any prime p the ring $\mathbb{Z}_{(p)}$ is not a field because p is a nonzero element of $\mathbb{Z}_{(p)}$ with no inverse in $\mathbb{Z}_{(p)}$.

Proposition 26. *A field is an integral domain.*

Proof. Let K be a field, and let a and b be nonzero elements of K. Then a and b are invertible so we have elements a^{-1} and b^{-1} in K and we can define $c = a^{-1}b^{-1}$. We then have $(ab)c = 1 \neq 0 = 0.c$, so we must have $ab \neq 0$. □

Proposition 27. *The ring $\mathbb{Z}_n$ is a field if and only if n is prime.*

Proof. If n is not prime, we have seen that $\mathbb{Z}_n$ is not a domain and thus is certainly not a field. Now suppose that p is a prime number, and that x is a nonzero element of $\mathbb{Z}_p$. Then $x = \overline{a}$ for some a with $0 < a < p$. It follows that a is coprime to p, so the Euclidean algorithm gives us integers u, v such that $au + pv = 1$. Thus $\overline{a}.\overline{u} = \overline{1 - pv} = \overline{1}$, so $\overline{u}$ is an inverse for x. It is clear that $1 \neq 0$, so $\mathbb{Z}_p$ is a field. □

3. Fields of fractions

Let R be an integral domain. The *field of fractions* of R is the set of "fractions" a/b, where a and b are elements of R, and $b \neq 0$. We regard two fractions a/b and c/d as the same if and only if $ad = bc$. We add, multiply, negate and invert fractions by the usual rules:

$$\begin{aligned} a/b + c/d &= (ad + bc)/(bd) \\ (a/b).(c/d) &= (ac)/(bd) \\ -(a/b) &= (-a)/b \\ (a/b)^{-1} &= b/a \qquad\qquad (\text{if } a \neq 0). \end{aligned}$$

We write $Q(R)$ for the set of all such fractions, which is a field, called the *field of fractions* of R.

There is a small subtlety here, which explains why we needed to start with an integral domain. When we say that a/b and c/d are "regarded as the same" if $ad = bc$, we really mean that a/b is an equivalence class for a suitable equivalence relation. We start with the set

$$F(R) = \{(a, b) \mid a, b \in R \text{ and } b \neq 0\}.$$

We then introduce a relation $\sim$ on the set $F(R)$, defined by $(a, b) \sim (c, d)$ if $ad = bc$. Clearly $ab = ab$ so $(a, b) \sim (a, b)$ for all $(a, b) \in F(R)$, so the relation is reflexive. Clearly also $ad = bc$ iff $da = cb$, so $(a, b) \sim (c, d)$ iff $(c, d) \sim (a, b)$; this means that our relation is symmetric. Finally, suppose that $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$, so $ad = bc$ and $cf = de$. Then

$$(af)d = (ad)f = (bc)f = b(cf) = b(de) = (be)d,$$

and $d \neq 0$ and R is a domain so we can cancel the d's to get $af = be$. Thus $(a, b) \sim (e, f)$, proving that our relation is transitive and thus an equivalence relation. We now define $Q(R) = F(R)/\sim$ and write a/b for the equivalence class of (a, b).

There are now some things to check to make sure that our definition of addition and so on are unambiguous. For example, suppose we have a fraction x and we want to add $y = c/d$ to it. Suppose we have two different expressions for x, say $x = a/b$ and $x = a'/b'$. The first expression gives $x + y = (ad + bc)/(bd)$, and the second expression gives $x + y = (a'd + b'c)/(b'd)$. For everything to be consistent we must have $(ad+bc)/(bd) = (a'd+b'c)/(b'd)$, or equivalently $(ad+bc, bd) \sim (a'd+b'c, b'd)$, or equivalently $(ad+bc)b'd = (a'd+b'c)bd$, or equivalently $ab'd^2 + bb'cd = a'bd^2 + bb'cd$. By assumption we have $a/b = a'/b'$ so $ab' = a'b$ so $ab'd^2 + bb'cd = a'bd^2 + bb'cd$ as required. Thus addition is well-defined, and one checks by similar methods that the other operations are also well-defined.

To convince you that this digression is necessary, we will exhibit some inconsistencies that arise if we try to work with fractions in a ring R that is not a domain. For concreteness, we consider $R = \mathbb{Z}_6$. As $\overline{3}.\overline{2} = \overline{0} = \overline{3}.\overline{4}$ we have $\overline{3}/\overline{3} = \overline{4}/\overline{2}$ but of course $\overline{3}/\overline{3} = \overline{1}/\overline{1}$ and $\overline{4}/\overline{2} = \overline{2}/\overline{1}$ so apparently we have $\overline{1}/\overline{1} = \overline{2}/\overline{1}$. This contradicts the rule that $a/b = c/d \Leftrightarrow ad = bc$, because $\overline{1} \neq \overline{2}$.

4. Ideals

Definition 28. An *ideal* in a ring R is a subset $I \subseteq R$ such that

(a) $0 \in I$
(b) if $a, b \in I$ then $a + b \in I$
(c) if $a \in R$ and $b \in I$ then $ab \in I$.

Example 29. Let I be the set of even integers; then I is an ideal in $\mathbb{Z}$. Indeed, 0 is even so (a) holds; the sum of two even integers is even so (b) holds; and the product of an even integer with any other integer is still even so (c) holds.

Example 30. Put $R = \mathbb{Z}[x]$ and $I = \{f \in \mathbb{Z}[x] \mid f(1) = 0\}$. For example $x^{10} - x \in I$ and $(x-3)(x-2)(x-1) \in I$ but $x + 7 \not\in I$ (because $1 + 7 \neq 0$) and $\frac{1}{2}x^2 - x + \frac{1}{2} \not\in I$ (because the coefficients are not integers, so $\frac{1}{2}x^2 - x + \frac{1}{2} \not\in \mathbb{Z}[x]$). Clearly the zero polynomial is an element of I, so (a) holds. If $f, g \in I$ then $f(1) = g(1) = 0$ so $(f + g)(1) = f(1) + g(1) = 0 + 0 = 0$, so $f + g \in I$; thus (b) holds. If $f \in I$ and g is any polynomial then $(gf)(1) = g(1)f(1) = g(1).0 = 0$ so $gf \in I$; thus (c) holds. This shows that I is an ideal in R.

Example 31. Put $R = \mathbb{Z}[x]$ and $I = \{$ constant polynomials $\} \subseteq R$. Then I is *not* an ideal. Axioms (a) and (b) certainly hold. Moreover, if a and b are elements of I then $ab \in I$ also. However, axiom (c) says more than this: it says that if $b \in I$ and a is *any* element of R, not necessarily in I, then ab must be in I. However, $x \in R$ and $1 \in I$ but $x.1 \not\in I$ so axiom (c) is violated.

Example 32. Let K be a field. It is easy to see that $\{0\}$ and K itself are ideals in K; I claim that these are the only ideals. Indeed, let I be an ideal in K. If $I \neq \{0\}$ then we have some nonzero element $b \in I$. For any element $c \in K$ we have $cb^{-1} \in K$ and $b \in I$ so axiom (c) tells us that $(cb^{-1})b \in I$, or in other words $c \in I$. This means that $I = K$, as required.

Example 33. Let R be any ring, and let x be any element of R. Define $Rx = \{ux \mid u \in R\}$. I claim that this is an ideal (called the *principal ideal* generated by x). First, we have $0 = 0.x \in Rx$, so axiom (a) holds. Second, if $a, b \in Rx$ then there exist u, v such that $a = ux$ and $b = vx$ so $a + b = (u + v)x$, so $a + b \in Rx$. This shows that (b) holds. Finally, if $a \in R$ and $b \in Rx$ then $b = vx$ for some $v \in R$ so $ab = (av)x \in Rx$, so (c) holds.

In example 29, the ideal I is just $\mathbb{Z}.2$. In example 30, the ideal I is just $\mathbb{Z}[x].(x - 1)$.

Definition 34. A *principal ideal domain* or *PID* is a domain R such that every ideal in R is a principal ideal.

If K is a field then the only ideals in K are $\{0\}$ and K. Clearly $\{0\} = K.0$ and $K = K.1$, so both of these ideals are principal. Thus K is a principal ideal domain.

We will see later that if I is a nonzero ideal in $\mathbb{Z}$ and d is the smallest strictly positive number in I then $I = \mathbb{Z}.d$; this implies that $\mathbb{Z}$ is a PID. From this one can deduce that $\mathbb{Z}[\frac{1}{p}]$ and $\mathbb{Z}_{(p)}$ are PID's.

We will also see in the next section that for any field K, the polynomial ring $K[x]$ is a PID. However, $\mathbb{Z}[x]$ is not a PID, because the set $I = \{f \in \mathbb{Z}[x] \mid f(0)$ is even $\}$ is a non-principal ideal (for example).

5. Divisibility

Definition 35. We say that a *divides* b (and write $a|b$) if there is an element $x \in R$ such that $b = ax$. We say that elements $a, b \in R$ are *associates* (and write $a \sim b$) if $a|b$ and $b|a$.

Example 36. In $\mathbb{Z}$, integers n and m are associates if and only if $m = \pm n$. In $\mathbb{R}[x]$, polynomials f and g are associates if and only if each is a constant multiple of the other.

Proposition 37. *a divides b if and only if $Rb \leq Ra$.*

Proof. If a divides b then $b = ax$ for some x. Let c be an element of Rb. Then $c = by$ for some $y \in R$, so $c = axy$, so $c \in Ra$. Thus $Rb \leq Ra$.

Conversely, suppose that $Rb \leq Ra$. Clearly $b \in Rb$ so $b \in Ra$ so $b = ax$ for some x, in other words $a|b$. □

Proposition 38. *The following are equivalent:*

(i) *a and b are associates.*
(ii) *$Ra = Rb$.*
(iii) *There is a unit $u \in R^\times$ such that $b = ua$.*

Moreover, the relation $\sim$ is an equivalence relation.

Proof. a and b are associates iff $a|b$ and $b|a$. By the previous proposition this is the same as to say that $Rb \leq Ra$ and $Ra \leq Rb$, or in other words $Ra = Rb$. Thus (i) is equivalent to (ii).

If (iii) holds then $b = ua$ and $a = u^{-1}b$ so $a|b$ and $b|a$ so (i) holds. Conversely, if (i) holds then $b = ax$ and $a = by$ for some $x, y \in R$. It follows that $axy = by = a$ so $a(xy - 1) = 0$. If $a \neq 0$ then (as R is a domain) we conclude that $xy = 1$, so x is a unit, so (iii) holds. In the exceptional case where $a = 0$ we note that $b = ax = 0$ also so $b = 1.a$ and 1 is a unit so again (iii) holds. This proves that (i) is equivalent to (iii).

We now know that $a \sim b$ iff $Ra = Rb$. We have $Ra = Ra$ so $a \sim a$. If $a \sim b$ then $Ra = Rb$ so $Rb = Ra$ so $b \sim a$. If $a \sim b$ and $b \sim c$ then $Ra = Rb = Rc$ so $a \sim c$. Thus, $\sim$ is an equivalence relation. □

Definition 39. We say that an element $d \in R$ is a *greatest common divisor* (or *gcd*) of a and b if $Rd = Ra + Rb$.

Proposition 40. *Any two elements a, b have a gcd. If d and d' are two different gcd's for a and b then $d \sim d'$.*

Proof. $Ra + Rb$ is an ideal in R and R is a principal ideal domain, so $Ra + Rb$ has the form Rd for some d, and this d is then a gcd for a and b. If d' is another gcd then $Rd = Ra + Rb = Rd'$, so $d \sim d'$ by the previous proposition. □

Remark 41. In $\mathbb{Z}$, the gcd's of -6 and 15 are $+3$ and -3. In this context it is natural to think of $+3$ as "the" gcd, but for an arbitrary PID there will be many different gcd's and there will be no way to pick out one of them as the "right" one.

The next proposition justifies the term "greatest common divisor".

Proposition 42. *An element $d \in R$ is a gcd of a and b if and only if*

(i) *$d|a$ and $d|b$*
(ii) *if c is some other element such that $c|a$ and $c|b$ then $c|d$.*

Proof. First suppose that d is a gcd of a and b, so $Rd = Ra + Rb$. Clearly $a \in Ra + Rb$ (as $a = 1.a + 0.b$) so $a \in Rd$ so $d|a$. Similarly $d|b$, so (i) holds. On the other hand, we have $d \in Rd$ so $d \in Ra + Rb$ so $d = ax + by$ for some x, y. Now suppose we have some other element c such that $c|a$ and $c|b$. This means that $a = ca'$ and $b = cb'$ for some a', b', so

$$d = ax + by = ca'x + cb'y = c(a'x + b'y),$$

so $c|d$. Thus (ii) holds.

Conversely, suppose that (i) and (ii) hold for some element d. Let d' be a gcd of a and b, so the previous paragraph shows that (i) and (ii) hold for d' as well. As (i) holds for d' we see that $d'|a$ and $d'|b$. As (ii) holds for d we deduce that $d|d'$. As (i) holds for d we see that $d|a$ and $d|b$. As (ii) holds for d' we deduce that $d'|d$. Thus $d \sim d'$ so $Rd = Rd'$. As d' is a gcd we have $Rd' = Ra + Rb$, so $Rd = Ra + Rb$, so d is a gcd as well. □

Definition 43. We say that two elements $a, b \in R$ are *coprime* if 1 is a gcd of a and b, or equivalently every gcd of a and b is a unit, or equivalently $Ra + Rb = R$.

Proposition 44. *If a is coprime to b and to c, then a is coprime to bc.*

Proof. By assumption we have $1 = ax + by = au + cv$ for some x, y, u, v. It follows that

$$1 = (ax + by)(au + cv) = a^2ux + acvx + abuy + bcyv = a(aux + cvx + buy) + (bc)(yv).$$

In other words, if we put $s = aux + cvx + buy$ and $t = yv$ we get $as + (bc)t = 1$, which shows that a and bc are coprime. □

Proposition 45. *If a and b are coprime and c is divisible by both a and b, then c is divisible by ab.*

Proof. By assumption there exist $u, v \in R$ such that $c = au = bv$. Morever, as a and b are coprime there exist $x, y \in R$ such that $ax + by = 1$. Thus $c = 1.c = axc + byc$. If we substitute $c = bv$ in the first term and $c = au$ in the second term we get $c = axbv + byau = (xv + yu)ab$, so c is divisible by ab as claimed. □

6. The Chinese remainder theorem

Theorem 46. *Let R be a PID, and let $a_1, \dots, a_n$ be elements of R such that a_i and a_j are coprime whenever $i \neq j$. Then there are elements $e_1, \dots, e_n \in R$ such that*

(a) $e_i = 1 \pmod{a_i}$ *for all i.*
(b) $e_i = 0 \pmod{a_j}$ *for all $j \neq i$.*
(c) $e_1 + \dots + e_n = 1$.

Example 47. Take $R = \mathbb{Z}$ and $a_1 = 3$, $a_2 = 4$ and $a_3 = 5$. As $(3, 4) = (3, 5) = (4, 5) = 1$, the theorem applies, so there must be integers e_1, e_2 and e_3 such that:

(1) e_1 is divisible by 4 and 5 and is congruent to 1 mod 3.
(2) e_2 is divisible by 3 and 5 and is congruent to 1 mod 4.
(3) e_3 is divisible by 3 and 4 and is congruent to 1 mod 5.

In fact, we can take $e_1 = 40$ and $e_2 = 45$ and $e_3 = -84$.

Proof. Put $a = a_1a_2 \dots a_n$ and $b_i = a/a_i = \prod_{j \neq i} a_j$. As a_1 is coprime to a_2, a_3 up to a_n, Proposition 44 tells us that a_1 is coprime to the product $a_2a_3 \dots a_n = b_1$. Similarly, a_i is coprime to b_i for all i. This means that there exist $u_i, v_i \in R$ such that $u_ia_i + v_ib_i = 1$. Put $e_i = v_ib_i = 1 - u_ia_i$. Because $e_i = 1 - u_ia_i$ we have $e_i = 1 \pmod{a_i}$. Now suppose that $j \neq i$. Then b_i is divisible by a_j and $e_i = v_ib_i$ so $e_i = 0 \pmod{a_j}$. Thus (a) and (b) in the theorem are satisfied. However, (c) may not be satisfied until we make a slight adjustment.

To see this, put $e = e_1 + \dots + e_n$. For any i, we can consider e modulo a_i. The term e_i is 1 mod a_i and the other terms are 0 mod a_i, so altogether we get $e = 1 \pmod{a_i}$. This means that $1 - e$ is divisible by a_i for all i, and the various different a_i's are coprime, so $1 - e$ is divisible by $a = \prod_i a_i$, by Proposition 45.

Now put $e'_n = 1 - e_1 - \dots - e_{n-1} = (1 - e) + e_n$. Then $e'_n = e_n \pmod{a}$, so $e'_n = e_n \pmod{a_i}$ for all i, so properties (a) and (b) still hold if we replace e_n by e'_n. After making this replacement we clearly have $e_1 + \dots + e_n = 1$, so (c) holds. □

7. Primes and factorisation

Definition 48. Let R be a PID. An element $p \in R$ is *prime* if it is a non-unit, but it cannot be factored as $p = ab$ where both a and b are non-units.

Remark 49. If p is a prime number in the usual sense, then the only possible factorisations of p in $\mathbb{Z}$ are $p = p.1$ and $p = (-p).(-1)$. As 1 and -1 are the units in $\mathbb{Z}$, we see that p is prime in the sense of the above definition. It is not hard to see that $-p$ is also prime. In fact, the prime elements of $\mathbb{Z}$ are precisely the numbers of the form p or $-p$ where p is a prime in the usual sense.

Remark 50. It is usual to define primes in commutative rings in a rather different way, and to prove that the other definition is equivalent to the above definition for PID's. As we will only be concerned with PID's, we have used the definition which is most obviously analogous to the ordinary definition of prime numbers.

Proposition 51. *The primes in $\mathbb{C}[x]$ are the nonconstant linear polynomials, or in other words the polynomials of the form $ax + b$ with $a, b \in \mathbb{C}$ and $a \neq 0$.*

Proof. Consider a polynomial $p(x) = ax + b$ with $a \neq 0$. The only units in $\mathbb{C}[x]$ are the nonzero constants, so p is a non-unit. If $p(x) = f(x)g(x)$ then $\deg(f) + \deg(g) = \deg(p) = 1$ so either f or g is constant. As neither f nor g can be zero and nonzero constants are invertible, we see one of f and g is a unit. This proves that p is prime.

Conversely, suppose that $q(x)$ is a prime in $\mathbb{C}[x]$. If $q = 0$ then the factorisation $q = x.0$ shows that q is not prime. If q is a nonzero constant then q is a unit and thus is not prime. Now suppose that $\deg(q) > 1$. The fundamental theorem of algebra then tells us that q has a root, say $q(\alpha) = 0$ for some $\alpha \in \mathbb{C}$. This means that $q(x)$ is divisible by $x - \alpha$, say $q(x) = r(x)(x - \alpha)$ for some polynomial $r(x)$. Clearly $\deg(r) = \deg(p) - 1 > 0$, so r is nonconstant and thus is not a unit. The factorisation $q(x) = r(x)(x - \alpha)$ now shows that q is not prime, contrary to hypothesis. Thus we must have $\deg(q) = 1$, so q is a nonconstant linear polynomial. □

In the ring $\mathbb{Z}$ we do not usually consider all prime elements but instead we concentrate on the positive ones. (If we did use all prime elements then we would not have unique factorisation, because for example $3.5 = (-3).(-5)$.) Similarly, in the ring $\mathbb{C}[x]$ we generally use only primes of the form $x-\alpha$. We do not need the prime $2x - 4$ (for example) because it is a unit multiple of the "standard" prime $x - 2$. These examples motivate the following definition.

Definition 52. Let R be a PID. A *standard set of primes* in R is a set $\mathcal{P}$ of prime elements such that

(a) For each prime element $p' \in R$, there is a prime $p \in \mathcal{P}$ that is a unit multiple of p'.
(b) If $p, q \in \mathcal{P}$ and $p \neq q$ then p is not a unit multiple of q.

In other words, $\mathcal{P}$ contains one element from each equivalence class of primes under the equivalence relation $\sim$ of Definition 35.

For any PID R we can certainly choose such a set $\mathcal{P}$. In the case $R = \mathbb{Z}$ we take $\mathcal{P}$ to be the set of positive primes, and in the case $R = \mathbb{C}[x]$ we take $\mathcal{P} = \{x - \alpha \mid \alpha \in \mathbb{C}\}$.

Lemma 53. *If p is prime and p does not divide a then p and a are coprime.*

Proof. Let d be a gcd of p and a, so $p = du$ and $a = dv$ for some $u, v \in R$. Because p is prime, either d or u must be a unit. If u is a unit then $a = p.(vu^{-1})$ so a is divisible by p, contrary to assumption. Thus d must be a unit, so a and p are coprime. □

Lemma 54. *If p is prime and p does not divide a or b then p does not divide ab.*

Proof. By the previous lemma we see that p is coprime to a and to b, so p is coprime to ab by Proposition 44, so p does not divide ab. □

8. Modules

Definition 55. Let M be an R-module and let $e_1, \ldots, e_n$ be elements of M. We say that:

(1) $e_1, \ldots, e_n$ are *independent* if the only way we can have $r_1e_1+\ldots+r_ne_n = 0$ is when $r_1 = \ldots = r_n = 0$.
(2) the elements $e_1, \ldots, e_n$ *generate* M if every element $x \in M$ can be written in the form $x = r_1e_1 + \ldots + r_ne_n$.
(3) the elements $e_1, \ldots, e_n$ form a *basis* for M if they are independent and they span M.
(4) M is *finitely generated* if there is some finite list $e_1, \ldots, e_n$ that generates M.
(5) M is *free of rank n* if there is a basis $e_1, \ldots, e_n$ for M.
(6) M is a *finitely generated free module* (or *FGF module*) if it is free of rank n for some n.

Example 56. In contrast to the situation for vector spaces, not every module has a basis. For example, consider the case where $R = \mathbb{Z}$ and $M = \mathbb{Z}/6$. For any element $x \in M$ we have $6x = 0$ so for any list $e_1, \ldots, e_n$ in M we have $6e_1 + \ldots + 6e_n = 0$. This relation shows that $e_1, \ldots, e_n$ is not independent and hence not a basis.

More generally, let M be a finite Abelian group, considered as a module over $R = \mathbb{Z}$. If $m = |M|$ then Lagrange's theorem tells us that every element $x \in M$ has order dividing m, so $mx = 0$. Thus for any list $e_1, \ldots, e_n$ we have $me_1 + \ldots + me_n = 0$, showing that $e_1, \ldots, e_n$ is not independent.

Remark 57. Here is a nice proof that $mx = 0$ avoiding Lagrange's theorem. Let $x_1, \ldots, x_m$ be the elements of M and put $y = x_1+\ldots+x_m$. The elements $x+x_1, \ldots, x+x_m$ are just the same as the elements $x_1, \ldots, x_m$ in a different order, so

$$(x + x_1) + \ldots + (x + x_m) = y.$$

On the other hand, we have

$$(x + x_1) + \ldots + (x + x_m) = (x + \ldots + x) + (x_1 + \ldots + x_m) = mx + y.$$

This means that $y = mx + y$ so $mx = 0$.

Example 58. Let M be $\mathbb{Z}[x]$ considered as an Abelian group (and thus a $\mathbb{Z}$-module) under addition. I claim that this is not finitely generated. To see this, let $f_1, \dots, f_n$ be any finite list of elements of M; we need to show that this list does not generate M. Let d be the maximum of the degrees of the polynomials f_i, and put $g(x) = x^{d+1}$. If $r_1, \dots, r_n$ are integers then it is clear that $r_1 f_1 + \dots + r_n f_n$ has degree at most d and thus is not equal to g. Thus g cannot be written in the form $r_1 f_1 + \dots + r_n f_n$, so $f_1, \dots, f_n$ does not generate.

Example 59. Let M be $\mathbb{Q}$ considered as an Abelian group (and thus a $\mathbb{Z}$-module) under addition. I claim that this is not finitely generated. To see this, let $e_1, \dots, e_n$ be any finite list of elements of M; we need to show that this list does not generate M. We can write $e_i = a_i/b_i$ with $a_i, b_i \in \mathbb{Z}$ and $b_i \neq 0$. Put $b = b_1 b_2 \dots b_n$ and note that the elements $e'_i := b e_i$ are integers. Put $x = 1/2b$. If $r_1, \dots, r_n$ are integers and $y = r_1 e_1 + \dots + r_n e_n$ then $by = r_1 e'_1 + \dots + r_n e'_n$ is an integer, but $bx = 1/2$ is not an integer, so $x \neq y$. This shows that x cannot be written in the form $r_1 e_1 + \dots + r_n e_n$, so $e_1, \dots, e_n$ does not generate M.

Proposition 60. *If M has a basis, then any two bases of M have the same number of elements.*

Proof. Let $e_1, \dots, e_n$ and $f_1, \dots, f_m$ be two bases of M. As the f's form a basis, we can write each e_i in terms of them, say

$$e_i = a_{i1} f_1 + \dots + a_{im} f_m = \sum_{j=1}^{m} a_{ij} f_j.$$

Similarly, we can write the f's in terms of the e's, say

$$f_j = b_{j1} e_1 + \dots + b_{jn} e_n = \sum_{k=1}^{n} b_{jk} e_k.$$

Let A be the matrix with entries a_{ij} and let B be the matrix with entries b_{jk}, so the matrix $C := AB$ has entries $c_{ik} := \sum_{j=1}^{m} a_{ij} b_{jk}$. We have

$$e_i = \sum_{j=1}^{m} a_{ij} f_j = \sum_{j=1}^{m} a_{ij} \sum_{k=1}^{n} b_{jk} e_k = \sum_{k=1}^{n} c_{ik} e_k.$$

By equating coefficients of the e's (which is valid because they form a basis) we see that $c_{ii} = 1$ and $c_{ik} = 0$ when $i \neq k$. This means that $C = AB$ is the identity matrix I_n. Similarly, we find that $BA = I_m$. We can regard A and B as matrices over the field of fractions QR, and we still have $AB = I_n$ and $BA = I_m$. It is a standard piece of linear algebra that when we have matrices over a field satisfying these conditions we must have $n = m$. Indeed, we may assume that $n \leq m$ (otherwise just switch everything around). Then any $n \times m$ or $m \times n$ matrix has rank at most n, so $\operatorname{rank}(A) \leq n$. Ranks always decrease under multiplication so $\operatorname{rank}(BA) \leq n$, but $BA = I_m$ so $\operatorname{rank}(BA) = m$ so $m \leq n$. As $n \leq m$ by assumption, we have $n = m$ as required.

For a more algorithmic proof, recall that by row and column reduction we can construct invertible matrices P and Q over QR such that the matrix $A' := PAQ$ is in normal form, or in other words

$$A' = \left(\begin{array}{c|c} I_r & 0_{(m-r)\times r} \\ \hline 0_{r\times(m-r)} & 0_{(m-r)\times(n-r)} \end{array}\right)$$

where r is the rank of A. Note that we could have $r = m$ in which case the two right hand blocks would have size zero and thus would not really be there. Similarly, we could have $r = n$ and then the bottom two blocks would have size zero. Define $B' = Q^{-1} A P^{-1}$, so

$$A'B' = PAQQ^{-1}BP^{-1} = PABP^{-1} = PI_nP^{-1} = I_n$$
$$B'A' = Q^{-1}BP^{-1}PAQ = Q^{-1}BAQ = Q^{-1}I_mQ = I_m.$$

We can partition B' into four blocks W, X, Y and Z of shapes rr, $(n-r)\times r$, $r\times(m-r)$ and $(n-r)\times(m-r)$. The equation $I_n = A'B'$ now becomes

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & I_{n-r} \end{array}\right) = \left(\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right) \left(\begin{array}{c|c} W & X \\ \hline Y & Z \end{array}\right) = \left(\begin{array}{c|c} W & X \\ \hline 0 & 0 \end{array}\right).$$

If the bottom right hand block is really there (ie it has nonzero size) then we find that $I_{n-r} = 0_{n-r}$ which cannot happen. Thus, this block must have size 0, in other words $n = r$. A similar argument using the equation $I_m = B'A'$ shows that $m = r$, so $n = m$. □

Definition 61. If M is an FGF module over R, then the *rank* of M is the number of elements in any basis of M.

Proposition 62. *Let F be a free module of rank n over R, and let M be a submodule of F. Then M is free of rank at most n.*

Proof. First, we know that F is isomorphic to R^n so we might as well assume that $F = R^n$. We will proceed by induction on n. The case $n = 0$ is clear: F is just the zero module, so M must be the zero module as well. For the case $n = 1$, recall that a submodule of R is just the same as an ideal, and all ideals are principal, so $M = Ra$ for some a. If $a = 0$ then M is free of rank 0. If $a \neq 0$ then we find that $\{a\}$ is a basis for M, so M is free of rank 1.

Suppose we have proved the proposition for free modules of rank $n-1$. Given a vector $x = (x_1, \ldots, x_n) \in R^n$, define $\pi(x) = x_n$. Put $G = \{x \in F \mid \pi(x) = 0\}$, so G is the set of vectors of the form $(x_1, \ldots, x_{n-1}, 0)$, which can be identified with R^{n-1}. Put $N = M \cap G$; this is a submodule of R^{n-1} and thus is free of rank at most $n-1$ by induction, so we can choose a basis $f_1, \ldots, f_r$ where $r \leq n-1$. Now put

$$I = \{a \in R \mid a = \pi(x) \text{ for some } x \in M\} = \pi(M).$$

I claim that this is an ideal. Indeed, if $a, b \in I$ and $c \in R$ then we can choose $x, y \in M$ with $\pi(x) = a$ and $\pi(y) = b$, so $x + y$ and cx lie in M and $a + b = \pi(x + y)$ and $cx = \pi(cx)$ so $a + b \in I$ and $cx \in I$. If $I = 0$ then $\pi(x) = 0$ for all $x \in M$, so $x \in G$ for all $x \in M$, so $M = N$, so M is free of rank $r < n$ as required. On the other hand, if $I \neq 0$ then $I = Ra$ for some $a \neq 0$. As $I = \pi(M)$ we can choose $f_{r+1} \in M$ with $\pi(f_{r+1}) = a$. I claim that $\{f_1, \ldots, f_{r+1}\}$ is a basis for M. We first prove that this set generates. Let x be an arbitrary element of M. Then $\pi(x) \in I = Ra$, so $\pi(x) = a_{r+1}a$ for some $a_{r+1} \in R$. Put $y = x - a_{r+1}f_{r+1}$ and note that $\pi(y) = \pi(x) - a_{r+1}\pi(f_{r+1}) = 0$, so $y \in G$. We also have $y \in M$ (because x and f_{r+1} lie in M) so $y \in M \cap G = N$. We know that $\{f_1, \ldots, f_r\}$ is a basis for N, so $y = a_1f_1 + \ldots + a_rf_r$ for some $a_1, \ldots, a_r$. As $y = x - a_{r+1}f_{r+1}$ this means that $x = a_1f_1 + \ldots + a_{r+1}f_{r+1}$, which shows that $\{f_1, \ldots, f_{r+1}\}$ is a generating set.

Finally, we need to show that the set $\{f_1, \ldots, f_{r+1}\}$ is independent. Suppose we have a relation $a_1f_1 + \ldots + a_{r+1}f_{r+1} = 0$. Then $\pi(a_1f_1 + \ldots + a_{r+1}f_{r+1}) = 0$, but for $i \leq r$ we have $f_i \in N$ so $\pi(f_i) = 0$, so this equation reduces to $\pi(a_{r+1}f_{r+1}) = 0$, or in other words $a_{r+1}a = 0$. As $a \neq 0$ we conclude that $a_{r+1} = 0$, so our relation only involves $\{f_1, \ldots, f_r\}$. By assumption these elements form a basis for N so they are independent, so $a_1 = \ldots = a_r = 0$ as required. □

Corollary 63. *If M is a finitely generated module and N is a submodule of M then N is also finitely generated.*

Proof. Choose a generating set $\{e_1, \ldots, e_n\}$ for M. For any element $x = (x_1, \ldots, x_n) \in R^m$, define

$$\phi(x) = x_1e_1 + \ldots + x_ne_n.$$

Put $L = \{x \in R^m \mid \phi(x) \in N\}$. If $x, y \in L$ then the element $\phi(x + y) = \phi(x) + \phi(y)$ lies in N, so $x + y \in L$. If $a \in R$ then $\phi(ax) = a\phi(x) \in N$, so $ax \in L$. Thus, L is a submodule of R^n. By the proposition, it is free of rank at most n, so we can find a basis $\{f_1, \ldots, f_m\}$ for L. Put $g_i = \phi(f_i) \in N$. I claim that the elements g_i generate N. To see this, suppose we have an element $y \in N$. Then in particular $y \in M$, and the e_i generate M so $y = x_1e_1 + \ldots + x_ne_n$ for some $x_1, \ldots, x_n$. In other words, we have $y = \phi(x)$ for some $x \in R^n$. As $\phi(x) = y$ lies in N by assumption, we have $x \in L$. As the elements f_i generate L, we have $x = \sum_i z_if_i$ for some $z_1, \ldots, z_m$. It follows that $y = \phi(x) = \sum_i z_i\phi(f_i) = \sum_i z_ig_i$, which lies in the span of the g_i as required. Thus, N is finitely generated. □

Definition 64. For any R-module M, an element $x \in M$ is a *torsion element* if there is some nonzero element $a \in R$ such that $ax = 0$. We write $\mathrm{tors}(M)$ for the set of all torsion elements. We say that M is *torsion-free* if $\mathrm{tors}(M) = \{0\}$. Equivalently, a module M is torsion-free if whenever $a \in R$ and $x \in M$ are both nonzero, their product ax is also nonzero.

Proposition 65. tors(M) *is a submodule of* M*.*

Proof. Suppose that $x, y \in \operatorname{tors}(M)$, so there exist nonzero elements a, b such that $ax = by = 0$. As R is a domain, the element ab is nonzero, and we have $ab(x + y) = b.(ax) + a.(by) = 0$, so $x + y \in \operatorname{tors}(M)$. Similarly, if $c \in R$ then $a.(cx) = c.(ax) = 0$, so $cx \in \operatorname{tors}(M)$. □

Proposition 66. *A finitely generated module is free if and only if it is torsion-free.*

Proof. First suppose that M is a free. Then $M \simeq R^n$ for some n, so we may assume that $M = R^n$. Suppose that $x = (x_1, \dots, x_n)$ is a torsion element. Then there is a nonzero element $a \in R$ such that $ax = 0$, so $(ax_1, \dots, ax_n) = (0, \dots, 0)$, so $ax_i = 0$ for all i. As R is a domain and $a \neq 0$ we must have $x_i = 0$ for all i, so $x = 0$. Thus M is torsion-free.

Conversely, suppose that M is torsion-free. Clearly, a set $\{e_1, \dots, e_n\}$ generates M iff $Re_1 + \dots + Re_n = M$. We will say that such a set *almost generates* M if there is a nonzero element $a \in R$ such that $Re_1 + \dots + Re_n \geq aM$. By assumption we can choose a finite list of elements that generates M, so certainly we can choose a finite list that almost generates M. Let $e_1, \dots, e_n$ be such a list which is as short as possible, and fix an element $a \neq 0$ such that $Re_1 + \dots + Re_n \geq aM$.

I claim that the elements $e_1, \dots, e_n$ are independent. If not, we have a relation $a_1e_1 + \dots + a_ne_n = 0$ where some coefficient a_k is nonzero. After reordering everything if necessary, we may assume that $a_n \neq 0$. For any $x \in M$ we know that ax can be written in the form $b_1e_1 + \dots + b_{n-1}e_{n-1} + b_ne_n$. After multiplying by a_n and using the substitution $a_nb_ne_n = -\sum_{i=1}^{n-1} a_ib_ne_i$, we see that a_nax lies in the span of $e_1, \dots, e_{n-1}$. Thus, $Re_1 + \dots + Re_{n-1} \geq a_naM$, so $\{e_1, \dots, e_{n-1}\}$ almost generates M, contradicting our assumption that the list $\{e_1, \dots, e_n\}$ was as short as possible. This contradiction shows that $\{e_1, \dots, e_n\}$ must be independent, after all. This implies that the module $N = Re_1 + \dots + Re_n$ is free. By assumption, the module $L := aM$ is contained in N, so it is free by Proposition 62, with basis $f_1, \dots, f_m$ say. As $L = aM$ we have $f_i = ag_i$ for some $g_i \in M$. If $\sum_i a_ig_i = 0$ then $\sum_i a_if_i = a\sum_i a_ig_i = 0$, and the f_i are independent so $a_1 = \dots = a_m = 0$. This shows that the g_i are independent. Moreover, if $x \in M$ then $ax \in L$ so $ax = \sum_i a_if_i$ for some $a_1, \dots, a_m$. This implies that

$$a(x - \sum_i a_ig_i) = ax - \sum_i a_if_i = 0,$$

but $a \neq 0$ and M is torsion-free so $x = \sum_i a_ig_i$. This shows that the elements g_i give a basis for M. □

9. Torsion modules

Definition 67. A module M is a *torsion module* if $M = \operatorname{tors}(M)$, or in other words for every $x \in M$ there exists $a \in R \setminus \{0\}$ such that $am = 0$. An *FGT module* is a finitely generated torsion module.

Our aim will be to give a complete classification of FGT modules up to isomorphism. We start by giving some basic building blocks. Throughout this section we fix a complete set of primes $\mathcal{P}$.

Definition 68. For any $p \in \mathcal{P}$ and $k \in \mathbb{N}$ we define B_p^k to be the cyclic module R/p^kR over R. A *basic* R-module is an R-module of the form B_p^k for some p and k. An R-module M is *standard* if it is isomorphic to the direct sum of a finite list of basic modules, with n_p^k copies of B_p^k, say.

Example 69. Here are some standard Abelian groups:

- $\mathbb{Z}/2 \oplus \mathbb{Z}/2 \oplus \mathbb{Z}/2$
- $\mathbb{Z}/3^{10} \oplus \mathbb{Z}/3^{20}$
- $\mathbb{Z}/2 \oplus \mathbb{Z}/3 \oplus \mathbb{Z}/5 \oplus \mathbb{Z}/7$.

The group $\mathbb{Z}/12$ is not obviously standard (because 12 is not a power of a prime). However, it will turn out that $\mathbb{Z}/12$ is isomorphic to $\mathbb{Z}/2^2 \oplus \mathbb{Z}/3$, so it is standard after all.

We will eventually show that *all* FGT modules are standard.

We would like to define the *multiplicity* of B_p^k in a standard module M to be the number of copies of B_p^k when we write M as a direct sum of these blocks. However, there is a possible ambiguity here: what if it happened that (say) $\mathbb{Z}/3 \oplus \mathbb{Z}/3 \oplus \mathbb{Z}/3$ was isomorphic to $\mathbb{Z}/3 \oplus \mathbb{Z}/9$; would the multiplicity of$\mathbb{Z}/3$ in this module be 1 or 2? Fortunately, this kind of thing never actually happens. To prove this, we need to introduce some numerical invariants.

Definition 70. Suppose that $p \in \mathcal{P}$ and $k \in \mathbb{N}$. For any FGT module M, we define

$$F_p^k(M) = \{x \in p^{k-1}M \mid px = 0\}.$$

This is easily seen to be a submodule of M. As M is finitely generated, we see from Corollary 63 that $F_p^k(M)$ is finitely generated. As $px = 0$ for all $x \in F_p^k(M)$, we can regard $F_p^k(M)$ as a module over R/p. As R/p is a field, every finitely generated module over it has a well-defined dimension, so we can define

$$f_p^k(M) = \dim_{R/p}(F_p^k(M)).$$

We also define

$$g_p^k(M) = f_p^k(M) - f_p^{k+1}(M).$$

Remark 71. It is easy to see that a pair $(x, y) \in M \oplus N$ lies in $F_p^k(M \oplus N)$ if and only if $x \in F_p^k(M)$ and $y \in F_p^k(N)$. It follows that $F_p^k(M \oplus N) = F_p^k(M) \oplus F_p^k(N)$ and thus that $f_p^k(M \oplus N) = f_p^k(M) + f_p^k(N)$ and $g_p^k(M \oplus N) = g_p^k(M) + g_p^k(N)$.

Remark 72. Suppose that $M \simeq M'$. I claim that $F_p^k(M) \simeq F_p^k(M')$, and thus that $f_p^k(M) = f_p^k(M')$ and $g_p^k(M) = g_p^k(M')$. Indeed, let $\phi\colon M \to M'$ be an isomorphism. Then if $x \in F_p^k(M)$ then $x = p^{k-1}y$ for some y and $px = 0$, so $\phi(x) = p^{k-1}\phi(y)$ and $p\phi(x) = 0$, so $\phi(x) \in F_p^k(M')$. Thus, ϕ gives a homomorphism from $F_p^k(M)$ to $F_p^k(M')$. Similarly, the homomorphism $\phi^{-1}\colon M' \to M$ restricts to give a homomorphism from $F_p^k(M')$ to $F_p^k(M)$. It is easy to see that these two maps are inverse to each other, so $F_p^k(M) \simeq F_p^k(M')$ as claimed.

Proposition 73. *We have*

$$g_p^k(B_q^j) = \begin{cases} 1 & \textit{if } p = q \textit{ and } k = j \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. We first prove that

$$f_p^k(B_q^j) = \begin{cases} 0 & \text{if } p \neq q \\ 0 & \text{if } p = q \text{ and } k > j \\ 1 & \text{if } p = q \text{ and } k \leq j. \end{cases}$$

First suppose that $p \neq q$. Then p^k and q^j are coprime, so $ap^k + bq^j = 1$ for some $a, b \in R$. If $x \in F_p^k(B_q^j)$ then $x = p^{k-1}y$ for some y and $px = 0$ so $p^k y = 0$. On the other hand, it is clear from the definition of B_q^j that $q^j z = 0$ for all $z \in B_q^j$, so $q^j y = 0$. We thus have $y = 1.y = ap^k y + bq^j y = 0$, and thus $x = p^{k-1}y = 0$. Thus $F_p^k(B_q^j) = 0$ and so $f_p^k(B_q^j) = 0$, as required.

Now suppose that $q = p$ and $j < k$. Then $k - 1 - j \geq 0$ and $p^{k-1}B_p^j = p^{k-1-j}p^j B_p^j = 0$ and $F_p^k(M) \leq p^{k-1}M$ so $F_p^k(B_p^j) = 0$. This means that $f_p^k(B_p^j) = 0$, as required.

Now suppose instead that $q = p$ and $k \leq j$. Let e be the element $1 + p^j R$ in B_p^j, so that $ae = (a + p^j R)$. Put $f = p^{j-1}e$, so that $f \neq 0$ and $pf = 0$. We also have $f = p^{k-1}(p^{j-k}e)$ so $f \in p^{k-1}B_p^j$, so $f \in F_p^k(B_p^j)$. Let u be another element of $F_p^k(B_p^j)$. We can write $u = ae = (a + p^j R)$ for some $a \in R$. As $pu = 0$ we have $pa = 0 \pmod{p^j}$, or in other words $pa = p^j b$ for some b, so $a = p^{j-1}b$ and thus $u = bf$. This shows that $\{f\}$ generates the vector space $F_p^k(B_p^j)$ over R/p, and $f \neq 0$ so the dimension must be exactly one. Thus $f_p^k(B_p^j) = 1$, as required.

It is now easy to deduce our description of $g_p^k(B_q^j)$. If $q \neq p$ then $f_p^k(B_q^j) = 0$ for all k and it follows easily that $g_p^k(B_q^j) = 0$. Suppose instead that $p = q$. If $k > j$ then $k + 1 > j$ as well so $f_p^k(B_p^j) = f_p^{k+1}(B_p^j) = 0$ so $g_p^k(B_p^j) = 0$ as claimed. If $k < j$ then both k and $k+1$ are less than or equal to j, so $f_p^k(B_p^j) = f_p^{k+1}(B_p^j) = 1$ so $g_p^k(B_p^j) = 0$ as claimed. If $k = j$ then $f_p^k(B_p^j) = 1$ and $f_p^{k+1}(B_p^j) = 0$ so $g_p^k(B_p^j) = 1$ as claimed. □

Corollary 74. *Let M be an FGT module. Then if there is any list of basic modules whose direct sum is isomorphic to M, then that list must contain precisely $g_p^k(M)$ copies of B_p^k.* □

We now start to prove that every FGT module is standard. The first step is to split each module M up into pieces $\mathrm{tors}_p(M)$ for each prime $p \in \mathcal{P}$. After that, we will split $\mathrm{tors}_p(M)$ up into pieces of the form B_p^k for various k.

Definition 75. Let p be a prime in $\mathcal{P}$. We say that an element $x \in M$ is a *p-torsion element* if $p^k x = 0$ for some $k \geq 0$. We write $\text{tors}_p(M)$ for the set of p-torsion elements, which is easily seen to be a submodule of M. We say that M is a p-torsion module if $\text{tors}_p(M) = M$, and that M is an FGT_pmodule if it is finitely generated and p-torsion.

Proposition 76. *If M is an FGT module, then there is a nonzero element $a \in R$ such that $aM = 0$. If M is an FGT_pmodule, then there exists $k \geq 0$ such that $p^k M = 0$.*

Proof. Choose a finite generating set $\{e_1, \ldots, e_n\}$ for M. As M is a torsion module, for each i we can choose $a_i \neq 0$ such that $a_i e_i = 0$. Put $a = \prod_i a_i$, so $ae_i = (\prod_{j \neq i} a_j)(a_i e_i) = 0$ for all i. As the elements e_i generate M, we deduce that $aM = 0$. In the p-torsion case we can choose each a_i to have the form p^{k_i} and then $a = p^k$ where $k = k_1 + \ldots + k_n$. More efficiently, we can put $l = \max(k_1, \ldots, k_n)$; clearly $p^l e_i = 0$ for all i and thus $p^l M = 0$. □

Proposition 77. *Let L be an R-module. Suppose that $b, c \in R$ are coprime and that $bcL = 0$. Put $M = \{y \in L \mid by = 0\}$ and $N = \{z \in L \mid cz = 0\}$. Then $L = M \oplus N$.*

Proof. As b and c are coprime there exist elements u, w such that $ub + wc = 1$. If $x \in M \cap N$ then $bx = cx = 0$ so $x = 1.x = ubx + wcx = 0$; thus $M \cap N = 0$. Now let x be an arbitrary element of M. Put $y = wcx$ and $z = ubx$, so $x = y + z$. We have $by = (bc)(wx)$ but $bcL = 0$ so $by = 0$ so $y \in M$. Similarly, $cz = (bc)(ux) = 0$ so $z \in N$. Thus $x = y + z \in M + N$, which shows that $M + N = L$. As $M \cap N = 0$, the sum is direct. □

Proposition 78. *Let M be an FGT module. Then there is a finite list of standard primes $p_1, \ldots, p_k$ (called the* support *of M) such that*

$$M = \text{tors}_{p_1}(M) \oplus \ldots \oplus \text{tors}_{p_k}(M),$$

and $\text{tors}_q(M) = 0$ *for all primes other than $p_1, \ldots, p_k$.*

Proof. Put $I = \{a \in R \mid aM = 0\}$. This is easily seen to be an ideal, and it is nonzero by Proposition 76. We thus have $I = Ra$ for some $a \neq 0$. We can factor a as $a = u \prod_{i=1}^k p_i^{v_i}$ as in Proposition **??**, and as u is a unit we have $I = R \prod_i p_i^{v_i}$.

Suppose that $m \in \text{tors}_{p_i}(M)$; I claim that $p_i^{v_i} m = 0$. Indeed, by the definition of $\text{tors}_{p_i}(M)$ we certainly have $p_i^w m = 0$ for some w. If $w \leq v_i$ then it follows easily that $p_i^{v_i} m = p_i^{v_i - w} p^w m = 0$ as required. If $w \geq v_i$ we note that $p_i^{v_i}$ is a gcd of p^w and a, so $p_i^{v_i} \in Ra + Rp^w = I + Rp^w$. We know that $Im = 0$ and $Rp^w m = 0$ so $(I + Rp^w)m = 0$ so $p_i^{v_i} m = 0$ as required.

Next, by the Chinese remainder theorem we can choose $e_1, \ldots, e_k \in R$ such that $e_i = 1 \pmod{p_i^{v_i}}$ and $e_i = 0 \pmod{p_j^{v_j}}$ for $j \neq i$ and $e_1 + \ldots + e_k = 1$.

Now suppose we have $m \in M$. Put $m_i = e_i m$ for $i = 1, \ldots, k$. I claim that $p_i^{v_i} m_i = 0$, so that $m_i \in \text{tors}_{p_i}(M)$. Indeed, e_i is divisible by $p_j^{v_j}$ for all $j \neq i$, so $p_i^{v_i} e_i$ is divisible by $p_j^{v_j}$ for all j including $j = i$, so $p_i^{v_i} e_i$ is divisible by a. As $am = 0$, it follows that $p_i^{v_i} e_i m = 0$, or in other words $p_i^{v_i} m_i = 0$, as claimed.

We also have $m = 1.m = (e_1 + \ldots + e_k)m = m_1 + \ldots + m_k$, and it follows that $M = \text{tors}_{p_1}(M) + \ldots + \text{tors}_{p_k}(M)$.

We next check that this sum is a direct sum. Suppose we have elements $n_1, \ldots, n_k$ with $n_i \in \text{tors}_{p_i}(M)$ and $n_1 + \ldots + n_k = 0$; we must show that $n_i = 0$ for all i. As $n_i \in \text{tors}_{p_i}(M)$ we have $p_i^{v_i} n_i = 0$, and $e_i = 1 \pmod{p_i^{v_i}}$ so $e_i n_i = n_i$. For $j \neq i$ we have $p_j^{v_j} n_j = 0$ and e_i is divisible by $p_j^{n_j}$ so $e_i n_j = 0$. Thus $n_i = e_i(n_1 + \ldots + n_k) = e_i.0 = 0$, as required. Thus $M = \text{tors}_{p_1}(M) \oplus \ldots \oplus \text{tors}_{p_k}(M)$.

Finally, let q be a prime other than $p_1, \ldots, p_k$. If $m \in \text{tors}_q(M)$ then $q^w m = 0$ for some w. This means that $(Ra + Rq^w)m = 0$, but q^w is coprime to a so $1 \in Ra + Rq^w$ so $m = 1.m = 0$. Thus $\text{tors}_q(M) = 0$ as claimed. □