

## 620-321 Algebra

Semester 1, 2003

### Hyam Rubinstein's Summary Notes for Algebra 620-321

#### CHAPTER ONE - Basic ring theory

We begin with basic topics in Ring theory - with results exactly like those in group theory as in Linear and Abstract algebra 222.

**Definition:** A ring  $(R, +, \cdot)$  is a set of elements  $R$  together with two binary operations  $+$  and  $\cdot$ , ie methods for combining two elements of  $R$  to form a third element.  $(R, +)$  is an Abelian group and there are two axioms involving multiplication.

- For elements  $x, y, z \in R$ ,  $x \cdot (y \cdot z) = (x \cdot y) \cdot z$  - *Associativity*
- For elements  $x, y, z \in R$ ,  $x \cdot (y + z) = x \cdot y + x \cdot z$   
and  $(y + z) \cdot x = y \cdot x + z \cdot x$  - *Distributivity*

**Definition:** A ring  $(R, +, \cdot)$  is commutative if  $x \cdot y = y \cdot x$  for all  $x, y \in R$ .

**Definition:** An element  $1$  is called a unit element for a ring  $(R, +, \cdot)$  if  $x \cdot 1 = 1 \cdot x = x$  for all  $x \in R$ .

Most of the rings we will consider in this course will be commutative with a unit element.

**Example 1** *The integers  $\mathbb{Z}$ , rationals  $\mathbb{Q}$ , reals  $\mathbb{R}$ , complex numbers  $\mathbb{C}$  and quaternions  $\mathbb{H}$  are all rings with usual addition and multiplication and all have units. They are also all commutative except for the quaternions.*

**Example 2** *Polynomials form an important class of rings; the coefficients of the polynomials can be in any fixed ring  $(R, +, \cdot)$ . So such a ring is denoted  $R[x]$ , where  $x$  is the variable. For instance,  $R$  could be any of  $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{H}$ .*

**Example 3** *The residue or congruence classes modulo  $n$ , where  $n$  is an integer  $\geq 2$ , form a ring denoted  $\mathbb{Z}_n$ . The elements of  $\mathbb{Z}_n$  are the classes  $\bar{0}, \bar{1}, \dots, \overline{n-1}$ , where  $\bar{0} = \{.. - 2n, -n, 0, n, 2n, ..\}$ ,  $\bar{1} = \{.. - 2n + 1, -n + 1, 1, n + 1, 2n + 1, ..\}$  etc. Addition of these classes is defined by  $\bar{i} + \bar{j} = \overline{i + j}$ , where we have to reduce  $i + j$  modulo  $n$ , ie replace it by  $i + j - n$  if  $i + j \geq n$ . Similarly multiplication is defined by  $\bar{i} \cdot \bar{j} = \overline{i \cdot j}$  where reduction modulo  $n$  takes place as necessary.*

**Example 4** *The final example is a 'mixed one', namely  $R = \mathbb{Z}_6[\sqrt{3}]$ . Notice that since  $(\sqrt{3})^2 = 3$ , powers of  $\sqrt{3}$  are either integers or integer multiples of  $\sqrt{3}$ . So this ring consists of the **constant and linear** polynomials only. So there are precisely 36 elements of the ring.*

Consider next special elements of rings.

**Definition:** A zero divisor in a ring  $R$  is a non zero element  $x$  so that there is a non zero element  $y$  with  $x \cdot y = 0$ . (Sometimes if the ring is not commutative this might be called a left zero divisor so that  $y$  is a right zero divisor. Here as usually  $R$  is commutative, we will not worry about this point).

**Definition:** A unit in a ring  $R$  with unit element 1, is an element  $x$  so that there is an element  $y$  and  $x \cdot y = y \cdot x = 1$ . Notice of course that 1 is a unit in this sense.  $y$  is called the inverse of  $x$  and can be denoted  $x^{-1}$ .

**Example 5** In the ring  $\mathbb{Z}_6$ , the elements  $\bar{1}$  and  $\bar{5}$  are units and are their own inverses. The elements  $\bar{2}, \bar{3}, \bar{4}$  are all zero divisors since  $\bar{2} \cdot \bar{3} = \bar{3} \cdot \bar{4} = \bar{0}$ .

Note that units and zero divisors are mutually exclusive, ie no element can be both, since if  $x \cdot y = 0$  and  $x$  has an inverse  $z$ , multiplying through by  $z$  shows that  $y = 0$ .  $R$  is called an *integral domain* if  $R$  is commutative, with an identity element and there are no zero divisors. Moreover  $R$  is a *field*, if  $R$  is commutative with an identity element and every non zero element of  $R$  is a unit.

**Example 6** Let  $M_n(F)$  denote the ring of  $n \times n$  matrices with entries in a field  $F$ . Then every non zero element is either a zero divisor or is a unit. The reason is that a unit is an invertible matrix. If  $A$  is not a unit, then the columns  $C_1, \dots, C_n$  of  $A$  are linearly independent. So we can find elements  $r_1, \dots, r_n$  of  $F$  so that  $\sum r_i C_i = 0$ . But then the matrix  $B$  defined by taking  $(r_1, \dots, r_n)$  for all  $n$  columns of  $B$ , satisfies,  $AB = 0$ . Hence  $A$  is a zero divisor.

**Lemma 1** In a ring  $R$ ,  $x \cdot 0 = 0$ .

**Proof:** Expand out  $(0 + 0) \cdot x = 0 \cdot x$  and cancel out one of the terms  $0 \cdot x$  from both sides of the equation by adding  $-0 \cdot x$ . ■

**Lemma 2** In an integral domain, cancellation can be performed, ie if  $x \cdot y = x \cdot z$  and  $x \neq 0$ , then  $y = z$ .

**Proof:** By the distributive law, we can factorise to get  $x \cdot (y - z) = 0$ . Since  $R$  is an integral domain, there are no zero divisors and since  $x \neq 0$ , it follows that  $y - z = 0$ . ■

**Corollary 1** In a ring, every unit has a unique inverse.

**Proof:** Suppose that  $x \cdot y = x \cdot z = y \cdot x = z \cdot x = 1$ , so that  $y, z$  are both inverse to  $x$ . Then  $z \cdot (x \cdot y) = z \cdot 1 = z$  and also  $z \cdot (x \cdot y) = (z \cdot x) \cdot y = 1 \cdot y = y$ . So the inverse is unique, if it exists. ■

**Definition:** A ring homomorphism is a map between rings  $\phi : R_1 \rightarrow R_2$  which preserves addition and multiplication, ie

- $\phi(x + y) = \phi(x) + \phi(y)$  and  $\phi(x \cdot y) = \phi(x) \cdot \phi(y)$ .

**Remark**

Notice that for a ring homomorphism, we get both  $\phi(0) = 0$  and  $\phi(-x) = -\phi(x)$ , for any  $x \in R_1$ . For the first, note that  $\phi(0 + 0) = \phi(0) + \phi(0)$  and so  $\phi(0) = 2\phi(0)$ . Cancelling one  $\phi(0)$  gives the result. Next,  $\phi(x + (-x)) = \phi(x) + \phi(-x) = \phi(0) = 0$ . So subtracting  $\phi(x)$  from both sides gives the second result.

**Definition:** A subring  $S$  of a ring  $R$  is a non empty subset which is itself a ring under the same addition, subtraction and multiplication as for  $R$ . It is easy to see this is the same as saying that  $S$  is closed under addition, subtraction and multiplication, ie  $x, y \in S$  implies  $x + y \in S$ ,  $x - y \in S$  and  $x \cdot y \in S$ .

It is easy to see that the kernel  $K$  and the image  $L$  of a ring homomorphism  $\phi$  are both subrings of  $R_1$  and  $R_2$  respectively. Recall that  $K = \{x : \phi(x) = 0\}$  and  $L = \{y : y = \phi(x)\}$ . Note also that  $\phi(0) = 0$  is easy to prove, expanding the expression for  $\phi(0 + 0)$ .

**Definition:** An ideal  $I$  in a ring  $R$  is a subring with the additional property that;

- If  $x \in I, r \in R$  then  $r \cdot x \in I$  and  $x \cdot r \in I$ .

**Remark**

Notice that an ideal is therefore a non empty subset of a ring which is closed under addition, subtraction and arbitrary multiplication of an element of the subset by any element of the ring.

It is very easy to check;

**Lemma 3** *The kernel of a ring homomorphism is always an ideal  $I$ . Also a ring homomorphism is one-to-one or an injection, if and only if it has kernel  $\{0\}$ .*

**Proof:** If  $\phi$  is a ring homomorphism and  $a \in I$ , then  $\phi(r \cdot a) = \phi(r) \cdot \phi(a) = \phi(r) \cdot 0 = 0$ . So  $r \cdot a \in I$ , whenever  $a \in I$ , ie  $I$  is an ideal. Moreover if  $I = \{0\}$ , then  $\phi(a) = \phi(b)$  implies  $\phi(a - b) = 0$  and so  $a - b \in I$ . Therefore  $a - b = 0$  and  $a = b$ . This shows that  $\phi$  is one-to-one. The converse is similar. ■

**Example 7**  $\phi : \mathbb{Z} \rightarrow \mathbb{Z}_n$ , given by  $\phi(i) = \bar{i}$ , is a ring epimorphism, or an onto map, for every  $n \geq 2$ . So the image of  $\phi$  is  $\mathbb{Z}_n$ . The kernel of  $\phi$  is the ideal  $n\mathbb{Z}$ , ie the set of multiples of  $n$ . Another example is  $\psi : \mathbb{Z} \rightarrow \mathbb{Z}$  defined by  $\psi(i) = mi$  for some fixed positive integer  $m$ . Here  $\psi$  is a monomorphism, or injection, or one-to-one, has kernel consisting of  $\{0\}$  and image is  $m\mathbb{Z}$ .

**Example 8** An important example is the evaluation map  $\alpha$  from a polynomial ring  $R[x]$  to the coefficient ring  $R$ . Choose any fixed element  $r_0$  of  $R$  and map any polynomial  $p(x)$  to the value  $p(r_0)$ . Notice we had better assume that  $R$  is commutative, or we have difficulties defining  $p(r_0)$ !. Notice that the kernel  $K$  is the ideal of polynomials which have  $r_0$  as a root. In the easiest case, that  $R = \mathbb{C}$ , then this is exactly the polynomials which have  $x - r_0$  as a factor. For general rings, it is a more difficult problem to determine  $K$  precisely.

Next we show the converse, that every ideal is the kernel of some homomorphism. This is exactly like the construction of residue or congruence classes modulo  $n$  in the integers  $\mathbb{Z}$  and the ring  $\mathbb{Z}_n$  of these. Let  $I$  be an ideal in a ring  $R$ .

**Definition:** The quotient ring, denoted  $R/I$  or  $\frac{R}{I}$  is defined as follows;

- $R/I$  is the collection of cosets of the ideal  $I$ . A coset is a set of the form  $x + I$  ie all elements  $x + i; i \in I$ , where  $x$  is fixed. We can also denote this set by  $[x]$ . Now it is easy to show that  $[x] = [y]$  if and only if  $x - y \in I$ . Otherwise the intersection  $[x] \cap [y] = \emptyset$ . So the cosets form a *partition* of  $R$ , ie divide  $R$  into disjoint sets. Addition, subtraction and multiplication of cosets are defined by  $[x] + [y] = [x + y]$ ,  $[x] - [y] = [x - y]$  and  $[x] \cdot [y] = [x \cdot y]$ . The natural projection  $\pi : R \rightarrow R/I$  is a ring epimorphism with kernel  $I$ .

**Theorem 1 The first ring isomorphism theorem.** If  $\phi : R_1 \rightarrow R_2$  is a ring homomorphism and  $I$  is the kernel and  $L$  the image of  $\phi$ , then  $R_1/I \cong L$ , where  $\cong$  is a ring isomorphism.

**Proof:** We define a map  $\psi : R_1/I \rightarrow R_2$  and show that  $\psi$  has kernel  $\{[0]\}$  and image  $L$ . Hence  $\psi$  is an isomorphism between  $R_1/I$  and  $L$ .  $\psi$  acts on a coset  $x + I$  by

$$\psi(x + I) = \phi(x)$$

. This is well defined, since  $x + I = x' + I$  exactly when  $x - x' \in I$  and then  $\phi(x - x') = 0$ . So  $\phi(x) = \phi(x')$ .

If  $x + I$  is in the kernel of  $\psi$ , then  $\psi(x + I) = \phi(x) = 0$ . But this implies that  $x \in I$  and so  $x + I = I$ . This shows the kernel of  $\psi$  is just the zero element of the quotient ring  $R_1/I$  which is the zero coset  $[0] = I = 0 + I$ . (Note the confusing notation - in the quotient ring we have three different ways of writing the zero element).

Next, the image of  $\psi$  is obviously the same as the image of  $\phi$  by definition. So this image is  $L$  and the result is proved. ■

We want to look at how subrings and ideals correspond under a ring homomorphism.

**Theorem 2 The third ring isomorphism theorem.** Assume  $\phi : R_1 \rightarrow R_2$  is a ring homomorphism and  $I$  is the kernel and  $L$  the image of  $\phi$ . If  $J_2$  is such a subring (respectively ideal) of  $L$ , then we can define a subring (respectively ideal)  $J_1$  of  $R_1$  containing  $I$  by  $J_1 = \phi^{-1}(J_2)$ . This is a bijection and moreover,  $J_1/I \cong J_2$ .

**Proof:** By definition,  $J_1$  consists of all elements of  $R_1$  mapped to  $J_2$  by  $\phi$ . It is easy to check that  $J_1$  is a subring (respectively ideal) of  $R_1$  since this is true for  $J_2$  in  $L$ . Also, if two different subrings (respectively ideals) of  $L$  are chosen, then the two subrings (respectively ideals) of  $R_1$  found this way will be different. This follows, since  $\phi(J_1) = J_2$ . Any subring (respectively ideal)  $J_1$  of  $R_1$  containing  $I$  is obtained by this construction, since we can take  $\phi(J_1)$  as the definition of  $J_2$ . Next clearly  $J_1 \supseteq I$ , since any element  $x \in I$  is mapped by  $\phi$  to  $0 \in J_2$ . So by definition  $x \in \phi^{-1}J_2 = J_1$ . So this shows that there is a bijection between subrings (respectively ideals) as claimed.

Finally, by the first isomorphism theorem applied to the restriction of  $\phi$  to the subring (respectively ideal)  $J_1$ , we get a ring homomorphism from  $J_1 \rightarrow J_2$  which is onto, with kernel  $I$  and so  $J_1/I \cong J_2$ . ■

**Example 9** Consider  $\phi : \mathbb{Z} \rightarrow \mathbb{Z}_6$  given by  $\phi(i) = \bar{i}$ . As is easy to check, the subrings of  $\mathbb{Z}_6$  are all ideals and there are four of them. These are  $\{\bar{0}\}$ ,  $\mathbb{Z}_6$ ,  $\{\bar{0}, \bar{3}\}$  and  $\{\bar{0}, \bar{2}, \bar{4}\}$ . The corresponding ideal of  $\mathbb{Z}$  which contain the kernel of  $\phi$  which is  $I = 6\mathbb{Z}$  are  $I$ ,  $\mathbb{Z}$ ,  $3\mathbb{Z}$  and  $2\mathbb{Z}$ . Note the third isomorphism theorem also gives results like  $2\mathbb{Z}/6\mathbb{Z} \cong \{\bar{0}, \bar{2}, \bar{4}\}$ . The latter ring is not isomorphic to  $\mathbb{Z}_3$ , since it has no multiplicative identity element!

Another useful result about ideals is the following;

**Theorem 3** Suppose that  $R$  is a ring and  $J, K$  are ideals with  $J \subseteq K$ . Then  $J$  (respectively  $K/J$ ) is an ideal in  $K$  (respectively  $R/J$ ) thought of as a ring and  $(R/J)/(K/J) \cong R/K$ .

**Proof:** To check that  $J$  (respectively  $K/J$ ) is an ideal in  $K$  (respectively  $R/J$ ) is easy - it follows directly from the closure properties. We need only prove the isomorphism result which will follow from the first isomorphism theorem. Notice that there is an induced homomorphism  $\phi : R/J \rightarrow R/K$  defined by  $\phi(x + J) = x + K$ . It is easy to show this is well defined. So it remains to observe that  $\phi$  is onto with kernel  $K/J$ , since the cosets of the form  $k + J$  for  $k \in K$  are exactly those elements in the kernel. Hence we get the required isomorphism. ■

**Definition:** The subring or ideal generated by a non empty set of elements  $S$  in a ring  $R$  is the smallest subring or ideal containing  $S$ . We denote this subring or ideal by  $\langle S \rangle$ . Note that by taking the intersection of all subrings or ideals including  $S$ , we can abstractly form  $\langle S \rangle$ . However we would like a practical way of constructing  $\langle S \rangle$ .

In case,  $S$  is a single element, ie  $S = \{x\}$ , and  $R$  is a commutative ring with a unit, we can easily find  $\langle S \rangle$ . To find the smallest subring containing  $x$ , notice that all polynomials with integer coefficients and ‘variable’  $x$  and no constant terms must be in this subring. On the other hand, this is a subring, so we have shown that  $\langle S \rangle$  is the polynomials with zero constant terms in  $\mathbb{Z}[x]$ , ie expressions like  $3x + 15x^2 - 4x^4$ . These polynomials all have  $x$  as a factor, ie lie in  $x\mathbb{Z}[x]$ . Hence  $\langle S \rangle = x\mathbb{Z}[x]$ .

To find the ideal generated by  $x$ , we must have all polynomials with variable  $x$  and arbitrary coefficients in the ring  $R$ , with zero constant term ie polynomials in  $xR[x]$ .

However this set can be considerably simplified. Notice that a polynomial of this form can be written as  $xr = rx$  for  $r \in R$ . On the other hand, every element of this form is in the ideal. Hence the ideal generated by  $x$  is  $xR$ . We call this a *principal ideal*. Principal ideals play a similar role in ring theory to cyclic subgroups in group theory.

One can also explicitly construct subrings and ideals generated by finitely many elements in a ring.

**Example 10** Consider the ideal generated by elements  $2, x$  in the ring  $\mathbb{Z}[x]$ . This is all polynomials in  $x$  with integer coefficients and even constant term, ie the set  $J$  of polynomials of the form  $a_0 + a_1x + \dots + a_kx^k$ , where  $a_i$  are integers and  $a_0$  is even. To see this, it is very straightforward to check  $J$  is an ideal. So the main point is why is this the smallest ideal containing  $2, x$ . Notice that as before, the ideal generated by  $x$  is  $x\mathbb{Z}[x]$ , ie all polynomials divisible by  $x$ , which have zero constant term. To such polynomials we can always add a multiple of  $2$ , ie any even constant term. So this shows that the ideal generated by  $2, x$  contains  $J$ . So these sets are equal. Note this ideal is **not** principal. For if  $J$  is all multiples of some fixed polynomial, then this polynomial would have to divide  $2$ . So the polynomial would have to be equal to  $2$  and then  $x$  would not be in  $J$ .

In the next chapter, we will be interested in rings with the property that every ideal is principal. Included in this category are all rings of the form  $F[x]$  of polynomials with coefficients in a field  $F$ . So this example shows that using a ring rather than a field destroys this characteristic of having all ideals principal.

*Last modified: 11 AM 26 March*

## CHAPTER TWO - Factorisation in rings, part I

A key idea in the study of rings is to imitate the factorisation of integers into primes. So we define the analogue of a prime element in a ring. Moreover, uniqueness of such factorisations is also crucial. It will turn out there is an interesting connection to the structure of the ideals in the ring. The Euclidean algorithm will also be extremely useful, as it underlies the factorisation of integers. We will find that since polynomials also exhibit a similar Euclidean algorithm, that they can be factorised uniquely.

In  $\mathbb{Z}$ , every positive integer  $a$  can be written as a product of prime numbers, ie  $a = p_1 p_2 \dots p_k$  where each  $p_i$  is prime. Moreover, if  $a = q_1 q_2 \dots q_l$  is another factorisation of  $a$  into primes, then  $k = l$  and the sets of primes  $\{p_1, \dots, p_k\}$  and  $\{q_1, \dots, q_l\}$  coincide. Alternatively we can say the primes are the same after reordering. Here a prime is an integer  $> 1$  which has only 1 and itself as positive integer divisors.

If we are working in the whole ring  $\mathbb{Z}$ , it makes sense to extend this idea to negative integers as well. In this case, we allow prime numbers  $p$  to be positive or negative, not equal to 0 or  $\pm 1$  and to have only  $\pm 1$  and  $\pm p$  as divisors.

We can now write any integer  $a$  as a product of prime numbers  $a = p_1 p_2 \dots p_k$ , so that if also  $a = q_1 q_2 \dots q_l$  is another factorisation of  $a$  into primes, then  $k = l$  and the primes in the two factorisations are the same after reordering and multiplication by  $\pm 1$ .

How is this proved rigorously? Firstly, we need to show a factorisation exists. Now either  $a$  is already a prime, or there is a divisor  $b$  which is different from  $\pm 1$  and  $\pm a$ . Now we get  $a = bc$  and clearly  $1 < |b| < |a|$ , with the same true for  $c$ . The argument can be continued, either  $b$  (and  $c$ ) is a prime, or there is a further factorisation. Since at each stage the absolute value of the factors decreases, the process must stop at primes.

Next we want to show that if  $a = p_1 p_2 \dots p_k = q_1 q_2 \dots q_l$  are two factorisations into primes, then the primes are the same after reordering and multiplication by  $\pm 1$ . A key property of integers is (\*) that if  $p|xy$  and  $p$  is a prime, then either  $p|x$  or  $p|y$ , where  $u|v$  means  $u$  divides  $v$ . We will consider how to prove property (\*) later - to extend the whole argument to other rings, this is a key step.

Now by property (\*), we can prove by induction that if  $p$  is a prime and divides a finite product, then it divides one of the factors. Hence applying this to the two factorisations of  $a$ , we see that  $p_1$  divides  $q_1 q_2 \dots q_l$  and so  $p_1|q_i$  for some  $i$ . But  $q_i$  is a prime and so a factor is  $\pm 1$  or  $\pm q_i$ . Hence since  $p_1$  is a prime it cannot be  $\pm 1$  and we see that  $p_1 = \pm q_i$ . This factor can now be cancelled from the two expressions and the  $\pm 1$  can be added to one of the prime factors. We can therefore assume that we have two different expressions  $p_2 p_3 \dots p_k = q_1 \dots q_{i-1} q_{i+1} \dots q_l$ . By induction on the length of the factorisations, we can assume now the remaining primes are the same after reordering and multiplication by  $\pm 1$ . This completes the argument.

Finally we need to establish (\*). Here the Euclidean algorithm is the key technique. Given two positive integers  $a, b$ , we can divide  $b$  by  $a$  and get a remainder  $r$  with  $0 \leq r < a$ . So  $b = qa + r$ . Now the Euclidean algorithm is to keep doing this until we end up with the greatest common divisor of  $a, b$ . So we next divide  $a$  by  $r$ . To make the notation better, let's write  $a = a_1$  and  $r = a_2$ . Then we get a sequence of equations of the following form:

$$b = q_1 a_1 + a_2, \quad 0 \leq a_2 < a_1$$

$$a_1 = q_2 a_2 + a_3, \quad 0 \leq a_3 < a_2$$

.....

$$a_{k-1} = q_k a_k, \quad 0 < a_k < a_{k-1}$$

Note there is no remainder in the last equation. Since the sequence  $a_1, a_2, \dots, a_k$  is strictly decreasing, it must terminate when such a zero remainder occurs.

Next, observe that  $a_k$  must be the greatest common divisor of  $a = a_1, b$ . For it is easy to see that  $a_k$  divides  $a_{k-1}, a_{k-2} \dots a_1, b$  using the equations above and so certainly divides both  $a, b$ . So it is a common divisor. On the other hand, if  $c$  divides both  $a, b$  it divides  $a_2$  and so  $a_3$  etc ie all the entries  $a_1, a_2, \dots, a_k$  and so is not larger than  $a_k$ . So  $a_k$  is the g.c.d as claimed.

Finally note that we get an equation of the form

$$a_k = xa + yb$$

by substituting in the above equations. To illustrate this, assume that  $k = 4$ , so the equations become  $b = q_1 a_1 + a_2$ ,  $a_1 = q_2 a_2 + a_3$  and  $a_2 = q_3 a_3 + a_4$ , with  $a_3 = q_4 a_4$ . Now we can substitute for  $a_3$  using the second equation, into the third one. This gives  $a_2 = q_3(a_1 - q_2 a_2) + a_4$  which can be rewritten as  $-q_3 a_1 + (1 + q_2 q_3) a_2 = a_4$ .

Next we will substitute for  $a_2$  in this equation using the first equation and will replace  $a_1 = a$ . Then  $-q_3 a_1 + (1 + q_2 q_3)(b - q_1 a_1) = a_4$  ie  $(-q_1 - q_3 - q_1 q_2 q_3)a + (1 + q_2 q_3)b = a_4$ . This is an equation of the form  $a_k = xa + yb$ .

Now we use this to deduce property (\*). Suppose a prime  $p$  divides  $ab$  and  $p$  does not divide  $a$ . (We may assume that all three numbers are positive without loss of generality by changing signs). Then the g.c.d of  $\{p, a\}$  must be 1, since the only positive divisors of  $p$  are 1 and  $p$ . Hence by the Euclidean algorithm, we can find integers  $x, y$  so that  $px + ay = 1$ . Multiply this equation by  $b$  to get  $pbx + aby = b$ . Now by assumption,  $p$  divides both terms on the right side of the equation. Hence  $p|b$  and property (\*) is proved.

Recall that an integral domain is a commutative ring with an identity element 1 for which there are no zero divisors. Throughout this chapter,  $R$  will always be an integral domain.

**Definition:**

- $a|b$ , ie  $a$  divides  $b$  if there is an element  $c$  with  $b = ac$ .
- $a$  is a unit if  $a|1$ , ie there is an element  $c$  with  $1 = ac$ .
- $a$  is irreducible if  $a$  is not a unit or 0 and whenever  $a = bc$ , then either  $b$  or  $c$  is a unit.
- $a$  and  $b$  are associates if  $a|b, b|a$  and neither  $a$  nor  $b$  is equal to 0 or a unit. In this case,  $b = ua$  and  $a = vb$  so that  $b = uvb$ . Since  $R$  is an integral domain, there are no zero divisors and so  $b(1 - uv) = 0$  with  $b \neq 0$  implies that  $1 - uv = 0$ . So  $u, v$  are units and associates differ by multiplication by units.
- $a$  is a proper divisor of  $b$  if  $a$  is not a unit or 0,  $a|b$  and  $b \neq 0$  but  $a, b$  are not associates, ie  $b$  does not divide  $a$ .

**Example 11** Let  $R = F[x]$  be the ring of polynomials with coefficients in a field  $F$ . A unit in this ring is a constant non zero polynomial. For certainly such a polynomial has an inverse and any non constant polynomial has no inverse, since the product of polynomials has degree the sum of the degrees of the two polynomials (see later!). So  $p(x)q(x) = 1$  implies both  $p(x), q(x)$  have degree 0, ie are constant. Hence associates are polynomials which differ by a constant factor, since clearly  $R$  is an integral domain ( see later also, again use the properties of degrees). Finding irreducible polynomials is quite tricky and this will be a major topic. For example,  $x^2 + 1$  is irreducible when  $F = \mathbb{R}$  but not when  $F = \mathbb{Z}_2$ . If we take a ring like  $\mathbb{Z}_6[x]$ , it is not even an integral domain, since  $\bar{2}x \cdot \bar{3}x = \bar{0}$

**Definition:** An integral domain  $R$  is called a unique factorisation domain (UFD) if:

- for every element  $a \in R$  which is not 0 or a unit, there is a factorisation of  $a$  into irreducible elements  $a = p_1 p_2 \dots p_k$ .

- If  $a = p_1 p_2 \dots p_k$  and  $a = q_1 q_2 \dots q_l$  are two such factorisations of a non zero element  $a$  which is not a unit into irreducible elements, then  $k = l$  and after reordering, the pairs  $p_i, q_i$  are associates.

**Example 12** The standard UFD is of course  $\mathbb{Z}$ . The ring  $\mathbb{Z}_8[x]$  of polynomials with coefficients in  $\mathbb{Z}_8$  is not a UFD. For instance, the polynomial  $x^2 - 1 = (x - 1)(x + 1) = (x + 3)(x - 3)$ . It is easy to check that any linear polynomial is irreducible, since any constant non zero polynomial is a unit so a linear polynomial can only have divisors which are associates or units. So these two factorisations are into irreducible elements which cannot be reordered to be associates. Two linear polynomials are associates exactly when they are constant multiples of each other. We will see shortly that  $\mathbb{Z}_7[x]$  is a UFD!

Now we would like to imitate the method of Euclid in a general ring.

**Definition:** A ring  $R$  is called a Euclidean domain if it satisfies;

-  $R$  is an integral domain  
 - there is a function  $\phi : R \setminus \{0\} \rightarrow N$ , where  $N = \{0, 1, 2, \dots\}$  such that for any pair of non zero elements  $a, b$  in  $R$ , we can find elements  $q, r$  in  $R$  satisfying

$$b = aq + r, \quad \phi(r) < \phi(a) \quad \text{or} \quad r = 0$$

It now follows that such a ring is a unique factorisation domain.

**Theorem 4** If  $R$  is a Euclidean domain then  $R$  is a unique factorisation domain

**Proof:** The proof of uniqueness of factorisations follows exactly the same pattern as for  $\mathbb{Z}$ . However we need to show factorisations exist. Note that when dealing with  $\mathbb{Z}$ , factors get smaller and so the process of splitting up an element terminates. We need to discuss why this works for a ED. The problem is that we may keep splitting indefinitely otherwise. So the difficulty is that we might have factors  $\dots a_i | a_{i-1} | \dots a_1 | a$  where each element is a proper

divisor of the next one, where we start with an element  $a$  which is not zero or a unit and seek a finite factorisation of it into irreducibles.

We now give an argument similar to one in the next chapter. There the method is phrased in terms of ideals, whereas here we only use ideals briefly. Suppose that  $a$  is non zero and not a unit, but there is no factorisation into a finite number of irreducibles. Clearly then we can find an infinite sequence of proper divisors  $\dots a_i | a_{i-1} | \dots | a_1 | a$ , by splitting  $a$  into two factors,  $a = a_1 b_1$ , then splitting  $a_1$  into two factors etc. Define  $\mathcal{S}$  to be the set of all multiples of elements in the sequence  $\dots a_i, \dots a_1, a$ . It is easy to see that  $\mathcal{S}$  is closed under multiplication by arbitrary elements of  $R$ . Moreover if  $b, c \in \mathcal{S}$ , it follows that  $b \pm c \in \mathcal{S}$ , since  $b = ra_i, c = r'a_j$  with say  $i \leq j$  implies that  $b = r''a_j$  as  $a_j | a_i$  implies  $a_i = xa_j$ . So  $b \pm c = (r \pm r'')a_j \in \mathcal{S}$ . So we have shown that  $\mathcal{S}$  is an ideal!

Next, choose an element  $y \in \mathcal{S}$  so that  $\phi(y)$  is smallest for all elements of this ideal. We claim that  $\mathcal{S} = yR$  ie all the elements of the ideal are multiples of  $y$ . Now if  $x \in \mathcal{S}$ , then by the Euclidean domain property, we solve  $x = qy + r$ , where either  $r = 0$  or  $\phi(r) < \phi(y)$ . But the latter would imply that  $r = x - qy$  is an element of the ideal  $\mathcal{S}$  with a smaller  $\phi$  value than  $y$ , a contradiction. We conclude that  $r = 0$ , ie  $y | x$ . So  $\mathcal{S} = yR$ .

Finally,  $y \in yR$  implies  $y = ra_i$  for some  $i$ , i.e  $a_i | y$ . But we also know that  $y | a_i$  for all  $i$ , so this shows that  $y, a_i$  are associates. If we look next at  $a_{i+1}$ , this is a proper divisor of  $a_i$  and hence of the associate  $y$ . But again,  $y | a_{i+1}$ , so this is a contradiction, as we would then conclude that  $y, a_{i+1}$  are also associates and so  $a_i, a_{i+1}$  would be associates and the second would not be a proper divisor of the first. This completes the proof of the existence of a factorisation. ■

**Example 13** *A standard example of a Euclidean domain is a polynomial ring  $R = F[x]$  where  $F$  is any field. So the coefficients of the polynomials could be rationals, reals, complexes or in  $\mathbb{Z}_p$ , for some prime  $p$ . Here the function  $\phi$  is the degree of a polynomial, ie if  $p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0$ , where the coefficients  $a_i \in F$  with  $a_n \neq 0$ , then  $\phi(p(x)) = n$ . It is easy to see that the usual long division of polynomials satisfies the property required of giving a remainder with smaller degree. Moreover there is an additional special property  $\phi(p(x)q(x)) = \phi(p(x)) + \phi(q(x))$  for  $p(x), q(x) \neq 0$  which implies that  $F[x]$  is an integral domain. So this gives unique factorisation in polynomial rings.*

We will see in the next chapter, that for the ring  $\mathbb{Z}[x]$  of polynomials with integer coefficients, there is in fact no Euclidean domain structure. However there is unique factorisation and this is an integral domain.

**Example 14** *The Gaussian integers  $\mathbb{Z}[i]$  is a Euclidean domain. To show this, we define  $\phi : \mathbb{Z}[i] \rightarrow \mathbb{N}$  by  $\phi(x + yi) = x^2 + y^2 = |x + yi|^2$ . Now we need to establish division. Notice that given  $a = x + iy, b = u + iv$  in  $\mathbb{Z}[i]$ , we can try to divide  $b$  by  $a$ , defining  $\frac{b}{a} = \frac{b\bar{a}}{|a|^2} = c + id$  as usual. However this is in  $\mathbb{C}$  not in  $\mathbb{Z}[i]$ . So we want to take the closest element in  $\mathbb{Z}[i]$  to this number. Choose  $q = c' + id'$ , where  $c', d'$  are integers and  $|c' - c| \leq \frac{1}{2}$  and  $|d' - d| \leq \frac{1}{2}$ . Then  $b - qa = a(\frac{b}{a} - q) = a(c + id - c' - id') = a((c - c') + (d - d')i)$ . It follows that  $\phi(b - qa) = |b - qa|^2 = |a|^2((c - c')^2 + (d - d')^2) \leq \frac{1}{2}|a|^2 < \phi(a)$ . It is obvious that  $\mathbb{Z}[i]$  is an integral domain.*

**Example 15** As a comparison,  $\mathbb{Z}[\sqrt{-5}]$  is not a unique factorisation domain. For we can write  $6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$ . Now the claim is that all the factors are irreducibles. To show this, define  $\phi : \mathbb{Z}[\sqrt{-5}] \rightarrow \mathbb{N}$  by  $\phi(x + y\sqrt{-5}) = x^2 + 5y^2 = |x + y\sqrt{-5}|^2$ . Now this ring is an integral domain, so if we could follow the same process in the previous example, ie establish division, then the ring would be a Euclidean domain. However this fails. First of all we show the only units in  $\mathbb{Z}[\sqrt{-5}]$  are  $\pm 1$  so that none of the four factors above are units. For given elements  $z_1, z_2$  in the ring, clearly  $\phi(z_1 z_2) = \phi(z_1)\phi(z_2)$ . So if  $z_1 z_2 = 1$ , then  $\phi(z_1)\phi(z_2) = 1$  and so both of the integers  $\phi(z_1), \phi(z_2)$  are 1. But then if  $z = a + b\sqrt{-5}$ ,  $\phi(z) = a^2 + 5b^2 = 1$  implies  $a = \pm 1, b = 0$ . This completes the discussion about units. Finally we can show that the factors  $2, 3, 1 + \sqrt{-5}, 1 - \sqrt{-5}$  are all irreducible, since their  $\phi$  values are respectively 4, 9, 6, 6 and these are not divisible by any other  $\phi$  values except for  $\pm$  themselves or  $\pm 1$ .

*Last modified, 11 AM March 26*

## Summary notes for Algebra 321

### CHAPTER THREE - Factorisation in rings, part II

We give a characterisation of Unique factorisation domains (UFDs) firstly, then go on to study another condition midway between Euclidean domains (EDs) and UFDs.

**Definition:** An element  $p$  in an integral domain is called a *prime* if  $p$  is not 0 or a unit and whenever  $p|ab$ , then  $p|a$  or  $p|b$ .

Notice this is just the key property (\*) we used in the previous chapter to ensure unique factorisation occurs in  $\mathbb{Z}$  and in general EDs.

**Theorem 5** *Let  $R$  be an integral domain in which every non zero element can be written as a product of a finite number of irreducible elements. Then every irreducible element in  $R$  is prime if and only if  $R$  is a UFD.*

**Proof:** By the last chapter, clearly if every irreducible is prime, then  $R$  is a UFD. Conversely, suppose that  $R$  is a UFD and  $p$  is an irreducible element with  $p|ab$ . By definition, it follows that  $ab = pq$ . Now write each of  $a, b, q$  as a (unique) product of irreducible elements. Then by uniqueness of the factorization of the element  $c = ab$ , we see that the irreducible element  $p$  must occur as an associate of one of the irreducible factors of either  $a$  or  $b$ . But then this means  $p$  divides this irreducible factor and so  $p|a$  or  $p|b$  follows. ■

**Lemma 4** *A prime element is always irreducible in an integral domain.*

**Proof:** Suppose that  $p$  is prime but not irreducible. So  $p = ab$  where neither  $a$  nor  $b$  is a unit. But now, by the primeness assumption, it follows that  $p|a$  or  $p|b$ . By symmetry, assume the first, ie  $a = pu$ . Then  $p = pub$  and so since we are in an integral domain,  $1 = ub$ . But this means  $b$  is a unit with inverse  $u$ , contrary to assumption. ■

We can summarise by saying that if elements can be factorised into irreducibles, then a UFD is characterised by primes being the same as irreducible elements.

As a consequence, we get some more simple properties of UFDs.

**Theorem 6** *In a UFD, if  $a|b$  and  $a = p_1p_2..p_k$  with  $b = q_1q_2..q_l$  are factorisations into irreducibles, then after reordering  $(p_1, q_1), (p_2, q_2)..(p_k, q_k)$  are pairs of associates and  $k \leq l$ . Moreover two non zero elements in a UFD always have a g.c.d.*

**Proof:** This follows immediately from the previous theorem. Note that since an irreducible is prime, we have  $p_1|a$ , so  $p_1|b = q_1(q_2..q_l)$  and so either  $p_1|q_1$  or  $p_1|q_2q_3..q_l$ . Continuing on ( or using induction) it follows that  $p_1|q_j$  for some  $j$ . But an irreducible is only divisible by a unit or an associate of itself. We conclude that  $p_1, q_j$  are associates for some  $j$ . We can then do this successively for  $p_2, \dots, p_k$  and see also that  $k \leq l$ .

Finally given  $a, b$ , we can find their factorisations into irreducibles and put together all common associate pairs to form a g.c.d  $d$ . Note then certainly  $d|a$  and  $d|b$  since the irreducibles of  $d$  are all associates of irreducibles of  $a$  and of  $b$ . Conversely if  $c|a$  and  $c|b$ , by the above argument, all the irreducibles in a factorisation of  $c$  must be amongst the associates of irreducibles of both  $a$  and  $b$  and hence are associates of some of the irreducibles of  $d$ . So  $c|d$ .

■

**Definition:** An integral domain is called a *principal ideal domain* (PID) if every ideal is principal, ie  $I = aR, a \in R$ .

Why is this idea relevant to factorisation? How do we check this condition?

**Example 16** We saw in chapter one, that  $\mathbb{Z}[x]$  was not a principal ideal domain, since the ideal  $I = \langle 2, x \rangle$  generated by the elements  $2, x$  consists of all polynomials with even constant term and this is clearly not a principal ideal. For if  $a \in I$ , then the principal ideal  $a\mathbb{Z}[x]$  does not contain 2 if  $a$  has zero constant term and it does not contain  $x$  if  $a$  has non zero constant term. So  $I = a\mathbb{Z}[x]$  is impossible. On the other hand, it is easy to see that  $\mathbb{Z}$  is a PID since any ideal is of the form  $n\mathbb{Z}$ .

**Lemma 5** Let  $R$  be an integral domain.

- (a)  $a|b$  if and only if  $aR \supseteq bR$
- (b)  $a$  is a unit if and only if  $aR = R$
- (c)  $a, b$  are associates if and only if  $aR = bR$
- (d)  $a$  is a proper divisor of  $b$  if and only if  $R \supset aR \supset bR$

**Proof:** We sketch this. (a)  $a|b$  means  $b = ac$  so all elements of the form  $br$  are also then  $acr = ar'$  and so all elements of  $bR$  are in  $aR$ . Next (b)  $a$  is a unit means  $a|1$ . By (a) then  $aR = R$ . Conversely,  $aR = R$  implies that  $ar = 1$  for some  $r$  and so  $a$  is a unit. For (c), just apply (a) twice, as associates divide each other. For (d), just apply (a) and (c).

■

**Theorem 7** Every PID is a UFD.

**Proof:** We would like to show first that any element  $a$  which is non zero and a not a unit in a PID  $R$ , can be written as a finite product of irreducibles. Now we would like to follow the same strategy as in  $\mathbb{Z}$ , namely to split  $a$  into proper factors and keep splitting these up until we reach irreducibles. However the same problem arises as in the previous chapter, where for a ED we found that there might be a sequence of proper divisors  $\dots a_i|a_{i-1}|a_{i-2}|\dots a_1|a$ . So the process of factorisation never ends in finitely many irreducibles.

To get around this, we use the PID property. Note to show factorisation into irreducibles in an ED, we were really using that an ED is a PID. Now consider the ideals

$\dots a_i R \supset a_{i-1} R \supset \dots a_1 R \supset aR$ . We can take the union of all these ideals, which are getting larger and larger. It is easy to check this union is itself an ideal, any two elements of the union will lie in some ideal and so their sum, difference and product are in the union. Similarly the product of an element by an arbitrary element of the ring is in the union. Finally, since  $R$  is a PID, we see that this union is  $bR$  for some element  $b$ . But then  $b \in a_i R$  for some  $i$  which implies that  $bR = a_i R$  and then the ideals are not strictly increasing. This is a contradiction and the finite factorisation into irreducibles follows!

Next, we show uniqueness of factorisations. We need to prove property (\*) again, ie if  $p$  is irreducible and  $p|ab$  then  $p|a$  or  $p|b$ , ie that every irreducible is prime. This completes the argument as before. Suppose that  $p$  does not divide  $a$ . Then form the ideal  $I = \langle p, a \rangle$  generated by  $p$  and  $a$ . We know that  $I = cR$  for some element  $c$ , since each ideal is principal. So  $c|p$  and  $c|a$ . If  $c, p$  are associates, then we would see that  $p|a$  contrary to assumption. Hence it must be the case that  $c$  is a unit, since  $p$  irreducible has either units or associates as divisors. But now we see that  $R = I = cR$  since any ideal containing a unit is the whole ring. So  $1 \in I$  and we see that  $1 = xp + ya$ , as  $I$  is the linear combinations of  $p, a$  since this is the (smallest) ideal generated by  $p, a$ . But now we are back in the same territory as with  $\mathbb{Z}$ , namely we can multiply by  $b$  and get  $b = bxp + yab$  and both terms on the right side are divisible by  $p$ . So  $p|b$  and (\*) is proved. ■

**Theorem 8** *Every ED is a PID.*

**Proof:** Let  $I$  be an ideal which is not  $\{0\}$ . We need to show that  $I = aR$  for some element  $a \in I$ . Choose  $a \in I$  with  $\phi(a)$  minimal amongst all elements of  $I \setminus \{0\}$ . Now suppose  $b \in I$ . By the assumption that  $R$  is an ED, we know that  $b = aq + r$  with  $\phi(r) < \phi(a)$  or  $b = aq$ . But also  $b - aq = r$  is in  $I$ , so we have found either a non zero element of  $I$  with smaller value of  $\phi$ , a contradiction, or else  $b = aq$ . So this shows that every element of  $I$  is in  $aR$  so the ideal is principal. ■

Note this immediately tells us why the ring  $\mathbb{Z}[x]$  is definitely not an ED since we know it is not a PID! This theorem is therefore a good way of telling if a ring is NOT an ED, by seeing if any ideals are not principal.

An important construction is how to embed an integral domain  $R$  into a field  $F$ , called the *field of fractions*. We will use this in the next chapter - it is based on how to embed the integers in the rationals.

Recall the idea of an equivalence relation  $\sim$ .

**Definition:** An equivalent relation on a set  $S$  satisfies;

- $x \sim x$  for all  $x \in S$  (reflexive)
- $x \sim y$  implies  $y \sim x$  (symmetric)
- $x \sim y$  and  $y \sim z$  implies  $x \sim z$  (transitive)

Given an equivalence relation, the equivalence classes form a partition of  $S$ , ie  $S$  is divided up into disjoint classes which are each all elements equivalent to a given element.

Cosets are a basic example of an equivalence relation, where we say  $x \sim y$  exactly when  $x - y \in I$  for an ideal  $I$ .

Start with the Cartesian product  $S = R \times R$ . We define an equivalence relation on  $S$  by  $(a, b) \sim (c, d)$  when  $ad = bc$ . Notice this means we are really thinking of a pair  $(a, b)$ , where  $b \neq 0$  as  $\frac{a}{b}$ . Let  $F$  denote the set of equivalence classes, so an element of  $F$  consists of all pairs  $(a, b)$  equivalent to some fixed pair  $(a_0, b_0)$ . We denote this equivalence class by  $[(a_0, b_0)]$ .

Next, define a ring structure on  $F$  as follows,

Then  $[(a, b)] \pm [(c, d)] = [(ad \pm bc, bd)]$  and  $[(a, b)] \cdot [(c, d)] = [(ac, bd)]$  define addition, subtraction and multiplication. It is now not hard to see that this defines a field.

The main problem is first to show that the operations are well defined. (We can write  $[(a, b)] = [(xa, xb)]$  for any  $x \in R$  with  $x \neq 0$ , by our definition of the equivalence relation). Notice that if  $[a, b] = [a', b']$  and  $[c, d] = [c', d']$ , then  $ab' = a'b$  and  $cd' = c'd$ . So  $[(a, b)] \pm [(c, d)] = [(ad \pm bc, bd)] = [(a'c'ad \pm a'c'bc, a'c'bd)] = [(a'cad' \pm ac'b'c, acb'd')] = [(a'd' \pm b'c', b'd')] = [(a', b')] \pm [(c', d')]$ . Also  $[(a, b)] \cdot [(c, d)] = [(ac, bd)] = [(a'c'ac, a'c'bd)] = [(a'c'ac, acb'd')] = [(a'c', b'd')] = [(a', b')] \cdot [(c', d')]$ . So this shows the operations on  $F$  work independent of choice of element in the equivalence classes. ( We should look at the extra cases where the multiplication of first and second terms might be by a zero factor. In this case, the argument is easy also).

We also check that elements have multiplicative and additive inverses. Notice that  $[(a, b)] + [(-a, b)] = [(ab - ab, b^2)] = [(0, b^2)] = [(0, 1)]$ , which is the additive zero in  $F$ . Next,  $[(a, b)] \cdot [(b, a)] = [(ab, ab)] = [(1, 1)]$ , the multiplicative identity element, so long as  $a \neq 0$ .

The final observation will be that  $R$  embeds in  $F$  by the map  $\psi : a \rightarrow [(a, 1)]$ . So it suffices to show that if  $[(a, 1)] = [(b, 1)]$ , then  $a = b$ . Now the former implies that  $[(a, 1) - (b, 1)] = [(0, 1)]$ , the 0 element in  $F$ . So  $[(a - b, 1)] = [(0, 1)]$ . But this implies that  $1 \cdot (a - b) = 1 \cdot 0$  and so  $a - b = 0$  as required.

**Example 17** *If we take  $R = \mathbb{Z}[i]$ , the Gaussian integers, then  $F = \mathbb{Q}[i]$ . An important example is when  $R = \tilde{F}[x]$  is the ring of polynomials with coefficients in a field  $\tilde{F}$ . In this case, the field  $F$  is all rational functions, ie quotients of polynomials in the variable  $x$  with coefficients in  $\tilde{F}$ .*

*Last modified 11 AM March 26*

## Summary notes for Algebra 321

### CHAPTER FOUR - Factorisation of polynomials

Our aim in this chapter is to study polynomials in  $\mathbb{Z}[x]$  and their factorisations. In particular, we will prove that  $\mathbb{Z}[x]$  is an UFD and also will find criteria for polynomials with integer coefficients to be irreducible. The key idea is to relate factorisation in  $\mathbb{Z}[x]$  to factorisation in  $\mathbb{Q}[x]$ . The methods extend quite easily to any polynomial ring  $R[x]$ , where  $R$  is an UFD. Then one replaces  $\mathbb{Q}[x]$  by  $F[x]$ , where  $F$  is the field of fractions of  $R$ .

We begin with a useful way of checking irreducibility of polynomials in  $\mathbb{Z}[x]$ . For a polynomial  $f(x) = a_0 + a_1x + \dots + a_nx^n$ , where  $a_n \neq 0$ , we call  $a_n$  the leading term of  $f(x)$ . The image of  $f(x)$  in  $\mathbb{Z}_p[x]$  is then  $\bar{f}(x) = \bar{a}_0 + \bar{a}_1x + \dots + \bar{a}_nx^n$ . Here  $\bar{a}$  is the congruence class of  $a \pmod{p}$ .

**Theorem 9** *Suppose that  $f(x)$  is a non constant polynomial in  $\mathbb{Z}[x]$  and let  $p$  be a prime which does not divide the leading coefficient  $a_n$  of  $f(x)$ . If  $\bar{f}(x)$  is irreducible, then so is  $f(x)$ .*

**Proof:** If  $f(x) = g(x)h(x)$  in  $\mathbb{Z}[x]$ , then  $\bar{f}(x) = \bar{g}(x)\bar{h}(x)$  in  $\mathbb{Z}_p[x]$ , since the projection  $\mathbb{Z}[x] \rightarrow \mathbb{Z}_p[x]$  is a ring homomorphism. If  $b_m, c_k$  are the leading coefficients of  $g(x), h(x)$  respectively, then by multiplying out,  $a_n = b_m c_k$ . (This also means that  $n = m + k$ .) Now since  $p$  does not divide  $a_n$ , then it does not divide either  $b_m, c_k$ . So the degrees of  $\bar{g}(x), \bar{h}(x)$  are the same as those of  $g(x), h(x)$  and we see that  $\bar{f}(x)$  is not irreducible. ■

Notice that we need to check that the polynomials in  $\mathbb{Z}_p[x]$  do not become constants, ie units, for otherwise we would not find a useful factorisation of  $\bar{f}(x)$ .

Since there are only finitely many polynomials in  $\mathbb{Z}_p[x]$  of bounded degree, it can be checked which are irreducible by a ‘sieve’ method. So we start with constants, linear, then quadratic polynomials etc and check which are divisible by lower degree ones. Notice that constants are always units ( $\mathbb{Z}_p$  is a field) and linear polynomials are always irreducible. So the first interesting case is quadratics. For  $p = 2$ , the polynomials in order are  $0, 1, x, x + 1, x^2, x^2 + 1, x^2 + x + 1, x^3, \dots$ . It is easy to see that  $x^2 + x + 1$  is the only irreducible quadratic, since  $x^2 + 1 = (x + 1)^2$ . As an example, the theorem then tells us that any quadratic polynomial  $a_0 + a_1x + a_2x^2$  where all coefficients are odd integers, is irreducible since it maps to  $x^2 + x + 1$ .

**Definition:** We call a polynomial  $f(x) = a_0 + a_1x + \dots + a_nx^n$  in  $\mathbb{Z}[x]$  *primitive* if the greatest common divisor of the coefficients  $\{a_0, a_1, \dots, a_n\}$  is 1. Given a polynomial  $g(x)$  in  $\mathbb{Q}[x]$ , we can always write  $g(x) = cg_0(x)$ , where  $g_0(x)$  is a primitive polynomial in  $\mathbb{Z}[x]$ .

We explain why the factorisation  $g(x) = cg_0(x)$  is unique up to sign of  $c$ . Suppose that  $g(x) = dg_1(x)$  where  $g_1(x)$  is primitive (in  $\mathbb{Z}[x]$ ). We can write  $c = c_1/c_2$  and  $d = d_1/d_2$  where the pairs of integers  $c_1, c_2$  and  $d_1, d_2$  are relatively prime. Then

$$c_1d_2g_0(x) = c_2d_1g_1(x).$$

If  $g_0(x) = a_0 + a_1x + \dots + a_nx^n$  and  $g_1(x) = b_0 + b_1x + \dots + b_nx^n$ , then we get

$$c_1d_2a_i = c_2d_1b_i,$$

for each  $1 \leq i \leq n$ . But since the greatest common divisor of all the  $a_i$  or all the  $b_i$  is 1 (since  $g_0, g_1$  are primitive), we have the greatest common divisor of all the  $c_1d_2a_i$  is  $c_1d_2$  which is the same (up to sign) as the greatest common divisor of all the  $c_2d_1b_i$  which is  $c_2d_1$ . Hence we see that  $c_1d_2 = \pm c_2d_1$  ie  $c = \pm d$  and  $g_0 = \pm g_1$  as claimed.

**Definition:**  $c$  is called the *content* of  $g(x)$ .  $g_0(x)$  is called the *associated primitive* to  $g(x)$ . Both of these are well defined up to sign. (Generally if we are working with an arbitrary UFD  $R$  as the ring of coefficients, then the content and associated primitive are well defined up to multiplication by a unit).

Note that if  $f(x)$  is in  $\mathbb{Z}[x]$ , then  $c$  is an integer and is the greatest common divisor of the coefficients of  $f(x)$ .

**Theorem 10 (Gauss lemma)** *A product of primitive polynomials in  $\mathbb{Z}[x]$  is primitive*

**Proof:** Suppose that  $h(x) = f(x)g(x)$  are polynomials in  $\mathbb{Z}[x]$ , where  $f(x), g(x)$  are primitive. Suppose that  $p$  is a prime and  $p$  divides all the coefficients of  $h(x)$ . Then using the ring homomorphism  $\mathbb{Z}[x] \rightarrow \mathbb{Z}_p[x]$ , we find that  $\bar{h}(x) = \bar{0}$ . Since  $\mathbb{Z}_p[x]$  is an integral domain, then  $\bar{f}(x)\bar{g}(x) = \bar{0}$  implies either  $\bar{f}(x) = \bar{0}$  or  $\bar{g}(x) = \bar{0}$ . But the latter implies either  $p$  divides all the coefficients of  $f(x)$  or  $g(x)$ , contradicting our assumption that these polynomials are primitive. We conclude that the greatest common divisor of the coefficients of  $h(x)$  is 1, ie  $h(x)$  is primitive. ■

**Corollary 2** *If  $f(x)$  is a primitive polynomial in  $\mathbb{Z}[x]$  and  $g(x)$  is another polynomial with integer coefficients, then  $f|g$  in  $\mathbb{Q}[x]$  if and only if  $f|g$  in  $\mathbb{Z}[x]$ .*

**Proof:** We need to show that if  $g = fh$  where  $h(x)$  is a polynomial with rational coefficients, then in fact the coefficients of  $h$  are integers. Write  $h(x) = ch_0(x)$ , where  $h_0$  is primitive (and so has integer coefficients) and  $c$  is its content. Now by the Gauss lemma, we know that  $fh_0$  is a primitive polynomial. But then  $g = cfh_0$  is (unique up to sign) the factorisation of  $g$  into its content and associated primitive. We conclude that since  $g$  has integer coefficients, that  $c$  is an integer and so  $h = ch_0$  is actually in  $\mathbb{Z}[x]$  as claimed. ■

**Corollary 3** *An irreducible non constant polynomial  $f(x)$  in  $\mathbb{Z}[x]$  is irreducible in  $\mathbb{Q}[x]$ .*

**Proof:** If  $f = gh$  where  $g, h$  are in  $\mathbb{Q}[x]$ , then we can write  $g = cg_0$  where  $g_0$  is the associated primitive and so  $f = g_0ch$ . By Corollary 1, since  $g_0$  is primitive and divides  $f$  in  $\mathbb{Q}[x]$  it divides  $f$  in  $\mathbb{Z}[x]$ , ie  $ch$  has integer coefficients. But this contradicts our assumption that  $f$  is irreducible in  $\mathbb{Z}[x]$ . ■

**Corollary 4** *If  $f, g$  are in  $\mathbb{Q}[x]$  and  $f_0, g_0$  are their associated primitives, then  $f|g$  in  $\mathbb{Q}[x]$  if and only if  $f_0|g_0$  in  $\mathbb{Z}[x]$ .*

**Proof:** Certainly if  $f_0|g_0$  in  $\mathbb{Z}[x]$  then  $f|g$  in  $\mathbb{Q}[x]$ . Conversely, if  $f|g$  in  $\mathbb{Q}[x]$  then  $f_0|g_0$  in  $\mathbb{Q}[x]$ . By Corollary 1,  $f_0|g_0$  in  $\mathbb{Z}[x]$ , proving the result. ■

**Theorem 11** *Irreducible polynomials in  $\mathbb{Z}[x]$  are either prime numbers or primitive polynomials which are irreducible in  $\mathbb{Q}[x]$ .*

**Proof:** Clearly for a constant polynomial in  $\mathbb{Z}[x]$ , irreducibility is the same as being a prime number. So we may assume that  $f(x)$  in  $\mathbb{Z}[x]$  is not constant. Now  $f = cf_0$ , where  $c$  is the content and  $f_0$  is primitive. Since  $f$  has integer coefficients,  $c$  is an integer and so  $c = \pm 1$  if  $f$  is irreducible. Finally by Corollary 2, if  $f$  is irreducible in  $\mathbb{Z}[x]$  then it is irreducible in  $\mathbb{Q}[x]$ . The converse is obvious, since irreducibility in  $\mathbb{Q}[x]$  is ‘stronger’ than irreducibility in  $\mathbb{Z}[x]$ . ■

**Theorem 12** *Irreducible polynomials in  $\mathbb{Z}[x]$  are prime.*

**Proof:** Suppose that  $f$  is irreducible and  $f|gh$  in  $\mathbb{Z}[x]$ . Let  $g = cg_0$  and  $h = dh_0$  be the usual factorisations in contents and primitives. By the Gauss lemma,  $g_0h_0$  is primitive and so  $f|cd(g_0h_0)$  is the factorisation into content and primitive. If  $f$  is constant, by Theorem 3,  $f$  is a prime and so  $f|cd$  implies either  $f|c$  or  $f|d$  and therefore  $f|g$  or  $f|h$ . (A number dividing a polynomial with integer coefficients clearly divides its content).

If  $f$  is not constant, by Theorem 3 again  $f$  is primitive and irreducible in  $\mathbb{Q}[x]$ . Therefore since we know that irreducibles are primes in the ED  $\mathbb{Q}[x]$ ,  $f|gh$  implies that  $f|g$  or  $f|h$  in  $\mathbb{Q}[x]$  and hence in  $\mathbb{Z}[x]$  by Corollary 1. ■

**Theorem 13**  *$\mathbb{Z}[x]$  is an UFD. Moreover any polynomial  $f$  in  $\mathbb{Z}[x]$  can be factorised as  $f(x) = p_1 \dots p_k q_1(x) \dots q_m(x)$ , where  $p_i$  are prime numbers,  $q_j(x)$  are primitive irreducible polynomials in  $\mathbb{Z}[x]$  and the factorisation is unique up to signs and reordering.*

**Proof:** We know that factorisations terminate in irreducibles in  $\mathbb{Z}[x]$ , since degrees go down. So existence of such a factorisation is easy and uniqueness follows from the previous theorem as we have seen before. ■

Notice the same methods work for  $R[x]$  and  $F[x]$ , where  $R$  is any UFD and  $F$  is its field of fractions. For instance, the content of a polynomial  $f(x)$  in  $R[x]$  is again the greatest common divisor of the coefficients and is defined up to multiplication by a unit. A primitive polynomial means that the greatest common divisor of the coefficients is a unit. The Gauss lemma is proved by mapping to  $R/pR[x]$ , where  $p$  is a prime or irreducible in  $R$  (which are identical notions).

To study irreducibility of a polynomial in  $\mathbb{Q}[x]$ , we may as well assume the polynomial lies in  $\mathbb{Z}[x]$  and is primitive, by adjusting it, multiplying by a fraction as necessary.

**Theorem 14 Rational root test** *If  $f(x) = a_0 + \dots + a_n x^n$  is a primitive polynomial in  $\mathbb{Z}[x]$  and  $r/s$  is a rational root of  $f$  with the g.c.d of  $\{r, s\}$  being 1, then  $r|a_0$  and  $s|a_n$ .*

**Proof:** If we substitute in for the root and multiply through by  $s^n$ , we get  $a_0 s^n + a_1 r s^{n-1} + \dots + a_n r^n = 0$ . Since  $r$  divides all terms after the first, it also divides the first term. Then since  $r, s$  have g.c.d 1, it follows that  $r|a_0$ . The argument for  $s$  is the same. ■

Note if  $f(x)$  is monic, ie has leading coefficient  $a_n = 1$ , then any rational root is an integer.

**Theorem 15 Eisenstein's criterion** Let  $f(x) = a_0 + \dots + a_n x^n$  be a polynomial in  $\mathbb{Z}[x]$ . Suppose there is a prime number  $p$  so that  $p|a_0, \dots, p|a_{n-1}$  but  $p$  does not divide  $a_n$  and  $p^2$  does not divide  $a_0$ . Then  $f$  is irreducible in  $\mathbb{Z}[x]$ .

**Proof:** Using the map to  $\mathbb{Z}_p[x]$ , we see that  $\bar{f} = \bar{a}_n x^n$ . If  $f = gh$  in  $\mathbb{Z}[x]$ , then  $\bar{g}\bar{h} = \bar{a}_n x^n$ . Hence both  $\bar{g}, \bar{h}$  are monomials, ie single powers of  $x$ . But then we see that all the coefficients of  $g, h$  except for the leading term (highest power) are divisible by  $p$ . Therefore both constant terms  $c_0, d_0$  of  $g, h$  are divisible by  $p$  and the constant term of  $f$  which is  $a_0 = c_0 d_0$  is divisible by  $p^2$ . This contradicts our hypotheses. ■

Notice that we can also do this 'backwards', ie there is a similar criterion with  $p|a_1 \dots p|a_n$  and  $p$  not dividing  $a_0$  nor  $p^2$  dividing  $a_n$ .

**Example 18** A polynomial like  $3 + 9x - 15x^2 - x^3$  satisfies the Eisenstein condition with  $p = 3$  and is therefore irreducible.

## Summary notes for Algebra 321

### CHAPTER FIVE - Modules

**Definition:** A module  $M$  over a ring  $R$  with identity 1 is an Abelian group  $M$  together with a scalar multiplication  $\phi : R \times M \rightarrow M$  denoted  $\phi(r, u) = ru$ . Scalar multiplication satisfies the following four rules;

- (1)  $r(u+v) = ru + rv$
- (2)  $(r+s)u = ru + su$
- (3)  $(rs)u = r(su)$
- (4)  $1u = u$

Here  $r, s \in R$  and  $u, v \in M$ . We will mainly consider modules where  $R$  is a commutative ring.

**Example 19** An Abelian group  $A$  is a  $\mathbb{Z}$  module. Scalar multiplication is interpreted as  $ng = g + \dots + g$ , the sum of  $n$  copies of  $g \in A$  if  $n > 0$ . Similarly  $0g = 0$  and  $(-n)g = n(-g)$ .

**Example 20** If  $F$  is a field, then a module  $V$  over  $F$  is just a vector space.

So modules incorporate features of both vector spaces and Abelian groups. They also include basic ideas from the theory of rings in the first two chapters of the notes.

**Example 21** A ring  $R$  with identity forms a module over  $R$ . So here  $M = R$ . Addition is the usual operation in  $R$  and scalar multiplication is the ring multiplication.

**Example 22** A key example which gives an important connection between the theory of linear transformations and matrices with module theory is as follows; let  $V$  be a vector space over a field  $F$  and let  $\alpha$  be a linear transformation acting on  $V$ . So  $\alpha(u+v) = \alpha(u) + \alpha(v)$  and  $\alpha(ku) = k\alpha(u)$  for all  $u, v \in V$  and  $k \in F$ . We can make  $V$  into a  $F[x]$  module by defining  $f(x)v = f(\alpha)v = a_0v + a_1\alpha(v) + \dots + a_n\alpha^n(v)$ , where  $f(x) = a_0 + a_1x + \dots + a_nx^n$ . Notice we are exploiting the similarity between the way a linear transformation acts on vectors and scalar multiplication in modules to make polynomials 'in  $\alpha$ ' act on the vector space. In fact we could also use the ring  $F[\alpha]$  as the scalars but it is more convenient to use the usual polynomial ring  $F[x]$ .

**Definition:** A non empty subset  $N$  of a module  $M$  is called a submodule if  $N$  is a subgroup of  $M$  (considered as an Abelian group) and given  $r \in R$  and  $u \in N$  then  $ru \in N$ .

In other words, it is sufficient to check that  $N$  is closed under addition and scalar multiplication. This is because, if  $u \in N$  then  $(-1)u = -u$  so  $-u \in N$ .

We immediately see another important example, connecting module theory and ring theory.

**Example 23** Let  $R$  be a commutative ring with 1 and consider  $R$  as a module over  $R$ . A submodule  $I$  of  $R$  is then precisely an ideal of the ring, since closure under addition and scalar multiplication defines an ideal.

**Definition:** A homomorphism of modules  $M, N$  over the same ring  $R$  is a map  $f : M \rightarrow N$  satisfying;

$$(1) f(u + v) = f(u) + f(v)$$

$$(2) f(ru) = rf(u)$$

where  $u, v \in M$  and  $r \in R$ . Notice the second condition requires the same ring for both modules.

It is easy to check that the kernel and image of a homomorphism are both submodules. We can also define quotient modules much like the situation in rings.

**Definition:** Given a pair  $M, K$  of a module and submodule, the quotient module  $M/K$  is defined as having elements given by the cosets  $u + K$  where  $u \in M$ . Addition and scalar multiplication of cosets is defined by  $(u + K) + (v + K) = (u + v) + K$  and  $r(u + K) = ru + K$ , for  $u, v \in M$  and  $r \in R$ . The quotient map  $\pi : M \rightarrow M/K$  is the onto homomorphism  $\pi(u) = u + K$  sending each element to its coset.

**Theorem 16 (The first isomorphism theorem)** *If  $f : M \rightarrow N$  is a homomorphism between  $R$ -modules,  $K$  is the kernel and  $L$  the image of  $f$  then  $M/K \cong L$ .*

Next we discuss generating sets (often called spanning sets), the same notion as that in vector spaces.

**Definition:** If  $S$  is a subset of a module  $M$ , then the submodule denoted  $\langle S \rangle$  generated by  $S$  is defined as all  $R$ -linear combinations of elements of  $S$ . So  $\langle S \rangle = \{r_1u_1 + \dots + r_ku_k; r_i \in R, u_i \in S\}$ . It is easy to check that this is indeed a submodule of  $M$ . If  $M = \langle S \rangle$  we say that  $S$  generates  $M$ . If  $S$  is a finite set, then  $M$  is called a finitely generated module.

An important special case is the submodule generated by a single element, ie when  $S = \{x\}$ . In this case, we say that  $\langle S \rangle = Rx$  is a cyclic submodule of  $M$  and consists of all scalar multiples of  $x$ . If  $M = Rx$  we say that  $M$  is a cyclic module.

**Definition:** Given  $R$ -modules  $M_1, M_2, \dots, M_k$ , we can define their direct sum  $M = M_1 \oplus M_2 \oplus \dots \oplus M_k$  by taking elements of  $M$  of the form  $(u_1, \dots, u_k)$ , where each  $u_i \in M_i$ . The operations are then defined as  $(u_1, \dots, u_k) + (v_1, \dots, v_k) = (u_1 + v_1, \dots, u_k + v_k)$  and  $r(u_1, \dots, u_k) = (ru_1, \dots, ru_k)$ . The direct sum of  $n$  copies of  $R$  is denoted by  $R^n$  and is called a free  $R$ -module.

We would like to characterise free  $R$ -modules in terms of bases, just like for vector spaces.

**Definition:** A basis of an  $R$ -module  $M$  is a linearly independent generating set  $S$ . Any module containing a basis is called free.  $S$  is called linearly independent if whenever  $r_1u_1 + \dots + r_ku_k = 0$ , where  $u_i \in S$ , then  $r_1 = \dots = r_k = 0$ .

Notice that  $S$  need not be finite, however for our purposes, we are mainly interested in the case of a free finitely generated  $R$ -module.

**Theorem 17** *A finitely generated free  $R$ -module  $M$  is isomorphic to  $R^n$ , where  $n$  is the number of elements in a basis for  $M$ .*

**Proof:** This is easy, we use the existence of a basis  $S = \{u_1, \dots, u_n\}$  to construct the isomorphism, representing module elements by vectors of coordinates, just as in vector spaces. So given  $v \in M$ , we can write  $v = r_1u_1 + \dots + r_nu_n$  since  $S$  is a generating set. Then define  $\psi : M \rightarrow R^n$  by  $\psi(v) = (r_1, r_2, \dots, r_n)$ . It is easy to check this is a well defined homomorphism which is onto, since any linear combination occurs of the elements of  $S$  in  $M$ .  $\psi$  is well defined since if  $v = r_1u_1 + \dots + r_nu_n = r'_1u_1 + \dots + r'_nu_n$  are two different representations of  $v$ , then  $(r_1 - r'_1)u_1 + \dots + (r_n - r'_n)u_n = 0$  and so since the elements of  $S$  are linearly independent, we have  $r_1 - r'_1 = \dots = r_n - r'_n = 0$ . Finally, if  $\psi(v) = 0$ , then all  $r_i = 0$  and so  $v = 0$ . This shows that the kernel of  $\psi$  is  $\{0\}$  and  $\psi$  is one-to-one. ■

Notice that a standard basis for  $R^n$  is given by the elements  $e_1 = (1, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0), \dots, e_n = (0, 0, \dots, 0, 1)$ . For it is easy to check these form a linearly independent generating set.

**Example 24** *An important example of a module over a non commutative ring in the theory of groups comes from the group ring  $\mathbb{Z}G$ . Let  $G$  be a group and let  $\mathbb{Z}G$  be the free Abelian group with basis all the elements of  $G$ . We will illustrate in the case of the symmetric group on 3 letters, which has permutations  $(1), (12), (13), (23), (123), (132)$ , which we denote by  $g_0 = 1, g_1, g_2, g_3, g_4, g_5$  respectively. So elements of  $\mathbb{Z}G$  are linear combinations  $n_0g_0 + \dots + n_5g_5$  where  $n_i \in \mathbb{Z}$ . Addition of elements is defined in the obvious way, for multiplication we use the distributive law to write  $(n_0g_0 + \dots + n_5g_5)(m_0g_0 + \dots + m_5g_5) = n_0m_0g_0g_0 + n_0m_1g_0g_1 + \dots + n_5m_5g_5g_5$ . So we can use properties of multiplication in  $G$  ( eg  $g_5^2 = g_4$  etc, to write the right side as a linear combination of elements of  $G$  again. In this way,  $\mathbb{Z}G$  becomes a non commutative ring and hence a module. If  $G$  is a finite group, this is a finitely generated free Abelian group ( forgetting multiplication), whereas if  $G$  is infinite, this is infinitely generated.*

**Definition:** A torsion element in a module  $M$  over a commutative ring  $R$  with identity 1 is an element  $u \neq 0$  for which  $ru = 0$  for some  $r \neq 0, r \in R$ .

It is easy to see that free  $R$  modules do not have any torsion elements, so long as the ring  $R$  is an integral domain.

In an  $R$ -module, where  $R$  is an integral domain, the torsion elements fit together to form a submodule called the torsion submodule of  $M$ .

**Lemma 6** *Let  $M$  be a module over an integral domain  $R$ . The set  $T$  of torsion elements, together with 0, form a submodule of  $M$ .*

**Proof:** Suppose that  $u, v$  are torsion elements of  $M$ . Then  $ru = 0$  and  $r'v = 0$ , where  $r, r' \in R$  are not zero. Then we see that  $rr'(u + v) = 0$ , but  $rr' \neq 0$  since  $R$  is an integral domain. Similarly it is clear that  $su$  is a torsion element for any  $s \in R$ , since  $rsu = 0$  and if  $s \neq 0$ , then  $rs \neq 0$ . So we see that  $T$  is a submodule. ■

We next want to analyse finitely generated free  $R$ -modules, before the main structure theorem for all finitely generated modules over a PID. Our ultimate aim will be to split a finitely generated module up into a direct sum of cyclic submodules in a 'canonical' way, so that the torsion modules is split up and the remainder is a free module. A first step is to show that the dimension or rank of a free module is an invariant, ie well defined.

**Theorem 18** *If  $R$  is a commutative ring with 1, then  $R^n \cong R^m$  if and only if  $m = n$ .*

We will deduce this from another result.

**Theorem 19** *If  $R$  is a commutative ring with 1, then any linearly independent set of  $R^n$  has at most  $n$  elements.*

**Proof:** Induction on  $n$ . We start with the case  $n = 1$ . Then the module is  $R$ . Given any set  $S = \{u_1, u_2, \dots\}$  with at least two elements, we find that  $u_2u_1 - u_1u_2 = 0$  and so there is a linear relation between the elements with non zero coefficients. (Of course if there are any zero elements in  $S$  then  $r0 = 0$  shows that  $S$  is linearly dependent. This shows that a linearly independent set in  $R$  has at most one element.

Next assume the result true for  $R^{n-1}$ . Define  $F \subset R^n$  by  $F = \{(0, r_2, r_3, \dots, r_n)\}$ . Clearly  $F \cong R^{n-1}$  and  $F$  is a submodule of  $R^n$ . Suppose that  $S$  is a linearly independent subset of  $R^n$  and let  $\pi : R^n \rightarrow R$  be the projection sending  $(r_1, r_2, \dots, r_n)$  to  $r_1$ . Clearly  $F$  is the kernel of  $\pi$ .

If all elements of  $S = \{u_1, \dots, u_m\}$  are in  $F$ , by our induction hypothesis, it follows that  $m \leq n - 1$  and so certainly  $m \leq n$ . So we can assume that the ordering is chosen so that  $u_1$  is not in  $F$ , ie its first coordinate is non zero.

Now by our argument for the case  $n = 1$ , it follows that  $\pi(u_1)$  and  $\pi(u_i)$  are linearly dependent, for  $i = 2, 3, \dots, m$ . So  $\pi(x_iu_1 + xu_i) = 0$ , where  $x_i$  is the negative of the first coordinate of  $u_i$  and  $x$  is the first coordinate of  $u_1$ . (Note  $x \neq 0$  since  $u_1$  is not in  $F$ ). Therefore  $x_iu_1 + xu_i \in F$ . By our induction hypothesis, since the  $m - 1$  vectors  $x_iu_1 + xu_i$ ,  $2 \leq i \leq m$  are linearly independent in  $F$ , it follows that  $m - 1 \leq n - 1$  and so  $m \leq n$  as required.

Notice that the assertion about linear independence is easy to check since if  $r_2(x_2u_1 + xu_2) + \dots + r_n(x_nu_1 + xu_n) = 0$  then  $(r_2x_2 + \dots + r_nx_n)u_1 + xr_2u_2 + \dots + xr_nu_n = 0$  and so by the linear independence of  $S$ ,  $xr_2 = \dots = xr_n = 0$ . Since  $x \neq 0$ , it follows that  $r_2 = \dots = r_n = 0$  and hence  $r_1 = 0$  also. This completes the argument. ■

To deduce Theorem 3 from Theorem 4, assume that  $R^m \cong R^n$ . Now the standard basis for  $R^m$  is mapped to a linearly independent subset of  $R^n$  and hence  $m \leq n$ . By symmetry, the inverse of the isomorphism maps a basis for  $R^n$  to a linearly independent subset of  $R^m$  and so  $n \leq m$ . Putting these together gives  $m = n$ . We call  $n$  the *rank* or sometimes *free rank* of  $R^n$ .

**Theorem 20** *Suppose that  $M$  is a submodule of  $R^n$ , where  $R$  is a PID. Then  $M$  is a free module of rank at most  $n$ .*

**Proof:** We use a similar approach to Theorem 4, with  $F \cong R^{n-1}$  given by taking vectors with first coordinate equal to 0. The proof is again by induction on  $n$ , starting with the case  $n = 1$ . If  $M$  is a submodule of  $R$ , then  $M$  is just an ideal. By our assumption that  $R$  is a PID, it follows that  $M = Rx$ , is cyclic. If  $x = 0$ , then clearly  $M = \{0\}$  is free of rank 0. If  $x \neq 0$ , the map  $\phi : R \rightarrow Rx$  given by  $\phi(r) = rx$  is an isomorphism, since  $R$  is an integral domain. Hence  $M$  is free of rank 1.

Next suppose the theorem is true in  $F \cong R^{n-1}$ . As before let  $\pi : R^n \rightarrow R$  be the projection onto the first coordinate, so that  $F$  is the kernel of  $\pi$ . By our induction

assumption, the submodule  $M \cap F$  of  $F$  is free of rank at most  $n - 1$ . Moreover  $\pi(M)$  is a submodule of  $R$  and hence free of rank at most 1. If this latter rank is 0, then  $M \subset F$  and the result follows. So we can assume that  $\pi(M)$  has generator  $\pi(u_0)$  for some  $u_0 \in M$ . Let  $u_1, u_2, \dots, u_m$  be a basis for  $M \cap F$ . It suffices to show that  $u_0, u_1, u_2, \dots, u_m$  is a basis for  $M$ .

Assume that  $v \in M$ . Now  $\pi(v) = r\pi(u_0)$  for some  $r \in R$ , since  $\pi(u_0)$  spans  $\pi(M)$ . Hence  $v - ru_0$  is in the kernel of  $\pi$  which is  $F$ . It is clearly also in  $M$  and so is in  $M \cap F$ . Consequently  $v - ru_0 = r_1u_1 + r_2u_2 + \dots + r_mu_m$  and this proves that  $u_0, u_1, u_2, \dots, u_m$  is a generating set for  $M$ .

Finally if  $r_0u_0 + r_1u_1 + r_2u_2 + \dots + r_mu_m = 0$ , applying  $\pi$  shows that  $r_0 = 0$ . But then since  $u_1, u_2, \dots, u_m$  is linearly independent, it follows that  $r_1 = \dots = r_m = 0$  and the result is complete. ■

**Example 25** Let  $R = \mathbb{Z}[x]$  be free of rank 1. Notice the ring is not a PID. Let  $M$  be the module (ideal) generated by  $2, x$ . We have seen previously that this is not a cyclic ideal and in fact it is easy to see it is not a free  $R$ -module of any rank. So we have an example of a submodule of a free module which is not free. To see this, assume  $\{f_1, f_2, \dots, f_k\}$  is a basis for  $M$ . Then  $f_2f_1 - f_1f_2 = 0$  and so the set is not linearly independent, unless it only has a single element, ie  $k = 1$ . So the only free submodules of  $R$  are cyclic.

We are now ready to prove the main structure theorem for finitely generated modules over PIDs.

**Theorem 21** Suppose that  $M$  is a finitely generated module over a ring  $R$  which is a PID. Then  $M \cong R \oplus R \oplus \dots \oplus R/d_1R \oplus R/d_2R \oplus \dots \oplus R/d_kR$ , where  $d_1|d_2|\dots|d_k$  and the elements  $d_1, d_2, \dots, d_k$  of  $R$  are not zero or units and are unique up to associates.

**Proof:** We can choose a finite generating set  $\{u_1, u_2, \dots, u_n\}$  for  $M$ . This can be used to construct a homomorphism  $\phi : R^n \rightarrow M$  where  $\phi(r_1, \dots, r_n) = r_1u_1 + \dots + r_nu_n$ . Clearly  $\phi$  is onto and we let  $K$  denote the kernel. By the first isomorphism theorem,  $M \cong R^n/K$ . Also by Theorem 5,  $K$  is free with rank at most  $n$ . So our strategy is to choose new bases of  $M$  and  $K$  which ‘match’ as closely as possible. In particular, a basis  $f_1, f_2, \dots, f_n$  of  $R^n$  will be found so that the basis of  $K$  is  $f_1, \dots, f_j, d_1f_{j+1}, \dots, d_kf_{k+j}$  where  $k + j \leq n$ . Assuming this, we sketch the completion of the proof.

Define  $\psi : R^n \rightarrow R \oplus \dots \oplus R/d_1R \oplus \dots \oplus R/d_kR$ , where the number of copies of  $R$  is  $n - k - j$ .  $\psi(r_1, r_2, \dots, r_n) = (r_{j+1} + d_1R, \dots, r_{j+k} + d_kR, r_{k+j+1}, \dots, r_n)$ . Clearly  $\psi$  is a well defined and onto homomorphism; it remains only to identify its kernel. But  $\psi(r_1, r_2, \dots, r_n) = 0$  exactly when  $r_{j+1} \in d_1R, \dots, r_{j+k} \in d_kR, r_{k+j+1} = \dots = r_n = 0$ . But these are precisely the vectors corresponding to elements of  $M$  lying in  $K$ , since the latter are linear combinations of  $f_1, \dots, f_j, d_1f_{j+1}, \dots, d_kf_{k+j}$ .

By the first isomorphism theorem applied to  $\psi$ , it follows that  $R^n/K \cong R \oplus \dots \oplus R/d_1R \oplus \dots \oplus R/d_kR$ , where there are  $n - j - k$  copies of  $R$ . So  $M \cong R^n/K$  is isomorphic to this direct sum of cyclic modules as claimed in the theorem.

Let  $\{e_1, e_2, \dots, e_n\}$  be the standard basis for  $R^n$  and let  $\{g_1, \dots, g_{j+k}\}$  be a basis for  $K$ . Each  $g_i$  can be written as a linear combination of the standard basis vectors and so gives

a vector of  $n$  elements of  $R$ . Writing these  $j + k$  vectors as rows, we get a  $(j + k) \times n$  matrix  $A$  with entries in  $R$ .

Now we can change the choice of the basis for  $K$  by interchanging two vectors, multiplying a vector by a unit of  $R$ , or adding a multiple of one vector  $g_i$  to another vector  $g_p$ , where  $i \neq p$ . (Note these operations need to be invertible, so we cannot multiply a basis vector by an element of  $R$  which is not a unit, in general). It is easy to see the effect on  $A$  is to perform precisely the same row operations. Moreover we can do similar changes to the basis for  $R^n$ . These give column operations on  $A$ . So the idea is to see how much simplification of  $A$  can be achieved by such a sequence of row and column operations.

Recall that for a PID  $R$ , it was proved that if two elements  $r, s$  have g.c.d equal to  $d$ , then one can find elements  $x, y$  so that  $xr + ys = d$ . It is easy to extend this to a finite set of elements  $\{r_1, \dots, r_k\}$ . So if  $d$  is the g.c.d of such a set, then we can find elements  $x_1, \dots, x_k$  so that  $x_1r_1 + \dots + x_kr_k = d$ . (Exercise: prove this by induction on  $k$ ). Our aim is to show that if  $d$  is the g.c.d of all the non zero elements of  $A$ , then by row and column operations, it can be achieved that one of the entries of  $A$  is  $d$ . After that, one can easily shift  $d$  to the top left hand entry by switching rows and columns. Finally since  $d$  divides all the non zero entries, we can reduce all the entries in the first row and column to zero except for the  $d$  entry. Deleting the first row and column gives a matrix  $A'$  with one fewer rows and columns and so the argument can be repeated. Note that the g.c.d  $d'$  of all the non zero entries of  $A'$  is divisible by  $d$  and so we can continue till the matrix has zero entries everywhere except along the main diagonal. The first  $j$  entries are units (possibly  $j = 0$ ) and the next  $k$  a sequence  $d_1|d_2|\dots|d_k$  of non units. We have then found a basis of  $R^n$  so that a basis of  $K$  is  $f_1, \dots, f_j, d_1f_{j+1}, \dots, d_kf_{j+k}$ .

To simplify the last step, we assume that  $R$  is an ED. For all our applications, this is the case. Given two non zero elements  $a, b$  in the matrix  $A$ , assume that these are adjacent, by switching rows and columns. By the Euclidean algorithm, either  $a|b$  or we can write  $b = aq + r$  where  $N(r) < N(a)$ . So by performing a row or column operation, either  $b$  can be replaced by 0 or by  $b - aq = r$  with smaller  $N$  value than  $a$ . So by choosing  $d$  as an element with least  $N$  value amongst all matrices which are obtained from  $A$  by row and column operations, we can either get a new element with smaller  $N$  value by doing more row and column operations or  $d$  divides all other non zero elements of  $A$ . Since the former contradicts our choice of  $d$ , it follows that  $d$  is the g.c.d of all non zero entries. ■

The diagonal entries  $\{1, 1, \dots, 1, d_1, d_2, \dots, d_k\}$  are called invariant factors of the module and the matrix with these diagonal entries is said to be the invariant factor matrix.

**Example 26** Find the invariant factor matrix over  $\mathbb{Z}$  for the matrix

$$A = \begin{pmatrix} -4 & -6 & 7 \\ 2 & 2 & 4 \\ 6 & 6 & 15 \end{pmatrix}$$

By subtracting 3 times row 2 from row 3, we get  $\begin{pmatrix} -4 & -6 & 7 \\ 2 & 2 & 4 \\ 0 & 0 & 3 \end{pmatrix}$  Next subtract 2

times row 3 from row 1 giving  $\begin{pmatrix} -4 & -6 & 1 \\ 2 & 2 & 4 \\ 0 & 0 & 3 \end{pmatrix}$ . Next switch columns 1 and 3 to give  $\begin{pmatrix} 1 & -6 & -4 \\ 4 & 2 & 2 \\ 3 & 0 & 0 \end{pmatrix}$ . Subtract row 3 from row 2 giving  $\begin{pmatrix} 1 & -6 & -4 \\ 1 & 2 & 2 \\ 3 & 0 & 0 \end{pmatrix}$ . Next subtract row 1 from row 2 and 3 times row 1 from row 3 giving  $\begin{pmatrix} 1 & -6 & -4 \\ 0 & 8 & 6 \\ 0 & 18 & 12 \end{pmatrix}$ . Next add 6 times column 1 to column 2 and 4 times column 1 to column 3 giving  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 8 & 6 \\ 0 & 18 & 12 \end{pmatrix}$ . This completes the first step. For the remaining  $2 \times 2$  matrix, subtract 2 times row 2 from row 3 and then switch rows 2 and row 3 giving  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 8 & 6 \end{pmatrix}$ . Finally subtract 4 times row 2 from row 3 giving  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix}$ . Note that this example also shows that the Abelian group with generators  $x, y, z$  and relations  $-4x - 6y + 7z = 0, 2x + 2y + 4z = 0, 6x + 6y + 15z = 0$  is isomorphic to the Abelian group  $\mathbb{Z}_2 \oplus \mathbb{Z}_6$ .

We give an example involving polynomials as the entries. So the ring is  $\mathbb{Q}[x]$  of polynomials with rational coefficients.

**Example 27** Find the invariant factor matrix for

$$A = \begin{pmatrix} x & 1 & -2 \\ -3 & x+4 & -6 \\ -2 & 2 & x-3 \end{pmatrix}$$

The first step is to notice, since there is an element 1, the g.c.d will be 1 and we can shift this to the top left hand corner by interchanging columns 1 and 2. This gives  $\begin{pmatrix} 1 & x & -2 \\ x+4 & -3 & -6 \\ 2 & -2 & x-3 \end{pmatrix}$ . The second step is to subtract  $x$  times column 1 from column 2

and add 2 times column 1 to column 3. This gives  $\begin{pmatrix} 1 & 0 & 0 \\ x+4 & -3-4x-x^2 & 2x+2 \\ 2 & -2-2x & x+1 \end{pmatrix}$ .

We can now subtract  $(x+4)$  times row 1 from row 2 and 2 times row 1 from row 3 giving  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -3-4x-x^2 & 2x+2 \\ 0 & -2-2x & x+1 \end{pmatrix}$ . The third step is to analyse the  $2 \times 2$  matrix remaining to identify the g.c.d. Clearly this is  $x+1$  which is one of the terms. So we shift this to the left hand corner by first interchanging columns 2 and 3 giving  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2x+2 & -3-4x-x^2 \\ 0 & x+1 & -2-2x \end{pmatrix}$ .

Now we interchange rows 2 and 3 giving  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & x+1 & -2-2x \\ 0 & 2x+2 & -3-4x-x^2 \end{pmatrix}$ . Finally we can add 2 times column 2 to column 3 giving  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & x+1 & 0 \\ 0 & 2x+2 & 1-x^2 \end{pmatrix}$ . Now subtract 2 times row 2 from row 3 ending with  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & x+1 & 0 \\ 0 & 0 & 1-x^2 \end{pmatrix}$ , which is the invariant factor matrix.

So the  $\mathbb{Q}[x]$  module with generators  $\alpha, \beta, \gamma$  and relations  $x\alpha + \beta - 2\gamma = 0$ ,  $-3\alpha + (x+4)\beta - 6\gamma = 0$ ,  $-2\alpha + 2\beta + (x-3)\gamma = 0$  is isomorphic to  $\mathbb{Q}[x]/(x+1)\mathbb{Q}[x] \oplus \mathbb{Q}[x]/(1-x^2)\mathbb{Q}[x]$ .

To finish this chapter, we make a remark about module isomorphisms and matrices. Here we need that  $R$  is commutative. As usual, given a free  $R$  module  $M$  and an isomorphism  $\phi : M \rightarrow M$ , we can pick any basis for  $M$  and represent  $\phi$  by a square matrix  $A$  relative to this basis. Since  $\phi$  is invertible, so is  $A$  and vice versa; if  $A$  is invertible so is the corresponding homomorphism. Over rings, we must be careful about checking if matrices are invertible. The key idea is to demand that the determinant of  $A$  is a unit in  $R$ , not just a non zero element.

**Lemma 7**  *$A$  is invertible if and only if  $\det A$  is a unit in  $R$ .*

**Proof:** If  $B$  is an inverse of  $A$ , then  $AB = I$ . So  $\det A \det B = 1$  in  $R$  and so  $\det A$  is a unit. Conversely if  $\det A$  is a unit, we can use Cramer's rule to find  $B$ . If  $A = [a_{ij}]$  are the entries of  $A$ , then the cofactor  $\text{cof}_{ij}$  is defined as  $(-1)^{i+j} \det A_{ij}$  where  $A_{ij}$  is the matrix obtained by deleting the  $i$ th row and  $j$ th column of  $A$ . Then  $B = \frac{1}{\det A} [\text{cof}_{ij}]$  defines the inverse of  $A$ . So the  $ij$ th entry of  $B$  is  $\text{cof}_{ij} / \det A$ .

Note that this definition makes sense in a ring  $R$ , so long as  $\det A$  is a unit so has an inverse and the ring is commutative. ■

**Example 28** *Note that if  $R = \mathbb{Z}$ , then an invertible square matrix  $A$  with entries in  $\mathbb{Z}$  must have determinant  $\pm 1$ . All such matrices form a group  $G$ , which is denoted by  $GL(n, \mathbb{Z})$  and is called the general linear group over the integers. Here the matrices are  $n \times n$ .*

April 10

## Summary notes for Algebra 321

### CHAPTER SIX - Applications to abelian groups and linear transformations

By the previous chapter, we have the following result about Abelian groups.

**Theorem 22** *Let  $A$  be a finitely generated Abelian group. Then  $A \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \dots \mathbb{Z} \oplus \mathbb{Z}/n_1\mathbb{Z} \oplus \mathbb{Z}/n_2\mathbb{Z} \oplus \dots \oplus \mathbb{Z}/n_k\mathbb{Z}$ , where  $n_1|n_2|\dots|n_k$  and the integers  $n_1, n_2, \dots, n_k$  can be chosen to be positive,  $> 1$  and are then unique.*

The integers  $n_1, n_2, \dots, n_k$  are called the torsion invariants of  $A$  since the torsion subgroup of  $A$  is clearly  $\mathbb{Z}/n_1\mathbb{Z} \oplus \mathbb{Z}/n_2\mathbb{Z} \oplus \dots \oplus \mathbb{Z}/n_k\mathbb{Z}$ . So for example, this subgroup has order  $n_1 n_2 \dots n_k$ . The number of copies of  $\mathbb{Z}$  is called the torsion-free rank of  $A$ .

For some purposes, it is useful to further decompose the cyclic subgroups  $\mathbb{Z}/n_i\mathbb{Z}$  into cyclic subgroups of prime power order. To do this, we require an important result;

**Lemma 8** *Let  $R$  be any PID and let  $a = bc$  where  $g.c.d$  of  $b, c$  is 1. Then  $R/aR \cong R/bR \oplus R/cR$ .*

**Proof:** Define  $\phi : R/aR \rightarrow R/bR \oplus R/cR$  by  $\phi(r + aR) = (r + bR, r + cR)$ . We need to show this is a well defined one-to-one and onto homomorphism. Since  $b|a$  and  $c|a$ , it is easy to see that if  $r + aR = r' + aR$ , then  $a|(r - r')$  and so  $b|(r - r')$  and  $c|(r - r')$ . So  $r + bR = r' + bR$  and  $r + cR = r' + cR$ . So the map is well defined. To show it is one-to-one, suppose that  $\phi(r + aR) = 0$ . Then  $r + bR = bR$  and  $r + cR = cR$ . So  $b|r$  and  $c|r$ . But then since we are in a PID, which is a UFD, it is easy to see that  $bc|r$ . Hence  $a|r$  and  $r + aR = aR$ . This shows that the kernel of  $\phi$  is zero and  $\phi$  is one-to-one. Finally to prove that  $\phi$  is onto, we need to find  $r$  so that  $r + bR = i + bR$  and  $r + cR = j + cR$  for any pair  $i, j$  in  $R$ . Hence  $r = i + xb$  for some  $x$ . So we need  $i + xb = j + yc$  for some element  $y$ . But this is easy, we know that there is a solution  $ub + vc = 1$  in a PID, for relatively prime elements  $b, c$ . So multiplying by  $i - j$  gives  $i - j = -xb + yc$  where  $x = (j - i)u$  and  $y = (i - j)v$ . So  $i + xb = j + yc$  as required.

**Corollary 5** *If  $m, n$  are relatively prime integers  $> 1$ , (i.e  $g.c.d(m, n) = 1$ ) then  $\mathbb{Z}_{mn} \cong \mathbb{Z}_m \oplus \mathbb{Z}_n$ .*

Using this result, in theorem 1 we can further split up each cyclic summand  $\mathbb{Z}/n_i\mathbb{Z}$  into summands  $\mathbb{Z}_{p^j} \oplus \mathbb{Z}_{q^l} \dots$  by factorising  $n_i = p^j q^l \dots$  into prime powers.

**Theorem 23** *Let  $A$  be a finitely generated Abelian group. Then  $A \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \dots \mathbb{Z} \oplus \mathbb{Z}_{p_1^{m_{11}}} \oplus \mathbb{Z}_{p_1^{m_{12}}} \oplus \dots \oplus \mathbb{Z}_{p_1^{m_{1q}}} \oplus \dots \mathbb{Z}_{p_s^{m_{st}}}$ , where  $1 \leq m_{11} \leq m_{12} \leq \dots \leq m_{1q}$  etc.*

The numbers  $m_{11}, m_{12}, \dots, m_{1q}$  are called the  $p_1$ -primary exponents of  $A$ , so  $m_{21}, m_{22}, \dots$  are the  $p_2$ -primary exponents etc. The numbers  $p_1^{m_{11}}$  etc are called the primary invariants. Notice it is very easy to determine the invariant factors (torsion invariants) if we know the primary invariants.

**Example 29**  $A \cong \mathbb{Z}_{20} \oplus \mathbb{Z}_{30}$ . Then decomposing into primary invariant factors, we get  $A \cong \mathbb{Z}_{2^2} \oplus \mathbb{Z}_5 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_5$ . This can be rearranged into  $A \cong \mathbb{Z}_2 \oplus \mathbb{Z}_{2^2} \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_5 \oplus \mathbb{Z}_5$  as primary invariant form. So the primary invariants are 2,  $2^2$  with 2-primary exponents 1, 2, then 3 with 3-primary exponent 1, finally 5, 5 with 5-primary exponents 1, 1.

Now to determine the torsion invariants (invariant factors), we take the maximum primary invariants for each prime. For this example, these will be  $2^2, 3, 5$ . We then form the cyclic group  $\mathbb{Z}_{2^2 \times 3 \times 5}$ . The next maximum primary invariants are combined to give 2, 5 and  $\mathbb{Z}_{2 \times 5}$ . So we conclude  $A \cong \mathbb{Z}_{60} \oplus \mathbb{Z}_{10}$  is the invariant factor decomposition. Note this method means the order of each cyclic group is divisible by all the following ones. So the order of the factors is reversed.

**Example 30** Determine all the Abelian groups of order 108 and their torsion invariants.

Since  $108 = 2^2 \times 3^3$ , the choices for 2-primary exponents are 1, 1 or 2. The choices for 3-primary exponents are 1, 1, 1 or 1, 2 or 3. So there are 6 possible groups, multiplying together possibilities. These are;

$$\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_3$$

$$\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_{3^2}$$

$$\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_{3^3}$$

$$\mathbb{Z}_{2^2} \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_3$$

$$\mathbb{Z}_{2^2} \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_{3^2}$$

$$\mathbb{Z}_{2^2} \oplus \mathbb{Z}_{3^3}$$

The corresponding torsion invariants and decompositions are;

$$6, 6, 3, \quad \mathbb{Z}_6 \oplus \mathbb{Z}_6 \oplus \mathbb{Z}_3$$

$$18, 6, \quad \mathbb{Z}_{18} \oplus \mathbb{Z}_6$$

$$54, 2, \quad \mathbb{Z}_{54} \oplus \mathbb{Z}_2$$

$$12, 3, 3, \quad \mathbb{Z}_{12} \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_3$$

$$36, 3, \quad \mathbb{Z}_{36} \oplus \mathbb{Z}_3$$

$$108, \quad \mathbb{Z}_{108}$$

Often Abelian groups are given by presentations. So for example, we could be given  $A$  is an Abelian group with presentation  $\langle a, b, c | 2a + 4b = 4b - 8c = 14a = 0 \rangle$ . To determine the structure of this Abelian group, we note that the relations which are linear combinations of the generators equal to zero, can be viewed as generators of the kernel of the homomorphism from  $\mathbb{Z}^3$  to  $A$ , mapping the standard basis of  $\mathbb{Z}^3$  to the chosen

generators  $a, b, c$ . So we get a relation matrix  $\begin{pmatrix} 2 & 4 & 0 \\ 0 & 4 & -8 \\ 14 & 0 & 0 \end{pmatrix}$  with rows corresponding

to the 3 relations. As usual we can perform row operations converting this to diagonal

form. This is  $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -56 \end{pmatrix}$  We can multiply the last row by the unit -1 and convert

all terms to be positive, ie  $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 56 \end{pmatrix}$ . The conclusion is that the Abelian group

$A \cong \mathbb{Z}_2 \oplus \mathbb{Z}_4 \oplus \mathbb{Z}_{56}$ . Alternatively in primary invariant form,  $A \cong \mathbb{Z}_2 \oplus \mathbb{Z}_{2^2} \oplus \mathbb{Z}_{2^3} \oplus \mathbb{Z}_7$ .

Next, we want to use module theory to describe canonical forms for linear transformations and matrices. There are several different canonical forms, but we will just do Jordan canonical form.

Let  $\alpha : V \rightarrow V$  be a linear transformation, where  $V$  is a finite dimensional vector space over  $\mathbb{C}$ . We want to use complex numbers to make sure that the number of eigenvalues (counted with multiplicity) for  $\alpha$  is the same as the dimension of  $V$ . This will come out as a consequence of our theory! As usual, by choosing a basis for  $V$  we get  $V \cong \mathbb{C}^n$  where  $n$  is the number of vectors in the basis, ie the dimension of  $V$ . We would like to use our module theory to choose a ‘better’ basis for  $V$ . To do this, consider  $V$  as a  $\mathbb{C}[x]$  module, where a polynomial  $f(x) \in \mathbb{C}[x]$  acts on a vector  $v \in V$ , by  $f(x)v = f(\alpha)v$ .

Now it is obvious that  $V$  is finitely generated as a  $\mathbb{C}[x]$  module, since it has a finite basis just as a vector space. So by the previous chapter, there is an invariant factor decomposition into cyclic submodules. Now  $\mathbb{C}[x]$  is an infinite dimensional vector space, since  $1, x, x^2, x^3, \dots$  are linearly independent. So we cannot get any copies of  $\mathbb{C}[x]$  in our decomposition of  $V$ . It follows therefore that  $V \cong \mathbb{C}[x]/d_1\mathbb{C}[x] \oplus \mathbb{C}[x]/d_2\mathbb{C}[x] \oplus \dots \mathbb{C}[x]/d_k\mathbb{C}[x]$ , where  $d_1|d_2|\dots|d_k$  are non constant polynomials. (Any constant polynomial is a unit, so the corresponding quotient would be zero).

We can then further decompose these cyclic modules  $\mathbb{C}[x]/d_i\mathbb{C}[x]$  into p-primary components, ie quotients by powers of prime (ie irreducible) polynomials. Now over  $\mathbb{C}$ , the only irreducible polynomials are linear, since any polynomial splits into linear factors. We can also arrange all our polynomials to be monic, since any complex number is a unit so we can always multiply by such a number. So a p-primary component will look like  $W = \mathbb{C}[x]/(x - \lambda)^m\mathbb{C}[x]$ . We can view  $W$  as a subspace of  $V$  as well as a submodule. We would like to pick a clever basis for  $W$  so that the action of  $\alpha$  on  $W$  is transformed into a familiar matrix.

Notice that the dimension of  $W$  is  $m$ , by the Euclidean algorithm. For given any coset  $f(x) + (x - \lambda)^m\mathbb{C}[x]$ , we can divide  $f(x)$  by  $(x - \lambda)^m$  to get  $f(x) = q(x)(x - \lambda)^m + r(x)$ , where the degree of  $r(x)$  is  $< m$ . Then  $f(x) + (x - \lambda)^m\mathbb{C}[x] = r(x) + (x - \lambda)^m\mathbb{C}[x]$ . So this shows that  $1 + (x - \lambda)^m\mathbb{C}[x], x + (x - \lambda)^m\mathbb{C}[x], \dots, x^{m-1} + (x - \lambda)^m\mathbb{C}[x]$  is a generating set for  $W$ . On the other hand, a linear relation between the elements of this generating set, would imply that some polynomial  $r(x)$  with degree  $< m$  satisfies  $r(x) + (x - \lambda)^m\mathbb{C}[x] = (x - \lambda)^m\mathbb{C}[x]$ . But this implies  $r(x) \in (x - \lambda)^m\mathbb{C}[x]$ , or equivalently  $(x - \lambda)^m|r(x)$ . This is impossible unless  $r(x) = 0$ .

Next we choose a better basis for  $W$ , namely  $1 + (x - \lambda)^m\mathbb{C}[x], x - \lambda + (x - \lambda)^m\mathbb{C}[x], \dots, (x - \lambda)^{m-1} + (x - \lambda)^m\mathbb{C}[x]$ . Certainly it suffices to show these elements are linearly independent, as we are in a subspace of dimension  $m$ . But this is easy, a linear relation will again give a polynomial of degree  $< m$  as a multiple of  $(x - \lambda)^m$  and this can only happen if the polynomial is zero. Moreover, since the only coset containing  $x^{m-1}$  in its leading term is the last, we see that the coefficient of this coset must be zero. Similarly all the coefficients are zero.

Finally it is easy to compute the matrix of  $\alpha$  relative to this basis. We illustrate with the case  $m = 3$ . Let  $w_1, w_2, w_3$  denote the basis elements  $1 + (x - \lambda)^3\mathbb{C}[x], x - \lambda + (x - \lambda)^3\mathbb{C}[x], (x - \lambda)^2 + (x - \lambda)^3\mathbb{C}[x]$ . Notice that the action of  $\alpha$  on this cyclic module is by definition multiplication by  $x$  on the cosets. So

$$\alpha(w_1) = x + (x - \lambda)^3\mathbb{C}[x] = \lambda w_1 + w_2,$$

$$\alpha(w_2) = x(x - \lambda) + (x - \lambda)^3 \mathbb{C}[x] = w_3 + \lambda w_2$$

$$\alpha(w_3) = x(x - \lambda)^2 + (x - \lambda)^3 \mathbb{C}[x] = \lambda w_3.$$

Note the last equation uses the fact that  $(x - \lambda)^3 \in (x - \lambda)^3 \mathbb{C}[x]$ . We see therefore that the matrix of  $\alpha$  relative to this basis of  $W$  is

$$\begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}$$

This is called an elementary Jordan matrix - note that  $\lambda$  is just an eigenvalue of  $\alpha$ . We see that the primary decomposition of  $V$  as a  $\mathbb{C}[x]$ -module with this choice of basis yields the Jordan canonical form matrix of  $\alpha$ . The primary exponents corresponding to the prime  $x - \lambda$  give the sizes of the elementary Jordan blocks where this eigenvalue  $\lambda$  occurs.

### Corollary 6 Cayley Hamilton theorem

A matrix  $B$  for  $\alpha$  satisfies the characteristic polynomial  $\det(B - \lambda I) = 0$ .

As an example, if  $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , the eigenvalue equation is

$$\begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = 0.$$

So  $\lambda^2 - (a + d)\lambda + (ad - bc) = 0$ . Then one can check easily that  $B^2 - (a + d)B + (ad - bc)I = 0$ .

Now the proof of the Cayley Hamilton theorem follows readily from our module decomposition picture. Notice that  $\alpha$  on  $W$  satisfies  $(\alpha - \lambda)^m = 0$ , since for any element  $r(x) + (x - \lambda)^m \mathbb{C}[x]$  in  $W$ , the action of  $(\alpha - \lambda)^m$  is multiplication by  $(x - \lambda)^m$ . But  $r(x)(x - \lambda)^m$  is in  $(x - \lambda)^m \mathbb{C}[x]$  so the action is to map every such element to zero.

We see then that if we take the largest primary exponent  $m_i$  for every choice of  $\lambda_i$  and multiply the polynomials  $(x - \lambda_i)^{m_i}$  together, this gives a polynomial  $f(x)$  for which  $\alpha$  is a root. It is easy to see this polynomial divides the characteristic polynomial, which is the product of all the primary factors for all the primary exponents, not just the largest. The polynomial  $f(x)$  is called the minimum polynomial for  $\alpha$ , since it is not hard to show it has the smallest degree of any polynomial with  $\alpha$  as a root and in fact divides every such polynomial. (All such polynomials form a principal ideal in  $\mathbb{C}[x]$  and  $f(x)$  is a generator.)

**Example 31** Suppose that the primary factors for  $\alpha$  are  $(x - \lambda_1)^2, (x - \lambda_1)^3, (x - \lambda_2)^1, (x - \lambda_2)^1, (x - \lambda_2)^4$ . Then the minimum polynomial for  $\alpha$  is  $(x - \lambda_1)^3(x - \lambda_2)^4$ .

## Summary notes for Algebra 321

### CHAPTER SEVEN - Field extensions

If  $F \subseteq K$  are fields, then  $K$  is said to be an extension field of  $F$ . We may view  $K$  as a vector space over  $F$  and the dimension of this vector space is denoted by  $[K : F]$ . This is called the degree of the extension.

The subfield of  $K$  generated by  $F$  and  $u$ , where  $u \in K$  is denoted by  $F(u)$ . It is formed by taking the smallest subfield of  $K$  containing  $u$  and  $F$ . Since it is easy to check that the intersection of subfields is a subfield ( a subfield contains 0, 1 and is closed under addition, subtraction, multiplication and inverses), then  $F(u)$  is the intersection of all subfields of  $K$  containing both  $F$  and  $u$ . Another more useful characterisation of  $F(u)$  is all rational functions in  $u$  with coefficients in  $F$ . A rational function is the quotient of two polynomials, ie of the form  $\frac{p(u)}{q(u)}$ , where  $p(u) = a_0 + a_1u + \dots + a_ku^k$ , and all  $a_i \in F$  and similarly for  $q(u)$ . Notice it is easy to see that this set of rational functions forms a field containing  $F, u$ , since it is closed under addition, subtraction, multiplication and inverses. On the other hand any subfield containing  $F, u$  will contain all polynomials in  $u$  with coefficients in  $F$  by closure under addition and multiplication and so will contain all rational functions because of closure under inverses and multiplication. So this gives a more useful description of  $F(u)$ .

**Example 32** A very important class of examples is of the form  $F(u)$  where  $F = \mathbb{Q}$  and  $u \in \mathbb{C}$ . So here  $K = \mathbb{C}$ . We consider various types of choices for  $u$ .

1)  $u = \sqrt{2}$ . In this case, polynomials in  $u$  with rational coefficients are all linear, since  $u^{2n} = 2^n$  and  $u^{2n+1} = 2^n u$ . Also  $\frac{1}{a+bu} = \frac{a-bu}{a^2-2b^2}$  and so quotients of linear polynomials in  $u$  are again linear polynomials in  $u$ . We conclude that  $\mathbb{Q}(u) = \mathbb{Q}[u]$ , ie the ring of polynomials in  $u$  and this is the same as the linear polynomials in  $u$ . Clearly then  $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$ , since a basis for  $\mathbb{Q}(\sqrt{2})$  as a vector space of  $\mathbb{Q}$  is  $1, \sqrt{2}$ .

2) A very easy example is where  $u = i = \sqrt{-1}$ . This is easier than 1) and we see immediately that  $\mathbb{Q}(i)$  is the set of expressions  $a + bi$  where  $a, b \in \mathbb{Q}$ . So  $[\mathbb{Q}(i) : \mathbb{Q}] = 2$  again.

3)  $u = 2^{\frac{1}{3}}$ . This is somewhat more complicated to analyse the same way as in 1). Certainly a polynomial in  $u$  can be written as a quadratic expression, using  $u^3 = 2$ . So we need only consider ratios of such quadratics. However we need to understand how to compute the inverse of a quadratic and this gets messy. Instead we use our previously developed ring theory to show that the inverse of a quadratic in  $u$  is a quadratic in  $u$  again, but dont give an explicit formula. Notice that the polynomial  $x^3 - 2$  is irreducible in  $\mathbb{Q}[x]$  by Eisenstein. So we know that the principal ideal  $I = (x^3 - 2)\mathbb{Q}[x]$  is maximal and  $\mathbb{Q}[x]/I$  is a field. We claim this is isomorphic to  $\mathbb{Q}(u)$ . The reason is we can define a map  $\phi : \mathbb{Q}(u) \rightarrow \mathbb{Q}[x]/I$  by  $\phi(f(u)/g(u)) = (f(x) + I)(g(x) + I)^{-1}$ . Now it is easy to see this is a well defined onto homomorphism and has kernel 0.

So the construction of inverses in  $\mathbb{Q}[x]/I$  gives the method for constructing inverses in  $\mathbb{Q}(u)$ . Namely by the Euclidean algorithm, if  $f(x)$  is a quadratic polynomial in  $\mathbb{Q}[x]$ , we can find polynomials  $a(x)$  and  $b(x)$  so that  $a(x)(x^3 - 2) + b(x)f(x) = 1$ . But then  $b(u)f(u) = 1$ , substituting in  $u = 2^{\frac{1}{3}}$  for  $x$ .

4) A quite different example is to consider  $u = \pi$ . By a deep result in number theory,  $\pi$  is transcendental, ie does not satisfy any polynomial equation with rational coefficients. Therefore  $f(\pi) \neq g(\pi)$  where  $f(x), g(x) \in \mathbb{Q}[x]$  are different polynomials. We claim that then any two rational functions  $\frac{f(\pi)}{g(\pi)}$  and  $\frac{r(\pi)}{s(\pi)}$  are different, so long as  $g.c.d.f, g = g.c.d.r, s = 1$ . The reason is that if these quotients are the same, then multiplying out gives  $f(\pi)s(\pi) - g(\pi)r(\pi) = 0$ . So since  $\pi$  is transcendental, it follows that  $f(x)s(x) - g(x)r(x) = 0$ . But then by unique factorisation of polynomials, this can only happen if  $f(x) = r(x)$  and  $g(x) = s(x)$ . We see that the field  $\mathbb{Q}(\pi)$  is isomorphic to the abstract field  $\mathbb{Q}(x)$  of rational functions in one variable  $x$  of quotients of polynomials with coefficients in  $\mathbb{Q}$ . The same is true for any other transcendental number, such as  $e$ .

5) Finite fields are not contained naturally in some large field like  $\mathbb{C}$ . However the same extension method of taking polynomials and constructing inverses as in 2) and 3) above, works well. Take  $F = \mathbb{Z}_2$  and  $u$  a root of the irreducible quadratic polynomial  $x^2 + x + 1$ . Note that there is only one irreducible polynomial in  $\mathbb{Z}_2[x]$ . Any polynomial in  $\mathbb{Z}[u]$  can be written as a linear polynomial, since  $u^2 = u + 1$  means that higher powers of  $u$  can be written as lower powers. Hence  $\mathbb{Z}[u] = \{\bar{0}, \bar{1}, u, u + \bar{1}\}$ , using  $\bar{i}$  to represent an element of  $\mathbb{Z}_2$ , ie a congruence class modulo 2. Finally we see that  $u(u + \bar{1}) = \bar{1}$  and so the elements  $u, u + \bar{1}$  are inverses of each element so  $\mathbb{Z}[u]$  is a field with 4 elements.

**Definition:** An element  $u$  in an extension field  $K$  over a subfield  $F$  is called *algebraic* over  $F$  if there is some non zero polynomial  $f(x) \in F[x]$  which has  $u$  as a root. If there is no such polynomial then  $u$  is called transcendental over  $F$ .

Notice that if  $[K : F]$  is finite, then every element  $u$  in  $K$  is algebraic. The reason is that the powers  $1, u, u^2, \dots$  cant be linearly independent, in fact if  $d$  is the degree of the extension then  $1, u, u^2, \dots, u^d$  must be linearly dependent. So this gives some polynomial with  $u$  as a root.

**Theorem 24** *Every algebraic element is a root of a unique monic irreducible polynomial called its minimal polynomial.*

**Proof:** This is easy; consider the map  $\phi : F[x] \rightarrow F[u]$ , which takes  $f(x)$  to  $f(u)$ , where  $u \in K$  is algebraic over the subfield  $F$ . Then the kernel  $I$  of  $\phi$  is the ideal of all polynomials with  $u$  as a root. Since  $I$  is an ideal and the ring  $F[x]$  is a PID, it follows that  $I = \langle g(x) \rangle$ , for some unique monic polynomial  $g$ . ( Any two generators of  $I$  must be associates, ie divide each other. So choosing a monic generator makes it unique.

We claim that  $g$  is irreducible. For if  $g = g_1g_2$ , where neither  $g_i$  is constant, then  $0 = g(u) = g_1(u)g_2(u)$ . Hence either  $g_1(u) = 0$  or  $g_2(u) = 0$ . But then either factor would be in  $I$  and so a multiple of  $g$ , which contradicts the fact they have smaller degrees than  $g$ . So  $g$  is the minimal polynomial sought. ■

**Example 33** *Consider the element  $u = \sqrt{2} + \sqrt{3}$ . To show this is an algebraic element over  $\mathbb{Q}$ , we take powers  $1, u, u^2, \dots$  and seek the smallest number giving a linear relation. Now  $u^2 = 5 + 2\sqrt{6}$ ,  $u^3 = 11\sqrt{2} + 9\sqrt{3}$ ,  $u^4 = 49 + 20\sqrt{6}$ . So it is easy to see that  $u^4 - 10u^2 + 1 = 0$  and there are no 'obvious' linear dependences between fewer powers of*

$u$ . So we expect that  $g(x) = x^4 - 10x^2 + 1$  should be the minimal polynomial of  $u$ . This would imply that  $u$  is indeed algebraic and that  $\mathbb{Q}[u]$  is a degree 4 extension of  $\mathbb{Q}$  with basis  $1, u, u^2, u^3$ . One way to make this rigorous is to note that this field has another basis given by  $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$ . It is clear that this set spans  $\mathbb{Q}[u]$ . So we need only prove these elements are linearly independent. Suppose that  $n_1 + n_2\sqrt{2} + n_3\sqrt{3} + n_4\sqrt{6} = 0$ . Note by multiplying through by a suitable integer, we can arrange that each  $n_i$  is an integer, rather than a rational number. Now write this as  $n_1 + n_2\sqrt{2} = -n_3\sqrt{3} - n_4\sqrt{6}$ . Multiply both sides by  $n_1 - n_2\sqrt{2}$ . The left side becomes an integer so the right side must also. But the right side is  $n_5\sqrt{3} + n_6\sqrt{6}$  for some integers  $n_5, n_6$ . Squaring both sides again, we get that  $6n_5n_6\sqrt{2}$  is an integer, which is impossible unless  $n_5 = 0$  or  $n_6 = 0$ . But this would imply that an integer multiple of  $\sqrt{3}$  or  $\sqrt{6}$  is an integer so  $n_5 = n_6 = 0$ . Working backwards, we get  $n_1 = n_2 = 0$  and so  $n_3 = n_4 = 0$  as well. Hence this is a basis, by this direct method. Usually we want a better method as this gets tedious!

For example 2, it is easy to check that the roots of  $g(x)$  are all in the field  $K = \mathbb{Q}[u]$ . In fact,  $g(x) = (x - \sqrt{2} - \sqrt{3})(x + \sqrt{2} - \sqrt{3})(x - \sqrt{2} + \sqrt{3})(x + \sqrt{2} + \sqrt{3})$  in  $K$ .

**Definition:** A splitting field for an irreducible polynomial  $g(x)$  with coefficients in a field  $F$  is defined as the field generated by all the roots  $\alpha_1, \dots, \alpha_k$  of  $g(x)$ . So  $K = F[\alpha_1, \dots, \alpha_k]$ . Splitting fields are often also called Galois extensions.

We will see in the next chapter, that such a splitting field has a group of symmetries  $G$  called the Galois group of  $K$  over  $F$ . In particular, a symmetry of  $K$  over  $F$  is a bijection  $\phi : K \rightarrow K$  which is a field homomorphism, ie  $\phi(u + v) = \phi(u) + \phi(v)$  and  $\phi(uv) = \phi(u)\phi(v)$ . Moreover we will require that  $\phi(f) = f$  for all  $f \in F$ . It is easy to show  $G$  is indeed a group. More difficult will be to show that the order  $|G|$  is the same as the degree of the extension  $[K : F]$ . For example 2, notice that any such a symmetry will permute the roots of  $g(x)$ . The reason is that  $\phi(r) = r$  for any rational number and so  $\phi(g(v)) = g(\phi(v))$  for any element  $v \in K$ . In particular, if  $v$  is a root of  $g(x)$ , then so is  $\phi(v)$ . Using this, it can be shown that the symmetries of  $K$  are  $I, \phi_1, \phi_2, \phi_3$  where

$$\begin{aligned}\phi_1(\sqrt{2}) &= -\sqrt{2}, & \phi_1(\sqrt{3}) &= \sqrt{3} \\ \phi_2(\sqrt{2}) &= \sqrt{2}, & \phi_2(\sqrt{3}) &= -\sqrt{3} \\ \phi_3(\sqrt{2}) &= -\sqrt{2}, & \phi_3(\sqrt{3}) &= -\sqrt{3}\end{aligned}$$

Then  $\phi_1\phi_2 = \phi_2\phi_1 = \phi_3$  and it is easy to see that  $G \cong \mathbf{Z}_2 \oplus \mathbf{Z}_2$ .

A key result is to show how degrees of multiple extensions are computed.

**Theorem 25** Suppose that  $u, v$  are algebraic elements over  $F$ . Let  $u$  have minimal polynomial  $f(x)$  over  $F$  and let  $v$  have minimal polynomial  $g(x)$  over  $F[u]$ . Then if  $k$  is the degree of  $f$  and  $m$  is the degree of  $g$ , we have  $[F[u, v] : F] = [F[u, v] : F[u]][F[u] : F]$ , ie the degree of  $F[u, v]$  over  $F$  is  $km$ .

**Proof:** Let  $b_1, \dots, b_k$  be a basis for  $F[u]$  as a vector space over  $F$  and let  $c_1, \dots, c_m$  be a basis for  $F[u, v]$  over  $F[u]$ . (We can take  $b_1 = 1, \dots, b_k = u^{k-1}$ ,  $c_1 = 1, \dots, c_m = v^{m-1}$ , but this is not needed. We claim that the set of elements  $b_i c_j$  for  $1 \leq i \leq k, 1 \leq j \leq m$  is a basis for  $F[u, v]$  over  $F$ .

Firstly we show this set spans the vector space  $F[u, v]$ . By definition, given  $u \in F[u, v]$ ,  $u = \lambda_1 c_1 + \dots + \lambda_m c_m$ , where the  $\lambda_j \in F[u]$ . So each  $\lambda_j = \mu_{j1} b_1 + \dots + \mu_{jk} b_k$ , where every  $\mu_{ji} \in F$ . Substituting this in, we obtain  $u = \sum_{i,j} \mu_{ji} b_i c_j$ . So we see that the set  $\{b_i c_j\}$  spans  $F[u, v]$  over  $F$ .

Next we have to show this set is linearly independent over  $F$ . So suppose that  $0 = \sum_{i,j} \mu_{ji} b_i c_j$  where every  $\mu_{ji} \in F$ . We can write this equation as  $0 = \lambda_1 c_1 + \dots + \lambda_m c_m$ , where  $\lambda_j = \mu_{j1} b_1 + \dots + \mu_{jk} b_k$ . Now since  $c_1, \dots, c_m$  are linearly independent over  $F[u]$ , it follows that  $0 = \lambda_1 = \dots = \lambda_m$ . So  $0 = \lambda_j = \mu_{j1} b_1 + \dots + \mu_{jk} b_k$  implies all  $\mu_{jk}$  as  $b_1, \dots, b_k$  are linearly independent over  $F$ . This completes the proof. ■

### Remarks

Notice that we can now estimate degrees of extensions like splitting fields. For if  $f(x)$  is an irreducible polynomial of degree  $n$  over  $F$ , adjoining the first root  $u_1$  to  $F$  to form  $F[u_1]$  gives an extension of degree  $n$ . In this extension certainly  $f(x) = g(x)(x - u_1)$  where  $g(x)$  has degree  $n - 1$ . If  $g(x)$  is irreducible, adjoining a root  $u_2$  gives an extension  $F[u_1, u_2]$  of degree  $n(n - 1)$  over  $F$ . If  $g(x)$  is not irreducible, then  $u_2$  satisfies a minimal polynomial of smaller degree  $d < n - 1$  and we see  $[F[u_1, u_2] : F] = nd$ . Continuing on we find;

**Corollary 7** *If  $f(x)$  is an irreducible polynomial of degree  $n$  over  $F$  then a splitting field for  $f(x)$  has degree at most  $n!$  over  $F$ .*

**Corollary 8** *If  $f(x)$  is an irreducible polynomial of degree  $n$  over  $F$  then the Galois group of a splitting field of  $f(x)$  over  $F$  has at most  $n!$  elements.*

Notice this corresponds nicely to the idea ( see later) that the Galois group is isomorphic to a subgroup of permutations of the roots of  $f(x)$ .

**Corollary 9** *If  $\{u_1, u_2, \dots, u_k\}$  are algebraic elements over a field  $F$ , then  $F[u_1, \dots, u_k]$  is an algebraic extension of  $F$  of degree at most  $d_1 d_2 \dots d_k$  where  $d_i$  is the degree of the minimal polynomial of  $u_i$  over  $F$ .*

### Remarks

In general it is hard to work out exactly the degree of such extensions. The problem is that each time we add some elements  $u_1, \dots, u_{i-1}$  to  $F$  to generate a field, we have to find the minimal polynomial for  $u_i$  over this new field. Of course this polynomial divides the minimal polynomial for  $u_i$  over  $F$  so  $d_i$  is an upper bound for the degree.

Notice also that any ‘combination’ of algebraic elements is algebraic. So for instance we can take any polynomial in two algebraic elements  $u$  and  $v$ , such as  $w = u^{11}v^6 - u^4v^8$  and this belongs to the algebraic extension  $F[u, v]$  and so is algebraic. Moreover the degree of the minimal polynomial of  $w$  is at most the degree  $d$  of  $F[u, v]$ , since we know

that  $1, w, w^2, \dots, w^d$  are linearly dependent over  $F$ . So we see that the degree of  $w$  is at most  $d_1 d_2$  where  $d_1$  is the degree of the minimal polynomial of  $u$  over  $F$  and  $d_2$  is similarly for  $v$  over  $F$ .

As an example, we could take  $2^{\frac{1}{5}} + \sqrt{-7}$ . This element must have degree at most  $5 \cdot 2 = 10$  over  $\mathbf{Q}$ , since  $2^{\frac{1}{5}}$  has minimal polynomial  $x^5 - 2$  (irreducible by Eisenstein) and  $\sqrt{-7}$  has minimal polynomial  $x^2 + 7$ .

**Corollary 10** *Suppose that  $F \subset K$  is an algebraic field extension of degree  $d$ . Then any subfield  $H$  of  $K$  which contains  $F$  must have degree  $d_1$  where  $d_1 | d$ .*

**Proof:** By the same method as in theorem 2, it follows that

$$[K : F] = [K : H][H : F],$$

where  $F \subseteq H \subseteq K$ . So it follows that the degree of the extension of  $H$  over  $F$  divides the degree of the extension of  $K$  over  $F$ . ■

Notice that if  $d$  is prime, then there are no subfields of  $K$  containing  $F$  except for  $F$  and  $K$  themselves. This is a very useful fact later on when we consider ruler and compass constructions and Galois theory.

## Summary notes for Algebra 321

### CHAPTER EIGHT - Ruler and compass constructions

Our aim is to first prove a positive result - that the coordinates of all constructible points by ruler and compass form a field and that also we can take square roots in this field.

Let  $\mathcal{P}$  be the collection of constructible points and let  $F$  be all their coordinates. Here we are allowed to perform the following operations.

- Given any two points  $P, Q \in \mathcal{P}$ , we can draw the straight line through  $P, Q$
- We can draw a circle with centre at a point  $P$  of  $\mathcal{P}$  and passing through a second point  $Q$  of  $\mathcal{P}$ .
- We also assume that two points are given; the first is taken as the origin  $(0, 0)$  of the coordinate system and the second as the points  $(1, 0)$ . So the second point determines the  $x$ -axis and the scale, ie measure of units of distance.

The set  $\mathcal{P}$  of constructible points is then obtained by performing repeated operations of drawing lines and circles, where new constructible points are intersection points of these curves.

**Example 34** *A basic construction is that of a perpendicular bisector. So if we are given two points  $A, B$ , we can draw circles  $C_A, C_B$  of radius  $r > |AB|$ , centred at  $A, B$  respectively. These circles will intersect at points  $P_1, P_2$ . Joining these points by a straight lines will give the perpendicular bisector of the interval between  $A$  and  $B$ .*

**Example 35** *Another basic construction is that of a parallelogram, starting with three points  $A, B, C$ . We assume  $AB, AC$  are sides of this parallelogram. Now using the perpendicular bisector construction, we can choose any two points on  $AB$  and can draw a perpendicular line  $L$  (the bisector of these two points) to  $AB$ . Next we want to draw an orthogonal line  $L'$  to  $L$  through  $C$ . Choose a radius  $r$  and a circle centred at  $C$  so that the circle meets  $L$  at two points. Now the perpendicular bisector of these two points is the required line  $L'$ . Finally we see that  $L'$  is parallel to  $AB$ . Similarly we can construct a line through  $B$  parallel to  $AC$ , completing the parallelogram. (Thanks to G. Zhang and A Cheeseman for this suggestion).*

Notice that by example 2, we can now draw a line interval parallel to  $AB$  at the point  $C$ , ie can translate line segments parallelly around in the plane. This enables us to use the compass to measure a distance and then draw a circle of this radius at any constructible point.

Next we show that the set of coordinates  $F$  is a field containing the rationals. Notice that addition and subtraction follow easily from the construction of parallelograms. By orthogonal projection as used in the construction of parallelograms, if we can construct a point  $P = (a, b)$ , then we can also construct  $(a, 0)$  and  $(0, b)$ . Moreover then we can construct the points  $(0, a)$  and  $(b, 0)$  by constructing a 45 degree angle. (It is easy to bisect angles - just take points equidistant from the corner and draw two circles of the same radius from these points. Then the intersection points give a bisecting line for the angle.)

To construct products and quotients, we use similar triangles. For products, assume points  $(1, 0)$ ,  $(b, 0)$  and a point  $(0, a)$  are all constructible. Now the line from  $(1, 0)$  to  $(0, a)$  is parallel to a line through  $(b, 0)$  which meets the  $y$ -axis at the point  $(0, ab)$  by similar triangles.

Similarly we can construct quotients by using constructible points  $(a, 0)$ ,  $(b, 0)$ ,  $(0, 1)$  and constructing a line through  $(a, 0)$  parallel to the line through  $(b, 0)$ ,  $(0, 1)$ . This meets the  $y$ -axis at the point  $(0, a/b)$  by similar triangles. So we conclude that  $F$  is a field.

Finally we want to show that  $F$  is closed under square roots, ie if  $u \in F$  and  $u > 0$  then  $\sqrt{u} \in F$ . To do this, assume that  $A = (1, 0)$ ,  $B = (u + 1, 0)$  are both constructible. By finding the midpoint  $C = (\frac{u+1}{2}, 0)$ , we can form a circle with centre  $C$  and radius  $\frac{u+1}{2}$ . Now construct a line through  $A$  parallel to the  $y$  axis and let  $D$  be the point where the line and circle cross, in the positive quadrant. It is easy to see by similar triangles  $ABD, ODA$  that  $AD$  has length  $\sqrt{u}$ . Here  $O = (0, 0)$ . We claim that this characterises  $F$  uniquely, ie there is a unique smallest subfield of  $\mathbb{R}$  with these properties and all constructible points give coordinates in this subfield.

Our main result in this chapter is that any coordinates of constructible points lie in an extension field of  $\mathbb{Q}$  of degree  $2^k$  for some  $k \geq 0$ . By the previous chapter, this will give a powerful criterion for which figures are constructible, since coordinates of all points must have minimal polynomials over  $\mathbb{Q}$  which have degrees some power of 2.

The idea is to show that the intersections between lines and circles have coordinates satisfying quadratic equations over  $F$ , where  $F$  is the field constructed by using all constructible points already found.

Suppose  $L$  is a line through constructible points  $A = (a, b)$ ,  $B = (c, d)$ . The equation of  $L$  can be written as  $\frac{x-a}{a-c} = \frac{y-b}{b-d}$ . Similarly the equation of a circle  $C$  with centre  $P = (p, q)$  and passing through  $Q = (r, s)$  is given by  $(x-p)^2 + (y-q)^2 = (r-p)^2 + (s-q)^2$ . So the equation of  $L$  becomes  $\alpha x + \beta y = \gamma$  and the equation of  $C$  becomes  $x^2 + \delta x + y^2 + \mu y = \nu$ , where all the coefficients are in  $F$ . So we see that solving for the intersection between two such lines gives a point with coordinates in  $F$ , while solving for the intersection of a line and a circle or two circles gives a quadratic equation with coefficients in  $F$ . Hence this gives a degree 2 field extension. Iterating gives an extension of degree a power of 2 as claimed.

### Remark

It is easy to see that solving a quadratic is the same as adjoining a square root of an element of  $F$ , using the quadratic formula. So we see  $F$  is obtained by adding a sequence of square roots. On the other hand, we have shown above that using ruler and compass we can add any square root of the coordinates of a constructible point. So it follows that ALL constructible points form the unique smallest field found by repeatedly adding square roots, starting from rationals. So elements like  $\sqrt{1 + \sqrt{1 + \sqrt{2}}}$  are in this field.

A beautiful case, due to Gauss is the construction of regular  $n$ -gons, ie polygons with all equal sides, by ruler and compass. So we are required to have  $\cos \frac{2\pi}{n}$  and  $\sin \frac{2\pi}{n}$  constructible, ie  $\exp \frac{2i\pi}{n}$  as a root of an irreducible polynomial of degree  $2^k$  over  $\mathbb{Q}$ . Now

this satisfies  $x^n = 1$ . So we have to find the irreducible factors of this polynomial, which are called cyclotomic polynomials.

As an example, consider  $x^5 - 1 = 0$ . This factorises as  $(x - 1)(x^4 + x^3 + x^2 + x + 1 = 0$ . Now we know by the trick involving substitution  $y + 1$  for  $x$ , that  $\frac{x^5 - 1}{x - 1} = y^4 + 5y^3 + \dots + 5$  is irreducible by Eisenstein's criterion. So we expect that the regular pentagon is constructible. It is easy to find an explicit expression for  $\cos(\frac{2\pi}{5})$  in terms of square roots, giving a ruler and compass construction of the regular pentagon.

The next prime which is of order  $2^k + 1$  is when  $k = 4$ , giving degree 17.

Gauss found the following expression giving an explicit way of finding the regular 17-sided polygon by ruler and compass.

$$16\cos\left(\frac{2\pi}{17}\right) = -1 + \sqrt{17} + \sqrt{34 - 2\sqrt{17}} + \sqrt{68 + 12\sqrt{17} - 2\sqrt{34 - 2\sqrt{17}} + 2\sqrt{17}\sqrt{34 - 2\sqrt{17}} - 16\sqrt{34 + 2\sqrt{17}}}$$

To complete this chapter, we now can show easily that various classical geometrical problems cannot be solved by ruler and compass alone.

### Duplication of the cube.

Given a cube, of side length one, we want to construct a second cube of side length  $x$  so that the second cube has twice the volume of the first one. So we have to solve  $x^3 = 2$ , ie construct a cube of side length  $x = 2^{\frac{1}{3}}$ . since  $x^3 = 2$  is irreducible over  $\mathbb{Q}$ , the degree of the extension field  $\mathbb{Q}[2^{\frac{1}{3}}]$  over  $\mathbb{Q}$  is 3. Hence if any sequence of ruler and compass constructions formed a side of length  $2^{\frac{1}{3}}$ , then we would have a field  $F$  of degree  $2^k$  over  $\mathbb{Q}$  containing a subfield  $\mathbb{Q}[2^{\frac{1}{3}}]$ . But this contradicts our result from the previous chapter that the degree 3 of the subfield should divide the degree  $2^k$  of  $F$ . So duplication cannot be achieved.

### Trisection of an angle

We know that any angle can be bisected by ruler and compass, hence divided into  $2^k$  equal pieces. It is easy to show that some angles can be trisected, such as  $\frac{\pi}{2}$ , since  $\frac{\pi}{6}$  is constructible. However we just need to find one angle which cannot be trisected.  $\frac{\pi}{3}$  turns out to be a good choice. If  $\frac{\pi}{3}$  can be trisected, then it is easy to see that the point  $(\cos\frac{\pi}{9}, \sin\frac{\pi}{9})$  is constructible. Now the first coordinate of this point satisfies the equation  $\frac{1}{2-4x^3-3x}$ , since this just expresses the formula  $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$ , using  $\theta = \frac{\pi}{9}$ . Now putting  $y = 2x$  we get the simpler equation  $1 = y^3 - 3y$ . Substituting in  $y = z + 1$  gives  $z^3 + 3z^2 - 3 = 0$ . This is irreducible by Eisenstein, so the field extension of  $\mathbb{Q}$  containing  $\cos\frac{\pi}{9}$  is a degree 3 extension and cannot lie in any field of degree  $2^k$  over  $\mathbb{Q}$ . Hence trisection cannot be achieved by ruler and compass alone.

### Squaring of the circle

Given a circle of radius 1, we are required to form a square enclosing the same area as the circle. So the side length  $x$  of the square satisfies  $x^2 = \pi$ , ie  $x = \sqrt{\pi}$ . Any

field extension of  $\mathbb{Q}$  containing  $x$  will clearly contain  $\pi$ . But such a field is NOT a finite dimensional vector space over  $\mathbb{Q}$ , since  $\pi$  is transcendental. So no ruler and compass construction can square the circle.

## Summary notes for Algebra 321

### CHAPTER NINE - Galois theory

We start with some general observations about fields, then specialise to studying the automorphisms of fields over the rationals. In the final chapter, solvability of polynomials in terms of radicals (ie  $n$ th roots) is related to the automorphism group of splitting fields of polynomials.

The first idea is the *characteristic* of a field. Any field  $F$  contains a multiplicative identity element denoted 1. We examine the possibilities for the smallest subfield  $K = \langle 1 \rangle$  containing 1. Clearly since  $1 \in K$ , so are all sums  $2 = 1 + 1, 3 = 1 + 1 + 1, \dots$ . Moreover then the negative integers and fractions are also in  $K$ , unless  $1 + 1 + \dots + 1 = 0$  for some smallest number  $n$  of copies of 1. In the latter case, we claim this  $n$  is a prime  $p$ . For if  $n = uv$ , then  $(1 + 1 + \dots + 1)(1 + 1 + \dots + 1) = 0$ , where the first bracket contains  $u$  1s and the second bracket  $v$  1s. But as  $F$  is an integral domain, we must have one or other of the brackets being 0, contradicting our choice of  $n$  as the smallest number of 1s which add to 0. To summarise, we have proved the following important result.

**Theorem 26** *Any field  $F$  either contains an isomorphic copy of the rationals  $\mathbb{Q}$ , or else it contains an isomorphic copy of  $\mathbb{Z}_p$  ( $p$  a prime) as the smallest subfield containing 1.*

**Definition:** In the first case, we say the field has characteristic 0 and in the second, characteristic  $p$ .

Note that an alternative approach would have been to define a ring homomorphism from  $\mathbb{Z}$  to  $F$  by mapping 1 to 1 and so  $n$  to  $n$  copies of 1 added together. If this map is one-to-one, it extends to a homomorphism of field taking  $\mathbb{Q}$  into  $F$  which is one-to-one (but not necessarily onto). If the map has a kernel, the kernel must be a maximal ideal so is generated by a prime  $p$ .

In the remainder of these notes, we always assume the fields under consideration are characteristic 0, ie contain a copy of the rationals. For the next idea, this turns out to be of critical importance.

**Theorem 27** *If  $F$  is a field of characteristic 0 and  $f(x)$  is an irreducible polynomial in  $F[x]$ , then all the roots of  $f(x)$  are distinct.*

**Proof:** This is very important when we are calculating symmetries of the splitting field of such a polynomial, to know that the roots are all different. This result is definitely not true for fields of characteristic  $p$ , so makes life more complicated there.

The trick is to study the derivative  $f'$  of  $f(x)$ ! for polynomials, we can define the derivative entirely formally, without reference to limits by assuming that if  $f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ , then

$$f'(x) = nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \dots + a_1.$$

Form the splitting field  $H = F[\eta_1, \eta_2, \dots, \eta_n]$ , where the  $\eta_i$  are the roots of  $f(x)$ . Then over  $H$ , we have that  $f(x) = (x - \eta_1)(x - \eta_2)\dots(x - \eta_n)$ . Moreover, since the product rule works fine for our ‘formal’ differentiation of polynomials, it is true that  $f'(x) = (x - \eta_2)\dots(x - \eta_n) + (x - \eta_1)(x - \eta_3)\dots(x - \eta_n) + \dots + (x - \eta_1)\dots(x - \eta_{n-1})$ . Now if some root is duplicated, we can order our roots so that  $\eta = \eta_1 = \eta_2$ . But then it is obvious that  $\eta$  is also a root of  $f'(x)$ . But this is a contradiction, since  $f(x)$  is the minimal polynomial of any of its roots over  $F$  so we can't find a smaller degree polynomial with coefficients in  $F$  and  $\eta$  as a root. We conclude that no root is repeated, as claimed.

■

Note this result can fail for characteristic  $p$  when  $p|n$ . For in this case, the derivative can ‘disappear’- it turns out that the bad situation is when all the powers of  $x$  in  $f(x)$  are actually powers of  $x^p$ .

**Definition:** A polynomial  $f(x)$  is *separable* if it has distinct roots in its splitting field.

So theorem 2 states that over a field of characteristic 0, any irreducible polynomial is separable.

Our main aim is to study automorphisms of field extensions. From now on, all fields have characteristic 0.

**Definition:** Given two fields  $E \subseteq F$ , an *automorphism of  $F$  over  $E$*  is a field homomorphism  $\phi : F \rightarrow F$  so that  $\phi$  is one-to-one and is the identity on  $E$ . In other words,  $\phi(u + v) = \phi(u) + \phi(v)$ ,  $\phi(uv) = \phi(u)\phi(v)$  and  $\phi(w) = w$  whenever  $w \in E$ .

We denote by  $Aut(F, E)$  the set of automorphisms of  $F$  over  $E$ . It is very easy now to show that  $Aut(F, E)$  is a group. For if  $\phi, \psi \in Aut(F, E)$ , it is obvious that  $\phi\psi \in Aut(F, E)$  and  $\phi^{-1} \in Aut(F, E)$ . For many extensions,  $Aut(F, E)$  will be the uninteresting group  $I$ , ie the only automorphism will be the identity map on  $F$ . A *Galois* extension will be one where there are lots of automorphisms; in fact the maximal number possible. These are often also called *normal* extensions.

**Definition:** The fixed set  $fix(\phi)$  of an automorphism  $\phi : F \rightarrow F$  is all elements  $u \in F$  so that  $\phi(u) = u$ . The fixed field of a group  $G$  of automorphisms of  $F$  is the elements of  $F$  which are in the fixed set of all members of  $G$ . We denote this by  $fix(G)$ .

**Lemma 9**  $fix(\phi)$  and  $fix(G)$  are subfields of  $F$ .

**Proof:** This is very easy. We just note that if  $\phi(u) = u$  and  $\phi(v) = v$  then  $\phi(u+v) = u+v$  and  $\phi(uv) = uv$ , using the fact that  $\phi$  is an automorphism. Then  $fix(G)$  is just the intersection of all the fixed sets of the elements of  $G$ . But it is easy to see that the intersection of subfields is a subfield and this completes the proof. ■

We can now make the key definition.

**Definition:** A field extension  $E \subseteq F$  is called Galois or normal if the degree  $d$  of the extension is finite and the Galois group  $Aut(F, E)$  has  $d$  elements.

Note it is obvious that  $E \subseteq \text{fixAut}(E, F)$ . If there are ‘not enough’ automorphisms of  $F$ , then this fixed set will be bigger than  $E$ . In fact an extension is Galois if and only if  $E = \text{fixAut}(E, F)$ .

It will turn out that the splitting field of a polynomial ( not necessarily irreducible) is precisely the same as a Galois extension.

Our first result in the direction of trying to find Galois extensions is to show that the order of  $\text{Aut}(F, E)$  is bounded by the degree of the extension of  $E$  by  $F$ .

**Theorem 28** *Suppose that  $E \subseteq F$  is a finite degree extension. Then  $[F : E] \geq |\text{Aut}(F, E)|$ .*

**Proof:** Assume that  $\text{Aut}(F, E)$  has elements  $\sigma_1, \dots, \sigma_n$ . We first prove that these automorphisms are ‘linearly independent’, ie the only choice of  $n$  elements  $a_1, \dots, a_n$  in  $F$  with  $a_1\sigma_1 + \dots + a_n\sigma_n = 0$  is for all  $n$  elements being 0. Notice it makes sense to take linear combinations of automorphisms; the resulting map takes  $F$  to  $F$  but is not an automorphism any longer, but just a linear transformation of  $F$  considered as a vector space over  $E$ .

We assume that none of the  $a_i$  are zero, otherwise we can drop the expression  $a_i\sigma_i$  from our sum. Now choose some elements  $u$  in  $F$  so that  $\sigma_1(u) \neq \sigma_2(u)$ . Such an element must exist since  $\sigma_1 \neq \sigma_2$ . Now we have  $a_1\sigma_1(uv) + \dots + a_n\sigma_n(uv) = 0$ , for every  $v \in F$ . Hence

$$a_1\sigma_1(u)\sigma_1(v) + \dots + a_n\sigma_n(u)\sigma_n(v) = 0 \quad (1)$$

Also  $a_1\sigma_1(v) + \dots + a_n\sigma_n(v) = 0$  can be multiplied by  $\sigma_1(u)$  to give

$$a_1\sigma_1(u)\sigma_1(v) + \dots + a_n\sigma_1(u)\sigma_n(v) = 0 \quad (2)$$

Subtracting (1) from (2) gives a new equation of the form  $b_2\sigma_2(v) + \dots + b_n\sigma_n(v) = 0$  for each  $v \in F$ , where  $b_2, \dots, b_n \in F$  are fixed elements. Continuing on, we can eliminate all except one automorphism, to get  $\sigma_n = 0$  contrary to assumption.

Now to complete the proof, assume that there is a basis  $u_1, \dots, u_m$  for  $F$  as a vector space over  $E$ , where  $m < n$ . Now consider the equations

$$x_1\sigma_1(u_1) + \dots + x_n\sigma_n(u_1) = 0$$

$$x_1\sigma_1(u_2) + \dots + x_n\sigma_n(u_2) = 0$$

.....

$$x_1\sigma_1(u_m) + \dots + x_n\sigma_n(u_m) = 0.$$

These represent  $m$  equations in  $n$  unknowns, where the  $x_i$  are variables in  $F$  and we view this as a collection of homogeneous equations in  $F$  as a vector space, rather than a field. As usual, such a system must have a non zero solution, so we can find a collection of  $x_i$ , not all zero, satisfying all these equations. Now any element  $v$  of  $F$  can be written as a linear combination of the basis elements  $u_i$ , with coefficients in  $E$ . Since each  $\sigma_i$  fixes  $E$ , by taking the appropriate linear combinations of the above system of equations, we end up with

$$x_1\sigma_1(v) + \dots + x_n\sigma_n(v) = 0$$

for all  $v \in F$ . But this contradicts the argument in the previous paragraph, since we have found a linear dependence between the automorphisms, which is impossible. ■

We now look at some examples, to see the difference between Galois extensions and non Galois extensions.

**Example 36** Consider the field extension  $F = \mathbb{Q}[\sqrt{2} + \sqrt{3}]$  over  $E = \mathbb{Q}$ . We saw before that the minimal polynomial  $f(x)$  for  $\sqrt{2} + \sqrt{3}$  is  $x^4 - 10x^2 + 1$ . So this field is a degree 4 extension. Moreover the roots of  $f(x)$  are  $\pm\sqrt{2} \pm \sqrt{3}$ . So  $F$  is a splitting field for  $f(x)$ . We claim that  $\text{Aut}(F, E)$  has 4 elements, so that  $[F : E] = |\text{Aut}(F, E)|$  and this is a Galois extension. Define  $\sigma_1(\sqrt{2}) = \sqrt{2}, \sigma_1(\sqrt{3}) = -\sqrt{3}, \sigma_2(\sqrt{2}) = -\sqrt{2}, \sigma_2(\sqrt{3}) = \sqrt{3}, \sigma_3(\sqrt{2}) = -\sqrt{2}, \sigma_3(\sqrt{3}) = -\sqrt{3}$ . Also the identity transformation is denoted by  $I$ . To show these all define automorphisms of  $F$  fixing  $E = \mathbb{Q}$ , note that  $F$  has basis  $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$  over  $\mathbb{Q}$ . So a linear transformation of  $F$  is defined by giving its action on such a basis. For example,  $\sigma_1$  has action on the first three basis vectors already given. For  $\sqrt{6}$ , we just use  $\sqrt{6} = \sqrt{2}\sqrt{3}$  and the fact that  $\sigma_1$  is an automorphism to see that  $\sigma_1(\sqrt{6}) = -\sqrt{6}$ .

Finally we have to explain why the linear transformation  $\sigma_1$  is really a field automorphism. Now any element of  $F$  can be also written as  $(a + b\sqrt{2})(c + d\sqrt{3})$ , where  $a, b, c, d \in \mathbb{Q}$ .  $\sigma_1$  fixes  $a + b\sqrt{2}$  and maps  $c + d\sqrt{3}$  to  $c - d\sqrt{3}$ . But it is easy to check that  $\sigma_1(c + d\sqrt{3})(c' + d'\sqrt{3}) = (c - d\sqrt{3})(c' - d'\sqrt{3}) = \sigma_1(c + d\sqrt{3})\sigma_1(c' + d'\sqrt{3})$ . Putting this together shows that  $\sigma_1$  is a field automorphism as claimed. The argument for the other maps is analogous.

Next we would like to give an example of a non Galois extension. In this case, we get to a better understanding of properties of automorphisms.

**Example 37** Consider  $F = \mathbb{Q}[2^{\frac{1}{3}}]$  over  $\mathbb{Q}$ . In this case, notice that only one root of the polynomial  $x^3 - 2$  is in  $F$  so this is not a splitting field. The degree of the extension is 3, so we know the number of automorphisms is at most 3. Here there is actually only 1, the identity! Let  $\phi$  be any automorphism of  $F$  fixing  $\mathbb{Q}$ . Consider the image  $\phi(2^{\frac{1}{3}})$ . This clearly satisfies the equation  $x^3 - 2 = 0$ , since we can apply  $\phi$  to this equation after substituting in  $2^{\frac{1}{3}}$ . Hence since the two complex roots of the polynomial are not in  $F$ , we see that  $\phi(2^{\frac{1}{3}}) = 2^{\frac{1}{3}}$ . But then since a basis for  $F$  as a vector space over  $\mathbb{Q}$  is  $1, 2^{\frac{1}{3}}, 2^{\frac{2}{3}}$ , we see that  $\phi$  maps each basis element to itself, so is the identity transformation. So  $I$  is the only automorphism in this case and the extension is not Galois.

To fix up example 2, let's extend the field to the splitting field of the polynomial  $x^3 - 2$ . So we also need to adjoin the roots  $\omega 2^{\frac{1}{3}}$  and  $\omega^2 2^{\frac{1}{3}}$ , where  $\omega = e^{2\pi i/3}$ . The minimal polynomial for  $\omega$  is  $x^2 + x + 1$ , since  $\omega$  is a cube root of unity, ie  $\omega^3 = 1$ .

**Example 38** The splitting field of  $x^3 - 2$  is a degree 6 extension of  $\mathbb{Q}$ , given by  $F = \mathbb{Q}[2^{\frac{1}{3}}, \omega]$ . A basis for  $F$  as a vector space over  $\mathbb{Q}$  is  $1, 2^{\frac{1}{3}}, 2^{\frac{2}{3}}, \omega, \omega 2^{\frac{1}{3}}, \omega 2^{\frac{2}{3}}$ . So we can view  $F$  as an extension of the subfield  $E = \mathbb{Q}[2^{\frac{1}{3}}]$  by adjoining  $\omega$ . We see that  $[F, E] = 2, [E, \mathbb{Q}] = 3, [F, \mathbb{Q}] = 6$ . Finally we would like to compute the Galois group  $G = \text{Aut}(F, \mathbb{Q})$ . We claim this is given by the permutations of the roots of  $x^3 - 2$ . Notice that any automorphism of

$F$  fixing  $\mathbb{Q}$  must permute these roots. We want to show there is exactly one automorphism corresponding to each such permutation. Certainly the information given by the images of the roots does completely determine the automorphism, since all the basis elements are powers of the roots. So the only problem is to show that the permutations actually give automorphisms. For simplicity, write  $\eta = 2^{\frac{1}{3}}$ . Write any element  $u \in F$  in the form

$$u = a_1 + a_2\eta + a_3\eta^2 + a_4\omega + a_5\eta\omega + a_6\eta^2\omega,$$

where  $a_i \in \mathbb{Q}$ . Now any automorphism  $\phi$  takes  $\eta$  to one of  $\eta, \omega\eta, \omega^2\eta$  and  $\omega$  to one of  $\omega, \omega^2$  (the two roots of the minimal polynomial  $x^2 + x + 1$  for  $\omega$ ). One can now check directly that the induced map on  $u$  preserves addition and multiplication. (Addition is easy - it follows directly by construction as  $\phi$  is a linear transformation over  $\mathbb{Q}$ . Multiplication is the tricky issue). We would like a more general method for showing this works, without all the effort.

**Theorem 29** A Galois extension  $E \subseteq F$  is the splitting field of a polynomial  $f \in E[x]$ .

**Proof:** We give the general ideas, not all the details. We first show that if  $E \subseteq F$  is Galois, then  $F$  is the splitting field of a polynomial  $f$ . We use a result that any finite degree extension is *simple*, ie can be generated by a single element. So  $F = E[u]$  for some element  $u \in F$ . Let  $f$  be the minimal polynomial for  $u$  over  $E$  and assume that the roots of  $f$  are  $u = \eta_1, \dots, \eta_d$ . We know the degree of  $f$  is the same as the degree of the extension  $F = E[u]$  over  $E$ , ie  $d$ . We also know that  $\text{Aut}(F, E)$  has  $d$  elements.  $\sigma_1, \dots, \sigma_d$ . Now apply each of these automorphisms  $\sigma_i$  to  $u$ . The result is a root  $\eta_j$ . We claim that all the images  $\sigma_i(u)$  are different. So this will show that all the roots of  $f$  lie in  $F$ , so  $F$  is the splitting field of  $f$ . Now if  $\sigma_i(u) = \sigma_r(u)$ , then we would have  $\sigma_s(u) = u$  for some element  $\sigma_s$  different from the identity. But since  $u$  generates the whole field  $F$ , this is a contradiction. We conclude that all the images of  $u$  are distinct and so all the roots of  $f$  lie in  $F$ .

Next assume that  $F$  is a splitting field of a polynomial  $f$ . It can be shown that  $F$  contains all the roots of any irreducible polynomial  $g$  with a root in  $F$  and coefficients in  $E$ . So let's assume again that we have written  $F$  as a simple extension  $F = E[u]$  and that  $f$  is the minimal polynomial for  $u$ . Now we claim that all the maps which take  $u$  to  $\eta_i$ ,  $1 \leq i \leq d$  can be made into automorphisms of  $F$  fixing  $E$ . For notice that the fields  $E[\eta_i]$  are all isomorphic, since every  $\eta_i$  has the same minimal polynomial  $f$ . It follows that  $\sigma_i$  is an isomorphism between  $E[u]$  and  $E[\eta_i]$ . But these two fields are both  $F$ , since they are both subfields of  $F$  of the same degree as  $F$ . Hence we have found  $d$  distinct members of  $\text{Aut}(F, E)$ . It is now easy to show there are no other automorphisms by theorem 3 above.

■

We can now state the fundamental theorem of Galois theory.

**Theorem 30** Suppose that  $E \subseteq F$  is a Galois extension and that  $G = \text{Aut}(F, E)$ . Then there is a bijection between subfields  $E \subseteq K \subseteq F$  and subgroups  $H \subseteq G$ , given by  $K = \text{fix}H$ .  $K \subseteq F$  is always a Galois extension and so is  $E \subseteq K$  exactly when  $H$  is a normal subgroup of  $G$ . In the latter case, the Galois group of  $E \subseteq K$  is the quotient group  $G/H$ .

**Example 39** Let's take the case of  $F$  as the splitting field of  $x^4 - 10x + 1$  again. The Galois group  $G \cong Z_2 \oplus Z_2$ . There are three proper subgroups,  $H_i = \{1, \sigma_i\}$ , for  $1 \leq i \leq 3$ . Here  $\sigma_1(\sqrt{2}) = \sqrt{2}, \sigma_1(\sqrt{3}) = -\sqrt{3}, \sigma_2(\sqrt{2}) = -\sqrt{2}, \sigma_2(\sqrt{3}) = \sqrt{3}, \sigma_3(\sqrt{2}) = -\sqrt{2}, \sigma_3(\sqrt{3}) = -\sqrt{3}$ . So we see immediately that  $K_1 = \mathbb{Q}[\sqrt{2}], K_2 = \mathbb{Q}[\sqrt{3}], K_3 = \mathbb{Q}[\sqrt{6}]$  are the corresponding fixed fields!

To show that all finite degree extensions are simple, is not too hard so we insert it here.

**Theorem 31** Let  $F$  be a field of characteristic 0 and assume that  $F \subseteq F[a, b]$  is a finite degree extension. Then  $F[a, b] = F[c]$ . By induction, it follows immediately that any finite degree extension of such a field  $F$  is simple.

**Proof:**

Let  $K$  be a splitting field for the minimal polynomials  $f, g$  for  $a, b$  over  $F$ , respectively. Hence  $F[a, b] \subseteq K$ . Notice that since the characteristic is 0, all roots of  $f$  and of  $g$  are distinct. Choose an element  $c$  so that  $c = a + rb$  but  $c \neq a_i + rb_j$  for any root  $a_i$  of  $f$  and  $b_j \neq b$  of  $g$ , where  $r \in F$ . Since there are infinitely many choices of  $r$ , we can take one to avoid any solution for which  $a + rb = a_i + rb_j$ , since then  $a - a_i = r(b_j - b)$  determines  $r$ .

We claim that  $F[a, b] = F[c]$ . It suffices to show that  $a \in F[c]$ , for then  $c - a = rb \in F[c]$  also, so  $b \in F[c]$ . Let  $h(x) = f(c - rx)$ . Now  $h(b) = f(c - rb) = f(a) = 0$ . So  $h, g$  have a common factor of  $x - b$ . We claim that this is a g.c.d of these two polynomials. For any larger common factor must contain other factors of  $g$  which are all  $x - b_j$  for a different root  $b_j$ . But if  $x - b_j$  divided  $h(x)$ , we would have  $h(b_j) = f(c - rb_j) = 0$ . But then  $c - rb_j = a_i$ , contrary to assumption. Finally, clearly  $h(x)$  is a polynomial with coefficients in  $F[c]$ , so by the Euclidean algorithm in  $(F[c])[x]$ , it follows that  $x - b = p(x)h(x) + q(x)g(x)$  is in  $(F[c])[x]$ . Hence the coefficients, ie  $b \in F[c]$  and the theorem follows. ■

## Summary notes for Algebra 321

### CHAPTER TEN - Solvability of polynomials

For quadratic polynomials, there is a general formula involving square roots only for the solution in terms of the coefficients. So if  $x^2 + bx + c$  is defined over some field  $F$ , we have the roots are  $x = \frac{-b \pm \sqrt{b^2 - 4c}}{2}$ . So if we take the field  $K = F(b, c)$  of rational functions (quotients of polynomials) in the two variables  $b, c$ , then the roots are obtained by adjoining the expression  $\sqrt{b^2 - 4c}$ . We say the solution is obtained *by forming radicals*.

For a cubic equation, the solution is called *Cardan's formula*. We start with  $x^3 + ax^2 + bx + c$ . Let

$$\begin{aligned}p &= b - a^2/3 \\q &= 2a^3/27 - ab/3 + c \\P &= (-q/2 + \sqrt{p^3/27 + q^2/4})^{1/3} \\Q &= (-q/2 - \sqrt{p^3/27 + q^2/4})^{1/3}.\end{aligned}$$

Here the cube roots must be chosen carefully. Then the roots of the cubic are

$$\begin{aligned}P + Q - a/3 \\ \omega P + \omega^2 Q - a/3 \\ \omega^2 P + \omega Q - a/3,\end{aligned}$$

where  $\omega = \exp(2\pi i/3)$ .

So again the general formula is obtained by adjoining square roots and then cube roots of rational functions of the coefficients, thought of as variables. We want to discuss the fact that there is no such general formula for polynomials of degree 5 or higher, although many such *particular* polynomials can be solved by radicals.

The key idea is to study the Galois group of the splitting field of the polynomials of degree  $n$ . Our first aim is to understand why there are such polynomials giving the full symmetric group as Galois group. Then we will define solvability for finite groups and show that  $S_n$  is solvable, exactly when  $n \leq 4$ . Finally we sketch why solvability of the Galois group is the same as a general solution in terms of radicals, as given by the quadratic and cubic formulae above.

**Example 40** *We want to show any irreducible polynomial of degree a prime  $p$  which has exactly two roots which are complex and not real numbers has Galois group  $S_p$ . Let  $f(x)$  be such a polynomial, with roots  $b_1, b_2$  which are complex and  $a_3, \dots, a_p$  which are real. Notice that complex conjugation  $\sigma$  sends  $f$  to itself, so permutes the roots. In particular all real roots  $a_i$  are fixed by  $\sigma$  and  $b_1, b_2$  are switched. Such a permutation is called a transposition. Next, we know that the subfield  $\mathbb{Q}[b_1]$  has degree  $p$  over  $\mathbb{Q}$  since this is the degree of the minimal polynomial  $f$  for  $b_1$ . Hence the degree of the splitting field  $K$  of  $f$  must be a multiple of  $p$ . We conclude since  $K$  is a Galois extension, that the order of the Galois group  $G$  is also divisible by  $p$ . Now by Cayley's theorem,  $G$  has an element of order  $p$ , ie*

$\tau^p = 1$  and no smaller power of  $\tau$  is 1. As a permutation of the roots of  $f$ , we can reorder the roots as  $u, \tau(u), \tau^2(u), \dots, \tau^{p-1}(u)$ , for any initial choice of root  $u$ . But then the effect of  $\tau$  as a permutation of these roots is to act as a  $p$ -cycle, ie  $1 \rightarrow 2 \rightarrow 3 \rightarrow \dots \rightarrow p \rightarrow 1$ . We can write  $\sigma$  as the permutation  $i \leftrightarrow j$  in terms of this ordering of the roots.

To complete this example, we show that the smallest group  $H$  containing  $\sigma, \tau$  is the whole symmetric group  $S_p$  of order  $p!$ . We illustrate with  $\sigma = (12345), \tau = (24)$ . Now taking expressions  $\sigma^i \tau \sigma^{-i}$  gives transpositions (13), (52), (41), (35) also in  $H$ . Next expressions like (13)(14)(13) = (34) give all the other transpositions are in  $H$ . But every permutation is a product of transpositions so every permutation is in  $H = S_5$ . The same method works generally.

### Definition

The commutator  $[a, b]$  of elements  $a, b \in G$  of a group  $G$  is defined by

$$[a, b] = a^{-1}b^{-1}ab.$$

The conjugate of an element  $a$  by an element  $g$  is denoted by  $a^g = g^{-1}ag$ . A subgroup  $H \subseteq G$  is *normal* if  $g^{-1}Hg = H$  for all  $g \in G$ . In this case, we can define the quotient group  $G/H$  as the set of left cosets  $gH$  for  $g \in G$ . There is a homomorphism from  $G$  onto  $G/H$  with kernel  $H$  given by sending  $g$  to the coset  $gH$ . Generally the kernel of any group homomorphism  $\phi$  is a normal subgroup and the analogue of the first isomorphism theorem holds for groups, ie the image of  $\phi$  is isomorphic to the quotient of the domain group by the kernel of  $\phi$ . Given a group  $G$ , the commutator subgroup  $[G, G] = G'$  is defined as the set of finite products of commutators of elements of  $G$ . Since  $[a, b]^{-1} = [b, a]$  and  $[a, b]^g = [a^g, b^g]$  it follows that  $G'$  is a normal subgroup of  $G$ . Moreover given any homomorphism  $\phi$  from  $G$  into an Abelian group  $A$ , clearly the kernel of  $\phi$  contains  $G'$  since  $\phi[a, b] = [\phi(a), \phi(b)] = 0$  in  $A$ . Conversely the quotient group  $G/G'$  is Abelian since  $a^{-1}G'b^{-1}G'aG'bG' = [a, b]G' = G'$  since  $G'$  contains all commutators. So we see that the cosets  $aG'$  and  $bG'$  commute, ie  $aG'bG' = bG'aG'$ .

### Definition

A group  $G$  is called *solvable* if its *derived series* of subgroups ends at the identity element 1. So there is a finite series of subgroups  $\{1\} = G^{(k)} \subset G^{(k-1)} \subset \dots \subset G' \subset G$ . Notice that each subgroup is normal in the next subgroup and the quotient groups are all Abelian.

We now check that  $S_3$  and  $S_4$  are solvable, but  $S_5$  is not solvable. Notice that  $S_3$  is the group of symmetries of an equilateral triangle, with vertices 1, 2, 3 and  $S_4$  is the symmetries of a regular tetrahedron with vertices 1, 2, 3, 4.

Our first observation is that the commutator subgroup of  $S_n$  is the alternating group of *even* permutations. We can define the *parity* of a permutation, by writing it as a product of transpositions and counting if there are an even or odd number. Now it is not obvious why this number is 'well defined', since there are a lot of different ways of writing permutations in this way. An easy way of proceeding is to use matrices to represent permutations.

If we think of 1, 2, 3 as basis vectors  $e_1, e_2, e_3$  for  $\mathbb{R}^3$ , then the transposition (12) has matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This has determinant  $-1$  as does any transposition. So any permutation has matrix with determinant  $\pm 1$  where an even product of transpositions gives  $+1$ . So we see that the alternating group  $A_n$  is well defined. Note  $A_n$  is a normal subgroup of  $S_n$ , since the conjugate of an even permutation is even. Counting transpositions in  $a^g = g^{-1}ag$  gives  $|a| + 2|g|$ , where the number of transpositions in  $a$  is  $|a|$  which is even. The same idea shows that all commutators are in  $A_n$  and in fact this is the commutator subgroup.

Now in case of  $S_3$ , we get  $A_3 = \{1, (123), (132)\}$ . So  $A_3$  is a cyclic group  $\mathbb{Z}_3$  and we have the derived series is  $\{1\} \subset A_3 \subset S_3$  and  $S_3$  is solvable.

In case of  $S_4$ ,  $|S_4| = 24$  and  $|A_4| = 12$ . Now the even permutations represent rotations of the regular tetrahedron. There are 8 such rotations of order 3, such as (123), which fix the vertex 4 and give a  $120^\circ$  rotation of the opposite face. There are 3 rotations through an axis between midpoints of opposite edges 12 and 34. So these are the permutation (12)(34), (13)(24), (14)(23). Together with 1, these give a subgroup of order 4 in  $A_4$ . Now it is easy to see this is a normal subgroup of  $A_n$ , since the order of an element  $a$ , ie the smallest  $n$  for which  $a^n = 1$ , is the same as the order of any conjugate  $a^g$ . For  $(a^g)^n = (a^n)^g$ . Hence  $g^{-1}Hg = H$  for any  $g \in A_4$ , since all elements of order 2 lie in  $H$ . So we see that the derived series for  $S_4$  must be

$$\{1\} \subset H \subset A_4 \subset S_4$$

(One should prove exactly that each is the commutator subgroup of the next. This can be shown, since in fact here, there is a unique normal subgroup of each subgroup which has Abelian quotient group, so must be the commutator subgroup. Generally this might not be the case.)

Finally we sketch why  $S_5$  is not solvable. The same method works for  $S_n$  for any  $n \geq 5$ . We have that the first commutator subgroup is  $A_5$ , but then  $A_5$  is its own commutator subgroup and the derived series get 'stuck' at  $A_5$ . Consider any normal subgroup of  $A_5$ . We will show that  $A_5$  is *simple* ie has no normal subgroups other than  $\{1\}$  or  $A_5$ . This is stronger than we need as the commutator subgroup is always normal and can only be  $\{1\}$  if all elements commute, ie the group is already Abelian.

It is easy to check that there are 15 elements of order 2 of the form (12)(34), 20 elements of order 3 of the form (123) and 24 elements of order 5 of the form (12345) in  $A_5$ , which has 60 elements ( ie half the number in  $S_5$ ). Now if there was a normal subgroup  $H$  of  $A_5$ , given an element of order 2, 3, 5, then all its conjugates will also be in  $H$ . But we can conjugate by all 60 elements and get all the elements of order 2 or 3 and at least half the elements of order 5 in  $H$ . But now multiplying together elements will show that we get all of  $A_5$ . So any non trivial normal subgroup is the whole group and  $A_5$  is simple as claimed.

## Solvable polynomials by radicals

To finish, we quickly sketch the connection between solvability for polynomials and for (Galois) groups. For convenience, we assume that the base field  $E$  contains all  $n$ th roots of unity required, ie all complex numbers  $\exp(2k\pi i/n)$ . So we take the splitting field  $E$  over  $\mathbb{Q}$  of  $(x^{n_1} - 1)(x^{n_2} - 1)(x^{n_3} - 1)\dots(x^{n_r} - 1)$  where we will need  $n_1$ th,  $n_2$ th, ...,  $n_k$ th roots of unity. It turns out this does not alter the basic problem, but makes it much easier. Assume that an irreducible polynomial  $f(x)$  over  $\mathbb{Q}$  has roots obtained by adjoining radicals of the coefficients of  $f(x)$  to  $\mathbb{Q}$ . Let  $F$  be the splitting field of  $f$ . Now each time we adjoin a radical, we form an expression  $a^{\frac{1}{n_1}}, b^{\frac{1}{n_2}}, \dots, e^{\frac{1}{n_k}}$ . Here  $a$  is a coefficient of  $f(x)$ , so is in  $\mathbb{Q}$ .  $b \in \mathbb{Q}[a]$ ,  $c \in \mathbb{Q}[a, b]$ , etc. If we work instead in  $E$ , adding one root of  $x^{n_1} - a$  to  $E$  gives the splitting field of this irreducible polynomial! For the roots are all  $a^{\frac{1}{n_1}}, \omega a^{\frac{1}{n_1}}, \dots$  where  $\omega$  is an  $n_1$ th root of unity. So adding radicals is like a sequence of splitting field extensions where each polynomial is of the form  $x^{n_i} - a$ . Now we know that a splitting field is a Galois extension and in this case, the Galois group is order  $n_1$  and permutes all the roots. So it must be a cyclic group, since the only subgroup of  $S_{n_1}$  which has order  $n_1$  and moves 1 to each point  $1, 2, 3, 4, \dots, n_1$  is cyclic. Now we get a sequence of subfields of  $F$ , namely

$$E \subset E[a] \subset E[a, b] \subset \dots \subset F$$

where each subfield is Galois in the next and all Galois groups are cyclic. But then by the fundamental theorem of Galois theory, the Galois group  $G = \text{Aut}(F, E)$  has a series of subgroups  $G \supset H_1 \supset \dots \supset H_k \supset \{1\}$ , where  $\text{fix} H_k = E[a]$  etc. Moreover the each subgroup must be normal in the next one ( as the extensions are Galois) and the quotient groups  $H_{i-1}/H_i$  are cyclic. This proves that  $G$  is a solvable group.

We can go backwards by a similar argument; if  $f$  has a solvable Galois group then we can find a sequence of subgroups with cyclic quotient groups as above. The corresponding fixed fields are then obtained as splitting fields of the previous ones and the polynomial extension can be shown to be of the form  $x^{n_i} - a$ , because the Galois group of the extension is cyclic.