Groups, Rings and Modules

Prof. C.J.B. Brookes

Lent 2005

LaTeXed by Sebastian Pancratz

These notes are based on a course of lectures given by Prof. C.J.B. Brookes in Part IB of the Mathematical Tripos at the University of Cambridge in the academic year 2004–2005.

These notes have not been checked by Prof. C.J.B. Brookes and should not be regarded as official notes for the course. In particular, the responsibility for any errors is mine — please email Sebastian Pancratz (`sfp25`) with any comments or corrections.

Contents

Introduction

Groups

This part of the course builds on the Part IA course *Algebra & Geometry*. The first punchline we aim for will be Sylow's theorems. We'll especially think p-groups, that is, groups of order p^a.

Rings

A ring is a set R together with two operations $+$ and $\times$, where multiplication is distributive over addition. This includes fields, but also the integers $\mathbb{Z}$ and polynomial rings $\mathbb{C}[X]$. We will concentrate on commutative rings.

Number theory

The Gaussian integers are $\mathbb{Z}[i] \subset \mathbb{Q}[i] \subset \mathbb{C}$. This is an example of adjoining roots of integral polynomials to $\mathbb{Q}$, or $\mathbb{Z}$. These rings analogous to integers do not necessarily have unique factorization. In fact, we do get unique factorization for Gaussian integers since Euclid's algorithm works.

These ideas will be continued in the Part II course *Number fields*.

Algebraic geometry

A family J of polynomials in $\mathbb{C}[x_1, \ldots, x_n]$ in n variables induces a subset of $\mathbb{C}^n$ which consists of $(\lambda_1, \ldots, \lambda_n)$ giving you zero for all polynomials in J, that is, a set of common roots. Hilbert's basis theorem implies that all but finitely many of the polynomials in J are redundant. The set of common zeros of J is equal to the set of common zeros of a finite subset of J.

Modules

Modules are a generalisation of vector spaces, using scalars from a ring rather than a field. Our aim is a structure theorem for rings in which Euclid's algorithm works, for example $\mathbb{Z}$. We also consider the structure of algebraic groups, which is important in algebraic topology, as well as $\mathbb{C}[X]$ and the Jordan normal form.

Lecture 1

E-mail

The lecturer can be contacted at brookes@dpmms.cam.ac.uk.

Books

- There is a long list of recommended books in the schedules.
- J.B. Fraleigh, *A First Course in Abstract Algebra*
- B. Hartley, T.O. Hawkes, *Rings, Modules and Linear Algebra*
- P.J. Cameron, *Introduction to Algebra*
- M. Artin, *Algebra*

Chapter 1

Groups

1.1 Basic concepts

Definition (Group)**.** A set G is a *group* if there is a binary operation $G \times G \to G$, satisfying

(i) If $g_1, g_2 \in G$ then $g_1 g_2 \in G$ *(closure).*

(ii) For all $g_1, g_2, g_3 \in G$, $(g_1 g_2) g_3 = g_1 (g_2 g_3)$ *(associativity).*

(iii) There exists $e \in G$ such that, for all $g \in G$, $eg = ge = g$ *(identity element).*

(iv) For all $g \in G$ there exists $g^{-1} \in G$, $gg^{-1} = g^{-1} g = e$ *(inverses).*

Definition (Subgroup)**.** A subset H is a *subgroup* of G if it is a group under the restriction of the operation. We write $H \leq G$.

Example. (i) Additive groups, e.g. $(\mathbb{Z}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$.

(ii) Matrix groups, such as the general linear group $GL(\mathbb{R}^n)$, i.e. the group of invertible $n \times n$ matrices, or the special linear group $SL(\mathbb{R}^n)$, i.e. matrices with determinant 1.

(iii) Permutation groups, i.e. the symmetric group S_n of permutations on $\{1, \ldots, n\}$ and the alternating group A_n containing even permutations.

(iv) Cyclic groups of order n, i.e. $\{e, g, \ldots, g^{n-1}\}$, and dihedral groups D_{2n} of order $2n$, i.e. the symmetries of the regular n-gon.

(v) Abelian groups, e.g. groups of order 4 including the Viergruppe V.

(vi) The quaternion group $\{\pm 1, \pm i, \pm j, \pm k\}, ij = k, ji = -k$ of order 8.

Given a subgroup $H \leq G$, we can define an equivalence relation on G by

$$g_1 \sim g_2 \iff g_1^{-1} g_2 \in H.$$

This partitions G into equivalence classes, called *left cosets* of H in G,

$$gH = \{gh : h \in H\}.$$

All these cosets are of the same size $|H|$. Counting the elements of G we obtain

Theorem 1.1 (Lagrange)**.** Let G be a finite group with subgroup H. Then

$$|G| = |H||G : H|$$

where $|G : H|$ is the number of left cosets of H in G, called the *index* of H in G.

Note 1. You can also do the same with the equivalence relation $g_1 \sim g_2 \iff g_2 g_1^{-1} \in H$, obtaining *right cosets* as equivalence classes. The number of right cosets is equal to the number of left cosets.

Definition (Order)**.** The *order* $o(g)$ of $g \in G$ is the least $n \in \mathbb{N}$ such that $g^n = e$. Note this implies that if $g^m = e$ then $o(g) \mid m$.

Lemma 1.2. The order $o(g)$ divides $|G|$.

Proof. The set $\{e, g, \ldots, g^{o(g)-1}\}$ is a subgroup of G. Now apply Theorem 1.1. □

1.2 Normal subgroups, quotient groups, homomorphism and isomorphisms

Example. The integers lie inside the real numbers, $(\mathbb{Z}, +) \leq (\mathbb{R}, +)$. We would like to add cosets,

$$(\mathbb{Z} + r_1) + (\mathbb{Z} + r_2) = \mathbb{Z} + (r_1 + r_2).$$

In general, we want to copy this procedure to define an operation on the cosets of a subgroup. If $K \leq G$ we want to define $g_1 K g_2 K = g_1 g_2 K$. But this requires $K g_2 K = g_2 K$, after multiplying on left by g_1^{-1}. In particular, we need $K g_2 \subset g_2 K$ for all $g_2 \in G$, and also $K g_2^{-1} \subset g_2^{-1} K$ which implies $g_2 K \subset K g_2$. So to make the definition work we need that $gK = Kg$ for all $g \in G$.

Definition (Normal subgroup)**.** A subgroup $K \leq G$ is *normal* in G if $gK = Kg$ for all $g \in G$. Equivalently, $gKg^{-1} = K$ for all $g \in G$. We write $K \triangleleft G$.

Example. In abelian groups all subgroups are normal. In general, $\{e\}$ and G are always normal. In the dihedral group D_{2n}, the subgroup of rotations is normal but $\{e, \tau\}$, for a reflection τ, is not normal if $n \geq 3$.

Proposition 1.3. Let $K \triangleleft G$. Then the set of cosets of K in G form a group under the operation $g_1 K . g_2 K = g_1 g_2 K$.

Definition (Quotient group)**.** This group is the *quotient group* G/K.

Proof. We need to check that the definition is well-defined, that is, if $g_1 K = g_1' K$ and $g_2 K = g_2' K$ then $g_1 g_2 K = g_1' g_2' K$. (See *Algebra & Geometry*.)

The identity element is K, the inverse of gK is $g^{-1}K$. Associativity is inherited from G. □

Definition (Homomorphism)**.** A map $\theta \colon G \to H$ where G and H are groups is a *homomorphism* if

$$\theta(g_1 g_2) = \theta(g_1)\theta(g_2).$$

Lemma 1.4. (i) The kernel $\ker\theta = \{g \in G : \theta(g) = e\} \lhd G$ is a normal subgroup.

(ii) The image $\operatorname{Im}\theta = \{\theta(g) : g \in G\} \leq H$ is a subgroup.

Proof. (i) Let $K = \{g \in G : \theta(g) = e\}$. Then the left coset gK is the set $\{g_1 \in G : \theta(g_1) = \theta(g)\}$ since

$$\begin{aligned} & \theta(g_1) = \theta(g) \\ \iff & \theta(g^{-1}g_1) = e \\ \iff & g^{-1}g_1 \in K \\ \iff & g_1 \in gK. \end{aligned}$$

Similarly, the right coset Kg is the set $\{g_1 \in G : \theta(g_1) = \theta(g)\}$. Thus $Kg = gK$.

(ii) We need to check that $\theta(g_1)\theta(g_2)^{-1} \in \operatorname{Im}\theta$ for all g_1, g_2. But this is $\theta(g_1 g_2^{-1})$. □

Definition (Isomorphism)**.** An *isomorphism* is a bijective homomorphism.

Observe that 'isomorphic to' defines an equivalence relation on groups. The notation used is $\cong$.

Theorem 1.5 (First isomorphism theorem)**.** Let $\theta\colon G \to H$ be a homomorphism. Then $\ker\theta$ is a normal subgroup and $G/\ker\theta \cong \operatorname{Im}\theta$.

Proof. We define a map $\Phi\colon G/K \to \operatorname{Im}\theta, gK \mapsto \theta(g)$ where $K = \ker\theta$.

Check that Φ is well-defined,

$$\begin{aligned} g_1K = g_2K & \implies g_2^{-1}g_1 \in K \\ & \implies \theta(g_2^{-1}g_1) = e \\ & \implies \theta(g_2)^{-1}\theta(g_1) = e \\ & \implies \theta(g_1) = \theta(g_2). \end{aligned}$$

Φ is a homomorphism,

$$\begin{aligned} \Phi(g_1Kg_2K) &= \Phi(g_1g_2K) \\ &= \theta(g_1g_2) \\ &= \theta(g_1)\theta(g_2) \\ &= \Phi(g_1K)\Phi(g_2K). \end{aligned}$$

Φ is injective. If $\theta(g_1) = \theta(g_2)$ then $\theta(g_1^{-1}g_2) = e$ and so $g_1^{-1}g_2 \in K$. Hence $g_1K = g_2K$.

Clearly, Φ is surjective. □

Example. Consider the map

$$\theta\colon (\mathbb{R}, +) \to (\mathbb{C}^\times, \times), r \mapsto e^{2\pi ir}.$$

This is a group homomorphism, $\theta(r_1 + r_2) = \theta(r_1)\theta(r_2)$. Its kernel is $(\mathbb{Z}, +)$. Theorem 1.5 implies $\mathbb{R}/\mathbb{Z} \cong \operatorname{Im}\theta = \{e^{2\pi ir} : r \in \mathbb{R}\}$. The map Φ as above is $\Phi(\mathbb{Z} + r) = e^{2\pi ir}$.

Theorem 1.6 (Second isomorphism theorem)**.** Let $H \leq G$ and $K \triangleleft G$. Then $HK \leq G$ and $H \cap K \triangleleft G$ where $HK = \{hk : h \in H, k \in K\}$ and $(HK)/K \cong H/(H \cap K)$.

Proof. First show that HK is a subgroup. Take $hk \in HK$. Then

$$(hk)(h_1k_1)^{-1} = h \underbrace{kk_1^{-1}}_{\in K} h_1^{-1}$$

But $Kh_1^{-1} = h_1^{-1}K$ since $K \triangleleft G$ and so $(kk_1^{-1})h_1^{-1} = h_1^{-1}k_3$ for some $k_3 \in K$. So $(hk)(h_1k_1)^{-1} = hh_1^{-1}k_3 \in HK$.

Define

$$\theta \colon H \to G/K, h \mapsto hK.$$

This is a group homomorphism,

$$\begin{aligned} \theta(h_1h_2) &= h_1h_2K \\ &= h_1Kh_2K \\ &= \theta(h_1)\theta(h_2). \end{aligned}$$

So Theorem 1.5 implies that $H/\ker\theta \cong \operatorname{Im}\theta$. Now

$$\begin{aligned} \ker\theta &= \{h \in H : hK = K\} = H \cap K \\ \operatorname{Im}\theta &= \{hK : h \in H\} = (HK)/K, \end{aligned}$$

cosets of K in HK. Thus $H/(H \cap K) \cong (HK)/K$, as required. □

In fact, there is a bijection between subgroups of G/K and the subgroups of G containing K. This is given by the maps

$$\begin{array}{ccc} \{\text{subgroups of } G/K\} & & \{\text{subgroups of } G \text{ containing } K\} \\ X & \longrightarrow & \{g \in G : gK \in X\} \\ L/K & \longleftarrow & L \end{array}$$

These maps are inverses of each other.

Take a normal subgroup $K \triangleleft G$ and form the quotient group G/K, with elements the cosets gK and multiplication defined by $g_1Kg_2K = g_1g_2K$. The correspondence

$$\{\text{subgroups of } G/K\} \quad \longleftrightarrow \quad \{\text{subgroups of } G \text{ containing } K\}$$

restricts to

$$\{\text{normal subgroups of } G/K\} \quad \longleftrightarrow \quad \{\text{normal subgroups of } G \text{ containing } K\}$$

For a subgroup $L/K \leq G/K$ we have that

$$(gK)(L/K)(gK)^{-1} = (gLg^{-1})/K$$

so

$$(gK)(L/K)(gK)^{-1} = L/K \iff gLg^{-1} = L,$$

that is,

$$L/K \triangleleft G/K \iff L \triangleleft G.$$

Lecture 3

Theorem 1.7 (Third isomorphism theorem)**.** If K and L are normal subgroups of G with $K \leq L \triangleleft G$ then $(G/K)/(L/K) \cong G/L$.

Proof. Apply Theorem 1.5 to the well chosen map

$$\theta\colon G/K \to G/L, gK \mapsto gL.$$

We need to check this is well-defined,

$$\begin{aligned} & g_1K = g_2K \\ \implies & g_1^{-1}g_2 \in K \leq L \\ \implies & g_1L = g_2L. \end{aligned}$$

θ is a homomorphism,

$$\begin{aligned} \theta(g_1Kg_2K) &= \theta(g_1g_2K) \\ &= g_1g_2L \\ &= g_1Lg_2L \\ &= \theta(g_1K)\theta(g_2K). \end{aligned}$$

Clearly, $\operatorname{Im}\theta = G/L$. Also $\ker\theta = \{gK : gL = L\} = L/K$ and so Theorem 1.5 implies

$$(G/K)/(L/K) \cong G/L. \qquad \square$$

Definition (Simplicity)**.** G is *simple* if the only normal subgroups in it are $\{e\}$ and G itself.

Example. The only simple abelian groups are C_p for p prime.

Proof. 'Normal' for abelian groups is easy to understand as all subgroups are normal because of commutativity. Take a non-trivial element $g \neq e$ in a simple abelian group G.

Then either g is of infinite order and G contains the subgroup $\{\dots, g^{-1}, e, g, g^2, \dots\}$. Simplicity implies that $G = \{\dots, g^{-1}, e, g, g^2, \dots\}$. But then there is also the subgroup $\{\dots, g^{-2}, e, g^2, g^4, \dots\}$, a contradiction.

Or g is of finite order and G has the subgroup $\{e, g, g^2, \dots, g^{o(g)-1}\}$. Simplicity implies that $G = \{e, g, g^2, \dots, g^{o(g)-1}\}$. But if $o(g) = rs$, say, then we also have the subgroup $\{e, g^r, g^{2r}, \dots, g^{r(s-1)}\}$. Now simplicity implies that there are no non-trivial factorisations of $o(g)$ so $G = \{e, g, \dots, g^{p-1}\}$ for some prime p. $\square$

We shall see shortly that the alternating group A_5 on $\{1, \dots, 5\}$ is simple. In fact, A_n is simple for $n \geq 5$, but we don't need to know the proof. Note that $|A_3| = 3$, so A_3 is isomorphic to C_3 and simple. $|A_4| = 12$ and A_4 contains the normal subgroup $V = \{e, (12)(34), (13)(24), (14)(23)\} \triangleleft A_4$, so A_4 is not simple.

There are finitely many families of finite simple groups, and 26 groups that do not fit into those general families, called the *sporadic* ones.

Note $|A_5| = 60$ and $|A_6| = 360$. There is another finite simple group of order 168, $GL_3(\mathbb{Z}$ mod 2), the group of invertible 3×3 matrices entries $\{0, 1\}$ modulo 2. (This is left as an exercise.)

Theorem 1.8. Let G be a finite group. Then there are subgroups $H_1, \ldots, H_s$ such that

$$G = H_1 \geq H_2 \geq \cdots \geq H_s = \{e\}$$

with $H_{i+1} \triangleleft H_i$ and H_i/H_{i+1} simple.

Warning. Not all H_i normal in G.

Proof. Let H_2 be normal in G with $|H_2|$ of largest order not equal to $|G|$. Then the correspondence between normal subgroups of G containing H_2 and the normal subgroups of G/H_2 implies that G/H_2 simple. Repeat this to find $H_3 \triangleleft H_2$ with H_2/H_3 simple. The process stops when we reach $H_s = \{e\}$. □

Remark. The simple groups H_i/H_{i+1} are essentially unique, although there may be lots of ways of picking the subgroups H_i, called *composition factors*.

Definition. G is *soluble* if all composition factors can be chosen to be cyclic of prime order.

Remark. The terminology comes from Galois theory, see the relevant Part II course. You all know a formula for solving quadratic equations. You can do something similar involving roots for cubics and quartics, the process is called 'solution by radicals'. But you cannot necessarily do it for a quintic, e.g. $t^5 - 6t + 3 = 0$. This can be shown by defining a *Galois group* associated with the polynomial. If you can solve the polynomial by radicals then the Galois group is soluble.

1.3 Permutation groups, actions and representations

The symmetry group S_n consists of permutations of $\{1, \ldots, n\}$. Useful notation is the disjoint cycle form, e.g. $(1\ 2\ 3)(4\ 5)(6)$ where the single cycle is often suppressed.

It contains a subgroup $A_n \leq S_n$, the alternating group of even permutations.

We recall that every permutation can be expressed as a product of transpositions. An element g is even if you need an even number of transpositions in such a product. Equivalently in disjoint cycle form, g is even if the number of even length cycles is even, e.g. $(1\ 2\ 3) = (1\ 2)(2\ 3)$ is even but $(1\ 2\ 3\ 4) = (1\ 2)(2\ 3)(3\ 4)$ is odd.

We note that the orders of these two groups are

$$|S_n| = n!,$$
$$|A_n| = \frac{n!}{2}.$$

We can quickly see that the alternating group $A_n \triangleleft S_n$ is a normal subgroup. There are two left cosets of A_n, $S_n \setminus A_n = gA_n$ for all permutations g and two right cosets of A_n, $S_n \setminus A_n = A_n g$. Thus the left cosets are equal to the right cosets.

More generally, for any set Ω you can define the *symmetric group* $\operatorname{Sym} G$ of all bijective maps $\Omega \to \Omega$.

Definition. A group G is a *permutation group* of degree n if $G \leq \operatorname{Sym} \Omega$ with $|\Omega| = n$.

Lecture 4

Example. $S_n, A_n, D_{2n} \leq \operatorname{Sym}\Omega$ where Ω is the set of vertices of the regular n-gon.

Definition. An *action* of a group G on a set Ω is a map $G \times \Omega \to \Omega$, $(g, \alpha) \mapsto g * \alpha$, often also written as $g(\alpha)$, such that

(i) $g * \alpha \in \Omega$ for each $\alpha \in \Omega$,

(ii) $g_1 * (g_2 * \alpha) = g_1 g_2 * \alpha$,

(iii) $e * \alpha = \alpha$.

Lemma 1.9. Let G act on Ω. Then for $g \in G$ the map $\Phi_g \colon \alpha \mapsto g * \alpha$ is a permutation of Ω and the map

$$\Phi \colon G \to \operatorname{Sym}\Omega, g \mapsto \Phi_g$$

is a homomorphism, a *permutation representation* of G.

The image $\Phi(G)$ is a subgroup of $\operatorname{Sym}\Omega$, denoted G^Ω. The kernel of Φ, $\ker\Phi = \{g \in G : g * \alpha = \alpha \text{ for all } \alpha \in \Omega\}$, also denoted $G_{(\Omega)}$, is a normal subgroup and $G/G_\Omega \cong G^\Omega$.

Proof. $\Phi_{g^{-1}}$ is the inverse of Φ_g and so Φ_g is a bijection. The map $\Phi \colon G \to \operatorname{Sym}\Omega$ is a homomorphism, from the second property of the definition of action. The rest is from Theorem 1.5. □

Example. (i) Let G be the symmetry group of the cube, $|G| = 48$, and Ω the set of diagonals, $|\Omega| = 4$. G acts on Ω,

$$G^\Omega = S_4,$$
$$G_{(\Omega)} = \{e, \text{symmetry sending each vertex to the opposite one}\}.$$

(ii) The left regular action of G on $\Omega = G$, given by

$$g * g_1 = g g_1.$$

Φ_g is left multiplication by g. The kernel is $\{e\}$, so Lemma 1.9 gives Cayley's Theorem 1.10.

(iii) For a subgroup H of G, let Ω be the set of left cosets of H. Consider the action of G on Ω given by

$$g * g_1 H = g g_1 H.$$

The kernel is $\bigcap_{g_1 \in G} g_1 H g_1^{-1}$ since, for all $g_1 \in G$,

$$\begin{aligned} & g g_1 H = g_1 H \\ \iff & g_1^{-1} g g_1 H = H \\ \iff & g_1^{-1} g g_1 \in H \\ \iff & g \in g_1 H g_1^{-1}. \end{aligned}$$

The kernel is the largest normal subgroup of G contained in H. If $K \triangleleft G, K \leq H$ then $g_1 K g_1^{-1} = K$ since K is normal. But $g_1 K g_1^{-1} \leq g_1 H g_1^{-1}$ so $K \leq g_1 H g_1^{-1}$ for all $g_1 \in G$.

Theorem 1.10 (Cayley)**.** Any group G is isomorphic to a subgroup of $\operatorname{Sym} G$.

Theorem 1.11. Let G be a finite group and H a proper subgroup of G of index n. Then there is a normal subgroup K of G contained in H such that G/K is isomorphic to a subgroup of S_n. In particular $|G : K|$ divides $n!$ and is at least n.

Moreover, if G is non-abelian simple then G is isomorphic to a subgroup of A_n and $n \geq 5$.

Proof. Let K be the kernel of the action of G on Ω, where Ω is the set of left cosets of H. Lemma 1.9 implies that $G/K \cong G^\Omega$, a subgroup of S_n. Lagrange's theorem 1.1 implies subgroups of S_n have order dividing $n!$. Since $K \leq H$ and H is of index n in G we have that $|G/K| \geq n$.

Now assume G is non-abelian simple. Then $K = \{e\}$, and so $G \cong G^\Omega$ subgroup of S_n. But $A_n \lhd S_n$ and so $G^\Omega \cap A_n \lhd G^\Omega$. Simplicity implies that either

$$G^\Omega \cap A_n = \{e\} \text{ or } G^\Omega \cap A_n = G^\Omega.$$

In the first case, the second isomorphism theorem implies

$$G^\Omega/(G^\Omega \cap A_n) \cong (G^\Omega A_n)/A_n \leq S_n/A_n$$

and hence $|G^\Omega| \leq 2$ since $G^\Omega \cap A_n = \{e\}$, a contradiction. Finally, we have $n > 4$ since A_4 has no non-abelian simple subgroups. □

Let us recall some definitions from the Part IA course *Algebra & Geometry*. Let G act on Ω.

Definition. The *orbit* on Ω containing α is $G(\alpha) = \{g * \alpha : g \in G\}$. The *stabiliser* of α is $G_\alpha = \{g \in G : g * \alpha = \alpha\}$.

Theorem 1.12 (Orbit-stabiliser theorem)**.** Let G act on Ω. Then $|G : G_\alpha| = |G(\alpha)|$.

Proof. The map

$$\{\text{cosets of } G_\alpha \text{ in } G\} \to G(\alpha), gG_\alpha \mapsto g * \alpha$$

is a bijection. □

1.4 Conjugacy classes, centralisers and normalisers

We consider the conjugation action of G on $\Omega = G$, defined by

$$g * x = gxg^{-1}.$$

for $x \in \Omega = G$. In this case the permutation $\Phi_g \colon x \mapsto gxg^{-1}$ as well as being bijective is a homomorphism $G \to G$. Thus they are *automorphisms*. The automorphism group $\operatorname{Aut} G$ of isomorphisms $G \to G$ is a group under the composition of maps. We have $\Phi_g \in \operatorname{Aut} G$.

Orbits are called *conjugacy classes*, $\operatorname{ccl}_G(x) = \{gxg^{-1} : g \in G\}$. Stabilisers are called *centralisers* of x, $C_G(x) = \{g \in G : gxg^{-1} = x\} = \{g \in G : gx = xg\}$.

The orbit-stabiliser theorem implies $|G : C_G(x)| = |\operatorname{ccl}_G(x)|$. The kernel of the action

$$G_{(\Omega)} = \{g \in G : \forall x \in \Omega\ gxg^{-1} = x\}$$

Lecture 5

$$= \{g \in G : \forall x \in \Omega\ gx = xg\}$$
$$= \bigcap_{x \in G} C_G(x)$$

is called the *centre* of G, $Z(G)$.

Observe that $|\mathrm{ccl}_G(x)|$ divides $|G|$, which is part of Lagrange's Theorem 1.1.

Lemma 1.13. Let G be a finite group. Then $1 = \sum 1/|C_G(x)|$, summing over distinct conjugacy classes.

Proof. We count the elements of G. $|G| = \sum |\mathrm{ccl}_G(x)|$, so

$$|G| = \sum \frac{|G|}{|C_G(x)|}$$

since $|\mathrm{ccl}_G(x)| = |G|/|C_G(x)|$. Now divide by $|G|$. □

Example (Conjugacy classes of S_n and A_n)**.** Recall from *Algebra & Geometry* that two permutations are conjugate in S_n precisely when they have the same cycle type when expressed in disjoint cycle form, e.g. $2^2\ 1$ denotes the cycle type of double transpositions in S_5, e.g. $(1\ 2)(3\ 4)(5)$.

In S_5, we have the following cycle types:

cycle type	1^5	$2^2\ 1$	$3\ 1^2$	5	$2\ 1^3$	$3\ 2$	$4\ 1$
# elements	1	15	20	24	10	20	30
	even permutations				odd permutations		

For $x \in A_n$ we have that $\mathrm{ccl}_{A_n}(x) \subset \mathrm{ccl}_{S_n}(x)$. We also have

$$|\mathrm{ccl}_{S_n}(x)| = |S_n : C_{S_n}(x)|,$$
$$|\mathrm{ccl}_{A_n}(x)| = |A_n : C_{A_n}(x)|$$

by the orbit-stabiliser theorem. But $C_{A_n}(x) = A_n \cap C_{S_n}$, and thus this set has index 1 or 2 in $C_{S_n}(x)$. To see this, use the second isomorphism theorem, Theorem 1.6,

$$C_{S_n}(x)/(A_n \cap C_{S_n}(x)) \cong (A_n C_{S_n}(x))/A_n \leq S_n/A_n.$$

If all permutations that commute with x are even, i.e. $C_{S_n}(x) = C_{A_n}(x)$,

$$|\mathrm{ccl}_{S_n}(x)| = \frac{|S_n|}{|C_{S_n}(x)|} = \frac{2|A_n|}{C_{A_n}(x)} = 2|\mathrm{ccl}_{A_n}(x)|.$$

If there is an odd permutation commuting with x, i.e. $|C_{S_n}(x) : C_{A_n}(x)| = 2$,

$$|\mathrm{ccl}_{S_n}(x)| = \frac{|S_n|}{|C_{S_n}(x)|} = \frac{|A_n|}{C_{A_n}(x)} = |\mathrm{ccl}_{A_n}(x)|.$$

Example. Consider S_5 and A_5, take $x \in A_5$. $(1\ 2)(3\ 4)(5)$ commutes with $(1\ 2)$, $(1\ 2\ 3)(4)(5)$ commutes with $(4\ 5)$. But conjugacy classes of 5 cycles do not split into 2 conjugacy classes of size 12. Consider $(1\ 2\ 3\ 4\ 5)$ and its conjugates.

$$g(1\ 2\ 3\ 4\ 5)g^{-1} = (g(1)\ \cdots\ g(5))$$
$$C_{S_n}(1\ 2\ 3\ 4\ 5) = \{e, (1\ 2\ 3\ 4\ 5), (1\ 3\ 5\ 2\ 4), (1\ 4\ 2\ 5\ 3), (1\ 5\ 4\ 3\ 2)\}$$

$C_{S_n}(1\ 2\ 3\ 4\ 5)$ is of order 5, all elements are even.

Proposition 1.14. A_5 is simple.

Proof. Consider a normal subgroup K of A_5. Thus $gkg^{-1} \in K$ for all $k \in K$, $g \in A_5$. Thus K must be a union of conjugacy classes in A_5. But by Lagrange's Theorem 1.1, $|K|$ divides $60 = |A_5|$. There is no way of taking a union of the conjugacy classes we've worked out to give $|K|$ dividing 60, unless $K = \{e\}$ or $K = A_5$. □

1.5 Finite p-groups

Definition (p-group)**.** A finite group G is a *p-group* if $|G| = p^n$ for prime p.

Theorem 1.15. Let G be a finite p-group. Then $Z(G) \neq \{e\}$.

Proof. We count the elements of G,

$$|G| = \sum |\mathrm{ccl}_G(x)|.$$

But the size of conjugacy classes divides $|G| = p^n$. So $|\mathrm{ccl}_G(x)| = p^a$ for some $0 \leq a \leq n$.

$$\begin{aligned} |G| = &(\text{no. of conjugacy classes of size } 1) \cdot 1 \\ &+ (\text{no. of conjugacy classes of size } p) \cdot p \\ &+ \cdots \end{aligned}$$

But p divides $|G|$. So p divides the number of conjugacy classes of size 1. But $\{x\}$ is a conjugacy class if and only if

$$\begin{aligned} &\forall g \in G\ gxg^{-1} = x \\ \iff &x \in Z(G) \end{aligned}$$

So $p \mid |Z|$. Thus $Z(G) \neq \{e\}$. □

Lemma 1.16. For any group G, if $G/Z(G)$ is cyclic then G is abelian.

Proof. Suppose gZ generates $G/Z(G)$ and thus each coset in Z is of the form $(gZ)^r = g^r Z$ for some r. Thus any element $x \in G$ is of the form $g^r z$ for some $z \in Z, r \in \mathbb{N}$.

$$\begin{aligned} x_1 x_2 &= (g^{r_1} z_1)(g^{r_2} z_2) \\ &= g^{r_1} g^{r_2} z_1 z_2 \\ &= (g^{r_2} z_2)(g^{r_1} z_1) \\ &= x_2 x_1 \end{aligned}$$

using that $z_1 \in Z$. Thus G is abelian and so $G = Z(G)$. □

Proposition 1.17. If $|G| = p^2$ for p prime then G is abelian.

Proof. By Theorem 1.15, $|Z(G)| = p$ or p^2 since by Lagrange's theorem $|Z(G)|$ is 1, p or p^2. By Lemma 1.16, $|Z(G)| \neq p$, for if $|Z(G)| = p$ then $|G/Z(G)| = p$ and so $G/Z(G)$ cyclic. So $|Z(G)| = p^2$. Hence G is abelian. □

Remark. Groups of order p^3 are considered on the example sheet.

Lecture 5

Recall the definitions of the *direct product*.

Definition (External direct product)**.** Given two groups G and H, $G \times H$ has a natural group structure $(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2)$. $G \times H$ has subgroups $G_1 = \{(g, e_H) : g \in G\} \cong G$ and $H_1 = \{(e_G, h) : h \in H\} \cong H$. Note that $G_1 \cap H_1$ is the trivial subgroup of $G \times H$ and $G_1 H_1 = G \times H$.

Definition (Internal direct product)**.** Given a group L and subgroups G_1, H_1 with $G_1 H_1 = L$, $G_1 \cap H_1 = \{e\}$, all elements $g \in G_1$ commute with all elements $h \in H_1$. Note that $L \cong G_1 \times H_1$.

Theorem 1.18. Let G be a p-group of order p^a. Then there is a subgroup of order p^b for any $0 \leq b \leq a$.

Proof. By induction on a. The result is clearly true for $a = 1$. Suppose $a > 1$ and pick an element $x \in Z(G), x \neq e$, noting that by Theorem 1.15 the centre is non-trivial. By taking a suitable power of it, we may assume that x is of order p. Hence $K = \{e, x, \ldots, x^{p-1}\}$ is a subgroup of G, and it is normal in G since x is central. Consider the quotient G/K. Its order is p^{a-1} and we can apply the inductive hypothesis to get $L/K \leq G/K$ of order p^{b-1}. The correspondence between subgroups of G/K and subgroups of G containing K gives a subgroup L of G containing K of order p^b. □

1.6 Finite abelian groups

For example, by Proposition 1.17 groups of order p^2 are abelian, and hence isomorphic to C_{p^2} or $C_p \times C_p$.

Theorem 1.19. Let G be a finite abelian group. Then

$$G \cong C_{d_1} \times C_{d_2} \times \cdots \times C_{d_k}$$

with $d_{i+1} \mid d_i$ and $d_1 \cdots d_k = |G|$.

Proof. Postponed until Section 3. □

Remark (Remark on proof)**.** One can be less sophisticated. As the first step, consider an abelian p-group G and the subgroup $\langle x \rangle$ generated by $x \in G$ of maximal order. If $\langle x \rangle = G$ we are done. Otherwise, if $\langle x \rangle \neq G$, the claim is that there exists a subgroup $H \leq G$ with $\langle x \rangle H = G$, $\langle x \rangle \cap H = \{e\}$, i.e. $G \cong \langle x \rangle \times H$.

Example. The abelian groups of order 8 are C_8, $C_4 \times C_2$ and $C_2 \times C_2 \times C_2$. The abelian groups of order 16 are C_{16}, $C_8 \times C_2$, $C_4 \times C_4$, $C_4 \times C_2 \times C_2$ and $C_2 \times C_2 \times C_2 \times C_2$.

Lemma 1.20. If m and n are coprime then $C_m \times C_n \cong C_{mn}$.

Proof. Take elements g of order m in C_m and h or order n in C_n.

Then (g, h) is of order mn in $C_m \times C_n$ since $(g, h)^r = (g^r, h^r)$ and so the order of (g, h) is the least common multiple of m and n which is mn since m and n are coprime.

But $|C_m \times C_n| = mn$ and so (g, h) is a generator for the group. □

Example. $C_2 \times C_3 \cong C_6$. Abelian groups of order $24 = 3 \times 8$ are C_{24}, $C_{12} \times C_2$, and $C_6 \times C_2 \times C_2$.

1.7 Sylow's theorems and applications to groups of small orders

Theorem 1.21 (Sylow)**.** Let G be a finite group of order $p^a m$ where p is prime and $p \nmid m$.

(i) There exists a subgroup P of order p^a, called the *Sylow p-subgroup*.

(ii) All Sylow p-subgroups are conjugate.

(iii) The number n_p of Sylow p-subgroups satisfies $n_p \equiv 1 \pmod{p}$ and divides $|G|$, and hence $n_p \mid m$.

Remark. G acts on the set Ω of Sylow p-subgroups via conjugation, $g * P = gPg^{-1}$. By Sylow's second theorem, there is exactly one orbit. The stabiliser, called the *normaliser* of P, is $N_G(P) = \{g \in G : gPg^{-1} = P\}$. By the orbit-stabiliser theorem, the size of the orbit is $n_p = |G : N_G(P)|$. Thus $n_p \mid |G|$.

Lemma 1.22. If $n_p = 1$ then the unique Sylow p-subgroup is normal in G.

Proof. Take the unique Sylow p-subgroup P. Then gPg^{-1} is also a Sylow p-subgroup. Thus $gPg^{-1} = P$ for all $g \in G$. □

Remark (continued)**.** By definition, $P \triangleleft N_G(P)$. So applying Sylow's second theorem to $N_G(P)$ we see that all Sylow p-subgroups of $N_G(P)$ are conjugate to P. But $P \triangleleft N_G(P)$ and so all its conjugates are P itself. Thus P is the unique Sylow p-subgroup of $N_G(P)$.

Corollary 1.23. Let G be a non-abelian simple group. Then $|G|$ divides $n_p!/2$, and $n_p \geq 5$ once we have shown that all non-abelian subgroups of S_4 are not simple.

Proof. G is acting on the set Ω of Sylow p-subgroups where $|\Omega| = n_p$. If G is non-abelian simple then by Theorem 1.1 G is isomorphic to a subgroup of A_{n_p}, and $n_p \geq 5$ subject to proviso. □

Example. Let $|G| = 1000 = 2^3 5^3$. Then G is not simple.

Proof. $n_5 \equiv 1 \pmod{5}$ and $n_5 \mid 8$. So $n_5 = 1$ and so there exists a unique Sylow 5-subgroup which is normal. Thus G is not simple. □

Example. Let $|G| = 300 = 2^2 \cdot 3 \cdot 5^2$. Then G is not simple.

Proof. $n_5 \equiv 1 \pmod{5}$ and $n_5 \mid 12$. Assume G is simple and so $n_5 \neq 1$. Then $n_5 = 6$. But 300 does not divide $6!/2$. Contradiction! Thus $|G| = 300$ implies that $n_5 = 1$. (G is not abelian simple since it is not of prime order.) □

Example. Let $|G| = 132 = 2^2 \cdot 3 \cdot 11$. Then G is not simple.

Proof. $n_{11} \equiv 1 \pmod{11}$ and $n_{11} \mid 12$. Assume G is simple and so $n_{11} \neq 1$. Then $n_{11} = 12$. $n_3 \equiv 1 \pmod{3}$ and $n_3 \mid 44$. By simplicity, $n_3 \neq 1$ and $n_3 \neq 4$ by Corollary 1.23, so $n_3 = 22$.

But since $n_{11} = 12$ and $n_3 = 22$ there are $12 \cdot (11-1)$ elements of order 11 and $22 \cdot (3-1)$ elements of order 3. But this gives too many elements, a contradiction! □

Lecture 7

This is a typical argument. We use Sylow's theorem to estimate the number of subgroups and then count elements.

Lemma 1.24. Suppose $|G| = 2p$ for an odd prime p. Then $G \cong C_{2p}$ or D_{2p}.

Proof. $n_p \equiv 1 \pmod{p}$ and $n_p \mid 2$, so $n_p = 1$. Thus there is a normal subgroup K of order p. Sylow's theorem also implies there is a subgroup H of order 2. Let $H = \{e, g\}$, $K = \{e, x, \ldots, x^{p-1}\}$. Note $H \cap K = \{e\}$ and G is a semidirect product of K by H. Observe that $|HK| = 2p$.]

Now consider the action of H on K by conjugation. In particular, consider conjugation by g — it permutes the elements of K. (In fact, conjugation by g gives an automorphism of K.) This is uniquely determined by the image of x, that is, $gxg^{-1} = x^r$ for some r. Then also

$$x = g^2xg^{-2} = gx^rg^{-1} = (gxg^{-1})^r = x^{r^2}.$$

So $r^2 \equiv 1 \pmod{p}$. This gives two choices. Suppose $r \equiv 1 \pmod{p}$, then g and x commute so gx is of order $2p$ and hence $G \cong C_{2p}$. Otherwise suppose $r \equiv -1 \pmod{p}$, then $gxg^{-1} = x^{-1} = x^{p-1}$ and $G \cong D_{2p}$. In the latter case, x is a rotation and g is a reflection. □

There is a problem on the examples sheet for the case $|G| = 15$.

Proof of Sylow's theorems

Let G be a finite group of order $|G| = p^a m$, where p is prime and $p \nmid m$.

(i) There exists subgroups of order p^a, called Sylow p-subgroups.

Proof. Consider the set Ω of subsets of G of size p^a and let $\{g_1, \ldots, g_{p^a}\} \in \Omega$. G acts on Ω via left multiplication,

$$g * \{g_1, \ldots, g_{p^a}\} = \{gg_1, \ldots, gg_{p^a}\}.$$

Consider an orbit $\Sigma \leq \Omega$. Take $g \in G$. If $\{g_1, \ldots, g_{p^a}\} \in \Sigma$ then

$$gg_1^{-1} * \{g_1, \ldots, g_{p^a}\} = \{g, gg_1^{-1}g_2, \ldots, gg_1^{-1}g_{p^a}\} \in \Sigma.$$

So g appears as an entry in one of the p^a-sets in orbit the Σ. Counting,

$$|\Sigma| \geq \frac{|G|}{p^a} = \frac{p^a m}{p^a} = m.$$

So we have two types of orbits Σ, (a) $|\Sigma| = m$, (b) $|\Sigma| > m$. By the orbit-stabiliser theorem, $|\Sigma|$ divides $|G| = p^a m$. So for orbits Σ of type (b), where $|\Sigma| > m$, we deduce that $p \mid |\Sigma|$.

The next step is to show that $|\Omega|$ is not divisible by p,

$$|\Omega| = \binom{p^a m}{p^a} = \frac{p^a m}{p^a} \frac{p^a m - 1}{p^a - 1} \cdots \frac{p^a m - p^a + 1}{1}.$$

Consider $(p^a m - k)/(p^a - k)$ for $0 \leq k \leq p^a - 1$. If $k = 0$ then

$$\frac{p^a m - k}{p^a - k} = \frac{p^a m}{p^a} = m.$$

If $k > 0$ then set $k = p^b q$, where $p \nmid q$, so that

$$\begin{aligned}\frac{p^a m - k}{p^a - k} &= \frac{p^a m - p^b q}{p^a - p^b q} \\ &= \frac{p^{a-b} m - q}{p^{a-b} - q}.\end{aligned}$$

Note that $b < a$, so p does not divide the numerator after this cancellation process, observing that $b < a \implies k < p^a$. So after the cancellation for the product, p does not divide the numerator. Thus $p \nmid |\Omega|$, as required.

Now count the elements of Ω,

$$\begin{aligned}|\Omega| &= \text{sum of sizes of orbits} \\ &= \text{sum of sizes of orbits of type (a)} \\ &\quad + \text{sum of sizes of orbits of type (b).}\end{aligned}$$

But $p \nmid |\Omega|$ and p divides the size of orbits of type (b), thus we do have orbits of type (a).

Consider the orbit-stabiliser theorem for an orbit Σ of type (a),

$$m = |\Sigma| = \frac{|G|}{|P|} = \frac{p^a m}{|P|}$$

where P is the stabiliser of an element chosen in Σ. So $|P| = p^a$ and thus there is a subgroup of order p^a. □

(ii) All Sylow p-subgroups are conjugate.

In fact, we will prove the following. If P is a Sylow p-subgroup and $Q \leq G$ with $|Q| = p^b$ for $1 \leq b \leq a$, then there exists $g_1 \in G$ such that $Q \leq g_1 P g_1^{-1}$. In particular, if Q is of order p^a then $Q = g_1 P g_1^{-1}$.

Proof. This time consider the action of Q on the set of left cosets of P via left multiplication,

$$g * g_1 P = g g_1 P$$

for $g \in Q, g_1 \in G$. We consider the orbits. By the orbit-stabiliser theorem, the size of an orbit divides $|Q|$, thus is either 1 or a power of p.

The number of cosets is equal to the index of P in G,

$$\frac{|G|}{|P|} = \frac{p^a m}{p^a} = m.$$

In particular, m is equal to the number of orbits of size 1 plus p^k times the number of orbits of size k.

Counting cosets, we see that we must have an orbit of size 1. Suppose $\{g_1 P\}$ is such an orbit. Thus $g g_1 P = g_1 P$ for all $g \in Q$ and hence $g_1^{-1} g g_1 P = P$. So $g_1^{-1} g g_1 \in P$ and $g \in g_1 P g_1^{-1}$. Thus $Q \leq g_1 P g_1^{-1}$, as required. □

Lecture 7

(iii) The number n_p of Sylow p-subgroups satisfies $n_p \equiv 1 \pmod{p}$ and $n_p \mid |G|$.

Last time we observed that n_p is the index of $N_G(P)$, the normaliser of a Sylow p-subgroup P, in G and so $n_p \mid |G|$.

Proof. Consider the action of P on the set of Sylow p-subgroups of G via conjugation. By the orbit-stabiliser theorem, orbits are of size 1 or of size divisible by p. We need to show that there is exactly one orbit of size 1.

Visibly $\{P\}$ is an orbit of size 1. Are there any others?

Suppose $\{Q\}$ is an orbit of size 1. Thus $gQg^{-1} = Q$ for all $g \in P$ and so $P \leq N_G(Q)$. But we observed last time that the normaliser of a Sylow subgroup has a unique Sylow subgroup. (This is a consequence of Sylow's second theorem applied to the normaliser.) So $P = Q$, since both are Sylow subgroups of $N_G(Q)$.

Thus $\{P\}$ is the only orbit of size 1.

Counting Sylow p-subgroups in G,

$$n_p = 1 + \text{sum of sizes of other orbits.}$$

But p divides the size of every other orbit. Therefore, $n_p \equiv 1 \pmod{p}$. □

Recall the definition of a soluble group, a group for which the composition factors can be chosen to be *abelian* simple.

Theorem (1904, Burnside)**.** If $|G| = p^a q^b$, where p and q are primes, then G is soluble.

Theorem (1937, Hall)**.** G is soluble if and only if whenever $|G|$ factorises as $|G| = mn$ with m and n coprime there is a subgroup of order m.

Theorem (1963, Feit, Thompson, "Odd Order Theorem")**.** If $|G|$ is odd then G is soluble.

Chapter 2

Rings

2.1 Definitions and examples

Definition (Ring)**.** A set R with two operations, $+$ and $\cdot$, forms a *ring* if

(i) $(R, +)$ is an abelian group,

(ii) multiplication is associative and there is a multiplicative identity $1_R \in R$ such that

$$\forall r \in R \quad 1_R \cdot r = r \cdot 1_R = r,$$

(iii) multiplication is distributive over addition, that is

$$r_1(r_2 + r_3) = r_1 r_2 + r_1 r_3,$$
$$(r_1 + r_2)r_3 = r_1 r_3 + r_2 r_3.$$

In this course all rings are assumed to be commutative, i.e. multiplication is commutative.

Definition (Subring)**.** A subset S of R is a *subring* if it is a ring under the restriction of the two operations. In particular, $1_R \in S$. The notation is $S \leq R$.

Example. We have the four nested rings $\mathbb{Z} \leq \mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}$. In the introduction we also met the Gaussian integers $\mathbb{Z}[i] \leq \mathbb{C}$. Also $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\} \leq \mathbb{R}$.

Remark. All fields are rings, e.g. $\mathbb{Z}$ modulo n, the set $\{0, \ldots, n-1\}$ with addition and multiplication modulo n.

Definition (Unit)**.** A *unit* of R is an element with a multiplicative inverse in R, e.g. 2 is a unit in $\mathbb{Q}$, but not a unit in $\mathbb{Z}$.

Warning. There are various conventions.

(i) Not everyone insists on a multiplicative identity 1_R.

(ii) Some people confusingly talk about 1_R as *the* unit of R.

Example. The zero ring $\{0\}$. In this case the multiplicative identity is the same as the additive one. But in all other rings $0 \neq 1$.

$$(0 + 0)r = 0r$$
$$(0 + 0)r = 0r + 0r$$

so $0r = 0$ for all $r \in R$. Hence 0 is *not* the multiplicative identity if R has more than one element.

Example. Let R be a ring. A *polynomial* f over R is of the form

$$f(X) = a_n X^n + \cdots + a_1 X + a_0$$

with $a_i \in R$. The *degree* of f is the largest n with $a_n \neq 0$. f is *monic* if $a_n = 1$, where n is the degree of f. $R[X]$ is the polynomial ring over R with ring operations

$$(f+g)(X) = \sum_i (a_i + b_i) X^i,$$

$$(fg)(X) = \sum_i \left(\sum_{j=0}^{i} a_j b_{i-j} \right) X^i$$

where $f(X) = \sum a_i X^i$, $g(X) = \sum b_i X^i$.

R is a subring of $R[X]$ by identifying R with the constant polynomials $f(X) = \sum a_i X^i$ with $a_i = 0$ for all $i \geq 1$.

$R[[X]]$ is the ring of *formal power series*

$$f(X) = a_0 + a_1 X + a_2 X^2 + \cdots$$

with the same addition and multiplication as for $R[X]$ but note that there is no restriction demanding that all but finitely many coefficients be zero.

Laurent polynomials,

$$f(X) = \sum_{i \in \mathbb{Z}} a_i X^i,$$

$$(fg)(X) = \sum_i \left(\sum_j a_j b_{i-j} \right) X^i$$

with the condition that all but finitely many a_i are zero, e.g. $X^{-1} + X$.

Laurent series,

$$\sum_{i \in \mathbb{Z}} a_i X^i$$

with the condition that all but finitely many of the a_i for $i \leq 0$ are zero. Addition and multiplication are defined in the same way as for Laurent polynomials.

Another example is given by the ring of all R-valued functions on a set A, $f \colon A \to R$. The ring operations can be defined by pointwise addition and multiplication,

$$(f+g)(a) = f(a) + g(a),$$
$$(fg)(a) = f(a)g(a).$$

For example, the set of continuous functions $f \colon \mathbb{R} \to \mathbb{R}$ forms a subring of the ring of all functions $\mathbb{R} \to \mathbb{R}$. The set of continuous functions $\mathbb{R} \to \mathbb{R}$ contains the subring of *polynomial functions*, $f \colon \mathbb{R} \to \mathbb{R}$, $a \mapsto f(a)$, where $f(X) \in \mathbb{R}[X]$.

We can similarly consider Laurent polynomial functions, power series functions $\mathbb{C} \to \mathbb{C}$ and Laurent series functions $\mathbb{C} \to \mathbb{C}$.

Lecture 9

2.2 Homomorphisms, ideals, quotient rings and isomorphisms

Definition (Ring homomorphism)**.** A map $\theta\colon R \to S$ is a *ring homomorphism* if

(i) $\theta(r_1 + r_2) = \theta(r_1) + \theta(r_2)$,

(ii) $\theta(r_1 r_2) = \theta(r_1)\theta(r_2)$,

(iii) $\theta(1_R) = 1_S$.

A bijective homomorphism is called an *isomorphism*. The kernel of θ is $\ker\theta = \{r \in R : \theta(r) = 0\}$.

Lemma 2.1. θ is injective if and only if $\ker\theta = \{0\}$.

Proof. θ is a homomorphism of the additive groups. □

Definition (Ideal)**.** A subset I of R is an *ideal*, written $I \lhd R$, if

(i) I is a subgroup of R under addition,

(ii) whenever $a \in I$ and $r \in R$ then $ar \in I$.

The second condition is called the strong closure property.

Warning. An ideal in general is not a subring. If $1 \in I$ then $r = 1r \in I$ for all $r \in R$ and so $I = R$ if $1 \in I$. So the only ideal which is a subring of R is R itself.

Lemma 2.2. The kernel of a ring homomorphism $\theta\colon R \to S$ is an ideal.

Proof. θ is a homomorphism of the additive groups $(R, +) \to (S, +)$ and so $\ker\theta$ is an additive subgroup of R.

If $a \in \ker\theta$ then $\theta(a) = 0$. Consider

$$\theta(ar) = \theta(a)\theta(r) = 0.\theta(r) = 0.$$

Thus $ar \in \ker\theta$, so we have the strong closure property. □

Example. (i) In a field $\mathbb{F}$, the only ideals are $\{0\}$ and $\mathbb{F}$. (Multiply any non-trivial element in the ideal by its inverse to see that 1 is in the ideal.)

(ii) In $\mathbb{Z}$ the ideals are of the form

$$n\mathbb{Z} = \{\ldots, -2n, -n, 0, n, 2n, \ldots\}.$$

Proof. Certainly each set $n\mathbb{Z}$ is an ideal of $\mathbb{Z}$.

Suppose I is a non-zero ideal. We show that all the non-zero additive subgroups of $\mathbb{Z}$ are of the form $n\mathbb{Z}$. Let n be the least positive element of I and use Euclid's algorithm. If $b \in I$ then $b = nq + r$ with $0 \leq r < n$ and some q. Note that $r = b - nq \in I$. Minimality of n implies that $r = 0$ and hence $b = nq$. Thus $I = n\mathbb{Z}$. □

Definition. Let R be a ring and $a \in R$. Then the *ideal generated by* a is $aR = \{ar : r \in R\}$. We often use the notation (a) for aR.

Note that this is the smallest ideal of R containing a. Such an ideal is called a *principal ideal*.

More generally, the ideal generated by $a_1, \ldots, a_k$ is

$$\begin{aligned}(a_1, \ldots, a_k) &= a_1R + \ldots a_kR \\ &= \left\{ \sum_{i=1}^{k} a_i r_i : r_i \in R \right\}.\end{aligned}$$

Even more generally, the ideal generated by a subset A of R is

$$(A) = \left\{ \sum_{a \in A} ar_a : \text{only finitely many } r_a \in R \text{ are non-zero } \right\}.$$

Examples of principal ideals include $n\mathbb{Z}$ in $\mathbb{Z}$ and (X) in $\mathbb{C}[X]$, where (X) is the ideal of polynomials with zero constant term.

Proposition 2.3. Let $I \lhd R$ be an ideal. Then the *quotient ring* R/I has elements consisting of the cosets $r + I$ with the operations

$$\begin{aligned}(r_1 + I) + (r_2 + I) &= (r_1 + r_2) + I, \\ (r_1 + I).(r_2 + I) &= r_1 r_2 + I.\end{aligned}$$

This defines a ring.

Proof. From Proposition 1.3 we have already checked that R/I forms a group under addition as defined. So we need to check that multiplication is well-defined. If

$$r_1 + I = r_1' + I, \qquad r_2 + I = r_2' + I$$

then

$$r_1' = r_1 + a_1, \qquad r_2' = r_2 + a_2$$

for some $a_1, a_2 \in I$, and therefore

$$\begin{aligned}r_1' r_2' &= (r_1 + a_1)(r_2 + a_2) \\ &= r_1 r_2 + (a_1 r_2 + r_1 a_2 + a_1 a_2)\end{aligned}$$

where the second term is contained in I by the strong multiplicative closure property. Thus $r_1' r_2' + I = r_1 r_2 + I$.

The multiplicative identity is $1 + I$ since $(1 + I)(r + I) = r + I$.

Associativity, closure under multiplication and distributivity follow from the respective properties in the ring R. (Check.) □

Example. (i) $n\mathbb{Z}$ is an ideal of $\mathbb{Z}$. Elements $m + n\mathbb{Z}$ of $\mathbb{Z}/n\mathbb{Z}$ may be expressed as one of

$$0 + n\mathbb{Z}, 1 + n\mathbb{Z}, \ldots, (n-1) + n\mathbb{Z}.$$

Addition and multiplication correspond to arithmetic modulo n.

Lecture 9

(ii) Consider the ideal $I = (X) \lhd \mathbb{C}[X]$. Elements $f(X) + I$ of the quotient ring $\mathbb{C}[X]/I$ may be expressed in the form $a + I$ where $a \in \mathbb{C}$ is the constant term of $f(X)$. Addition and multiplication correspond to that in $\mathbb{C}$,

$$(a + I) + (b + I) = (a + b) + I,$$
$$(a + I)(b + I) = ab + I.$$

Therefore, $\mathbb{C}[X]/I \cong \mathbb{C}$.

Proposition 2.4 (Euclid's algorithm for polynomials)**.** Let $\mathbb{F}$ be a field and consider polynomials $f(X), g(X) \in \mathbb{F}[X]$. Then we can write

$$f(X) = g(X)q(X) + r(X)$$

with $\deg r < \deg g$.

Proof. Write

$$f(X) = \sum_{i=0}^{n} a_i X^i, \qquad g(X) = \sum_{i=0}^{m} b_i X^i$$

with $\deg f(X) = n$ and $\deg g(X) = m$.

If $n < m$ then set $q(X) = 0$ and $r(X) = f(X)$. Now assume $n \geq m$, and argue by induction on n. Set $f_1(X) = f(X) - a_n b_m^{-1} X^{n-m} g(X)$ so $\deg f_1(X) < \deg f(X)$. If $m = n$ then $\deg f_1(X) < m$ and we can apply our first case to $f_1(X)$,

$$f(X) = a_n b_m^{-1} g(X) + f_1(X).$$

If $n > m$ then we have by induction

$$f_1(X) = g(X)q_1(X) + r_1(X)$$

for suitable $q_1(X)$ and $r_1(X)$, so

$$f(X) = g(X)(q_1(X) + a_n b_m^{-1} X^{n-m}) + r_1(X). \qquad \square$$

Note 2. We require $\mathbb{F}$ to be a field as we need b_m^{-1}.

Example. Consider the ideal

$$I = (X^2 + 1) \lhd \mathbb{R}[X].$$

For any $f(X) \in \mathbb{R}[X]$, we can write $f(X) = (X^2 + 1)q(X) + r(X)$ with $\deg r(X) \leq 1$. Thus $f(X) + I = r(X) + I$. The elements of $\mathbb{R}[X]/I$ are of the form $a + bX + I$.

Addition takes the form

$$(a_1 + b_1 X + I) + (a_2 + b_2 X + I) = (a_1 + a_2) + (b_1 + b_2)X + I,$$

and similarly multiplication is given by

$$\begin{aligned}(a_1 + b_1 X + I)(a_2 + b_2 X + I) &= a_1 a_2 + (b_1 a_2 + a_1 b_2)X + b_1 b_2 X^2 + I \\ &= (a_1 a_2 - b_1 b_2) + (b_1 a_2 + a_1 b_2)X + I.\end{aligned}$$

This corresponds to addition and multiplication in $\mathbb{C}$. In fact,

$$\mathbb{R}[X]/(X^2 + 1) \cong \mathbb{C},$$
$$a + bX + I \mapsto a + bi.$$

Theorem 2.5 (First isomorphism theorem)**.** Let $\theta\colon R \to S$ be a ring homomorphism. Then $\ker\theta \lhd R$ and $R/\ker\theta \cong \operatorname{Im}\theta \leq S$.

Example. The map

$$\theta\colon \mathbb{R}[X] \to \mathbb{C}, \sum a_j X^j \mapsto \sum a_j i^j$$

is a ring homomorphism with $\ker\theta = (X^2+1)$ and $\operatorname{Im}\theta = \mathbb{C}$.

Proof. Lemma 2.2 states that $\ker\theta$ is an ideal. $\operatorname{Im}\theta$ is a subring of S,

(i) θ is a homomorphism of additive groups and so Lemma 1.4 implies that $\operatorname{Im}\theta$ is an additive subgroup of S.

(ii) $\theta(r_1)\theta(r_2) = \theta(r_1 r_2) \in \operatorname{Im}\theta$, so we have closure under multiplication. Associativity is inherited from S. Moreover, $\theta(1_R) = 1_S$.

Let $\Phi\colon R/I \to \operatorname{Im}\theta, r+I \mapsto \theta(r)$ and note $I = \ker\theta$. We know that Φ is well-defined and bijective and an isomorphism of additive groups from proof of Lemma 1.4. It is left to check that Φ is a ring homomorphism.

$$\begin{aligned}\Phi((r_1+I)(r_2+I)) &= \Phi(r_1 r_2 + I)\\ &= \theta(r_1 r_2)\\ &= \theta(r_1)\theta(r_2)\\ &= \Phi(r_1+I)\Phi(r_2+I),\\ \Phi(1_R+I) &= \theta(1_R) = 1_S.\end{aligned}$$

Thus Φ is a ring homomorphism. □

Theorem 2.6 (Second isomorphism theorem)**.** Let R be a subring of S and $J \lhd S$ an ideal. Then $R \cap J \lhd S$ is an ideal and $\{r+J : r \in R\} = (R+J)/J \leq S/J$ and $R/(R\cap J) \cong (R+J)/J \leq S/J$.

Proof. Let

$$\theta\colon R \to S/J, r \mapsto r+J.$$

This is a ring homomorphism. (Check.) We have that

$$\begin{aligned}\ker\theta &= \{r \in R : r+J = J\} = R \cap J \lhd R,\\ \operatorname{Im}\theta &= \{r+J : r \in R\} = (R+J)/J \leq S/J.\end{aligned}$$

Theorem 2.5 implies that $R/\ker\theta \cong \operatorname{Im}\theta$, that is, $R/(R\cap J) \cong (R+J)/J \leq S/J$. □

There is a correspondence between additive groups,

$$\left\{\begin{array}{c}\text{additive subgroups of } R\\ \text{containing ideal } I\end{array}\right\} \longleftrightarrow \left\{\begin{array}{c}\text{additive subgroups}\\ \text{of } R/I\end{array}\right\}$$

as in Chapter 1. This correspondence induces

$$\left\{\begin{array}{c}\text{subrings of } R\\ \text{containing } I\end{array}\right\} \longleftrightarrow \left\{\begin{array}{c}\text{subrings}\\ \text{of } R/I\end{array}\right\},$$

Lecture 10

$$\left\{\begin{matrix}\text{ideals of } R\\ \text{containing } I\end{matrix}\right\} \longleftrightarrow \left\{\begin{matrix}\text{ideals}\\ \text{of } R/I\end{matrix}\right\}.$$

(This needs to be checked.)

Theorem 2.7 (Third isomorphism theorem)**.** Let I and J be ideals of R, with $I \subset J$. Then $(R/I)/(J/I) \cong R/J$ where $J/I = \{r + I : r \in J\}$.

Proof. Let θ be the map

$$\theta\colon R/I \to R/J, r + I \mapsto r + J.$$

This is well-defined as in Theorem 1.7. One can check it is a ring homomorphism and $\theta(1 + I) = 1 + J$.

The first isomorphism theorem implies that $R/\ker\theta \cong \operatorname{Im}\theta \leq R/J$. Noting that

$$\begin{aligned}\ker\theta &= \{r + I : r \in J\} = J/I \lhd R/I,\\ \operatorname{Im}\theta &= R/J,\end{aligned}$$

we get the required result. □

Example. Given a ring R, there is a unique ring homomorphism

$$\begin{aligned}\phi\colon \mathbb{Z} &\to R\\ 1 &\mapsto 1_R\\ m &\mapsto \underbrace{1 + \cdots + 1}_{m \text{ times}}\end{aligned}$$

The first isomorphism theorem implies $\mathbb{Z}/\ker\phi \cong \operatorname{Im}\phi \leq R$. $\operatorname{Im}\phi$ is the *prime subring* of R. $\ker\phi \lhd \mathbb{Z}$ is an ideal and so is of the form $n\mathbb{Z}$ for some n.

This n is the *characteristic* of R. If R is one of $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$ then the characteristic is 0. If $R = \mathbb{Z}/(p\mathbb{Z})$ then the chacteristic is p.

2.3 Integral domains, fields of fractions, maximal ideals and prime ideals

Definition (Integral domain)**.** A ring is an *integral domain* if $ab = 0$ implies $a = 0$ or $b = 0$ for all $a, b \in R$. A *zero divisor* a is non-zero and there exists $b \neq 0$ such that $ab = 0$. (In an integral domain there are no zero divisors.)

Example. $\mathbb{Z}$ is an integral domain. All fields are integral domains. Subrings of integral domains are integral domains, e.g. $\mathbb{Z}[i] \leq \mathbb{C}$.

Definition (Principal ideal domain)**.** A *principal ideal domain* is an integral domain in which all ideals are principal ideals, i.e. of the form $(a) = aR$ for some $a \in R$.

Example. $\mathbb{Z}$ is a principal ideal domain. We will see that any integral domain where Euclid's algorithm applies is a principal ideal domain.

Lemma 2.8. Let R be a finite integral domain. Then R is a field.

Proof. Consider the map $R \to R, r \mapsto ar$, that is, multiplication by $a \neq 0$.

Since R is an integral domain this map is injective. (If $ar_1 = ar_2$ then $a(r_1 - r_2) = 0$ and so $r_1 - r_2 = 0$, i.e. $r_1 = r_2$. Note that cancellation is valid in integral domains.) Since R is finite the map is also surjective.

Thus there exists $r \in R$ with $ar = 1$. Thus a has a multiplicative inverse. So R is a field. □

Lemma 2.9. If R is an integral domain then $R[X]$ is an integral domain.

Proof. If

$$f(X) = \sum a_i X^i, \qquad g(X) = \sum b_i X^i$$

with $\deg f = m$ and $\deg g = n$ then $f(X)g(X)$ has degree $m+n$ since $a_m b_n \neq 0$ as R is an integral domain. Therefore,

$$f(X) \neq 0, g(X) \neq 0 \implies f(X)g(X) \neq 0.$$

Thus $R[X]$ is an integral domain. □

Remark. By induction, $R[X_1, \ldots, X_n]$ polynomial ring in indeterminates $X_1, \ldots, X_n$ is an integral domain if R is. This is because $R[X_1, X_2]$ may be regarded as $(R[X_1])[X_2]$.

Theorem 2.10. Let R be an integral domain. Then there is a *field of fractions* $\mathbb{F}$ with the properties

(i) $R \leq \mathbb{F}$ is a subring,

(ii) $\mathbb{F}$ is a field,

(iii) every element of $\mathbb{F}$ has the form ab^{-1} where $a \in R$ and b^{-1} is the multiplicative inverse in $\mathbb{F}$ of $b \in R$.

Example. With the notation as above, $R = \mathbb{Z}$, $\mathbb{F} = \mathbb{Q}$.

Proof. Consider pairs (a, b) with $a \in R$, $b \in R$ and $b \neq 0$. Define an equivalence relation

$$(a, b) \sim (c, d) \iff ad = bc.$$

(For transitivity, note that $(a, b) \sim (c, d) \iff ad = bc, (c, d) \sim (e, f) \iff cf = de$ so $adf = bcf = bde$ as cancellation is valid in R. So $af = be$ and $(a, b) \sim (e, f)$.)

Write a/b for the equivalence class of (a, b). We define addition by

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

and multiplication by

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

One can easily check that these operations are well-defined.

Lecture 11

If we let $\mathbb{F} = \{a/b : a \in R, b \in R, b \neq 0\}$ then $\mathbb{F}$ is a ring under these operations.

R may be identified with the subring of elements of the form $r/1$. The multiplicative identity is $1/1$. a/b has multiplicative inverse b/a if $a \neq 0$,

$$\frac{a}{b}\frac{b}{a} = \frac{ab}{ab} = \frac{1}{1}.$$

Every element a/b is of the form $(a/1)(b/1)^{-1}$ since $(b/1)^{-1} = 1/b$. □

Lemma 2.11. A non-zero ring R is a field if and only if its only ideals are $\{0\}$ and R.

Proof. Suppose that R is a field and take an ideal $I \triangleleft R$. If $I \neq \{0\}$ then pick $a \neq 0 \in I$. Then $1 = aa^{-1} \in I$. Hence $I = R$.

Conversely, assume that $\{0\}$ and R are the only ideals. Take $a \in R$ with $a \neq 0$. Then $(a) \neq \{0\}$ and so $(a) = R$. Hence $1 \in (a)$ and there exists r such that $ar = 1$. □

Definition (Maximal ideal)**.** An ideal I in R is *maximal* in R if $I \neq R$ but whenever $I \subset J \triangleleft R$ then $I = J$ or $J = R$.

Lemma 2.12. For an ideal $I \triangleleft R$, the quotient ring R/I is a field if and only if I is maximal in R.

Proof. R/I is a field if and only if I/I and R/I are the only ideals in R/I. By the usual correspondence, this is equivalent to that the only ideals in R containing I are I and R, which is the statement that I is maximal in R. □

Definition (Prime ideal)**.** An ideal I of R is a *prime ideal* in R if whenever $ab \in I$ then $a \in I$ or $b \in I$.

Lemma 2.13. For an ideal $I \triangleleft R$, the quotient ring R/I is an integral domain if and only if I is a prime ideal in R.

Proof. Suppose R/I is an integral domain and $ab \in I$. Then $I = ab + I = (a + I)(b + I)$ and so either $a + I = I$ or $b + I = I$. So $a \in I$ or $b \in I$.

Conversely, suppose I is prime and $(a + I)(b + I) = I$. Then $ab + I = I$ and so $ab \in I$. Primeness implies $a \in I$ or $b \in I$ and so $a + I = I$ or $b + I = I$. □

Example. In $\mathbb{Z}$, the maximal ideals are $p\mathbb{Z}$ where p is a prime number. Prime ideals are $p\mathbb{Z}$ for p prime and $\{0\}$.

Lemma 2.14. The characteristic of an integral domain is 0 or a prime number p.

Proof. Recall that we have a unique ring homomorphism $\phi\colon \mathbb{Z} \to R$. The prime subring $\operatorname{Im}\phi \cong \mathbb{Z}/n\mathbb{Z}$, where n is the characteristic of R. If n properly factorises as $n = st$ then working modulo n we have $st = 0 \pmod{n}$ and so we have zero divisors.

So if R is an integral domain, its prime subring is an integral domain and so n is prime or 0. □

2.4 Factorisation in integral domains — units, primes, irreducibles

Throughout this section R is assumed to be an integral domain.

Definition. An element $a \in R$ is a *unit* if it has a multiplicative inverse. Equivalently, $(a) = R$. We say a *divides* b, written $a \mid b$, if there is $c \in R$ such that $b = ac$. Equivalently, $(b) \subset (a)$. Elements a and b are *associates* in R if $a = bc$ for some unit $c \in R$. Equivalently, $(a) = (b)$. $r \in R$ is *irreducible* in R if it is non-zero, not a unit and whenever $r = ab$ with $a, b \in R$ then a or b is a unit. $r \in R$ is *prime* in R if it is non-zero, not a unit and if $r \mid ab$ then $r \mid a$ or $r \mid b$.

Remark. These definitions do depend on the ring R, e.g. 2 is prime and irreducible in $\mathbb{Z}$, but not in $\mathbb{Q}$. $2X$ is irreducible in $\mathbb{Q}[X]$ but not in $\mathbb{Z}[X]$, as $2X = 2 \cdot X$ and 2, X are not units in $\mathbb{Z}[X]$.

Lemma 2.15. (r) is a prime ideal of R if and only if r is prime or $r = 0$.

Proof. If $r \neq 0$ and $r \mid ab$ then $ab \in (r)$ and so if (r) is a prime ideal then $a \in (r)$ or $b \in (r)$. Thus $r \mid a$ or $r \mid b$.

(0) is a prime ideal in an integral domain. If r is prime and $ab \in (r)$ then $r \mid ab$ and hence $r \mid a$ or $r \mid b$. Thus $a \in (r)$ or $b \in (r)$. □

Lemma 2.16. If r is prime in R then r is irreducible in R.

Proof. Suppose that r is prime in R and $r = ab$ with $a, b \in R$. Then $r \mid ab$ and so $r \mid a$ or $r \mid b$. Without loss of generality suppose $r \mid a$. So $a = qr$ for some $q \in R$. So $r = qrb$. Cancellation in integral domains gives $1 = qb$, so b is a unit. □

Example. To show that the converse of Lemma 2.16 does not hold in general, consider $R = \mathbb{Z}[\sqrt{-5}] \leq \mathbb{C}$. R is an integral domain since it is a subring of a field. Define a norm,

$$N(a + b\sqrt{-5}) = a^2 + 5b^2 \in \mathbb{Z}_{\geq 0}.$$

Thus $N(z) = z\bar{z}$. N is multiplicative, that is, $N(z_1 z_2) = N(z_1)N(z_2)$. The units of R are precisely the elements of norm 1, namely ± 1. Suppose z is a unit and so there exists z_1 with $zz_1 = 1$ so $N(z)N(z_1) = N(zz_1) = N(1) = 1$ hence $N(z)$ is a unit in $\mathbb{Z}$ and also $N(z) \geq 0$. So $N(z) = 1$ is the only possibility.

There are no elements in R of norm 2 or 3. (We cannot solve $a^2 + 5b^2 = 2$ or 3.)

Consider the identity $6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$ in R.

To see that 2 is irreducible, express $2 = z_1 z_2$ and consider norms, $4 = N(z_1)N(z_2)$. Since there are no elements of norm 2 one of the $N(z_j) = 1$ so z_j is a unit.

But 2 is not prime in R since $2 \mid (1 + \sqrt{-5})(1 - \sqrt{-5})$ but 2 does not divide either $1 \pm \sqrt{-5}$. (Consider norms, $N(2) = 4$, $N(1 \pm \sqrt{-5}) = 6$ and $4 \nmid 6$.)

Similarly, 3, $1 + \sqrt{-5}$, $1 - \sqrt{-5}$ are irreducible.

Definition (Euclidean domain)**.** An integral domain R is a *Euclidean domain* (ED) if there is a function $\phi \colon R - \{0\} \to \mathbb{Z}_{\geq 0}$ called Euclidean function such that

(i) $\phi(ab) \geq \phi(a)$ for all $a, b \in R \setminus \{0\}$,

(ii) if $a, b \in R$ with $b \neq 0$ then there exist $q, r \in R$ with $a = qb + r$ with either $r = 0$ or $\phi(r) < \phi(b)$. (Euclid's algorithm.)

Example. (i) $\mathbb{Z}$ is a Euclidean domain with $\phi(n) = |n|$.

(ii) $\mathbb{F}[X]$ with $\mathbb{F}$ a field is a Euclidean domain with $\phi(f(X)) = \deg f$. (Section 2.4 gives Euclid's algorithm here.)

(iii) $R = \mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\} \leq \mathbb{C}$ is an integral domain. As with $\mathbb{Z}[\sqrt{-5}]$ we can define a norm via $N(z) = z\bar{z}$ for $z \in R$. Thus $N(a + bi) = a^2 + b^2$. In this case the norm is a Euclidean function.

N is multiplicative and so property (i) holds,

$$N(z_1 z_2) = N(z_1)N(z_2) \geq N(z_1),$$

since $N(z_2) \geq 1$. To verify property (ii) take $z_1, z_2 \in R$ with $z_2 \neq 0$. Consider $z_1/z_2 \in \mathbb{C}$. Then in the complex plane it is distance less than 1 from the nearest element of R.

$z_1/z_2 = q + z_3$ with $q \in R$, $z_3 \in \mathbb{C}$ and $|z_3| < 1$. So $z_1 = qz_2 + z_2 z_3$. Set $r = z_2 z_3$. Thus $z_1 = qz_2 + r$.

$$N(r) = |z_2 z_3|^2 = |z_2|^2 |z_3|^2 < |z_2|^2 = N(z_2).$$

Remark. Similar arguments apply for other subrings of $\mathbb{C}$ and one can sometimes show that the norm is a Euclidean function but one needs that any point in the complex plane is distance less than 1 from a lattice point of R. Note for $\mathbb{Z}[\sqrt{-5}]$ this is not true.

Proposition 2.17. If an integral domain R is an ED then R is a PID.

Proof. Let R be an ED with Euclidean function $\Phi\colon R \setminus \{0\} \to \mathbb{Z}_{\geq 0} = \{n \in \mathbb{Z} : n \geq 0\}$. Let $I \triangleleft R$ be an ideal and suppose that I is non-zero. Pick $b \in I \setminus \{0\}$ with $\Phi(b)$ minimal. Then take $a \in I$ and use Euclid's algorithm: $a = bq + r$ with $r = 0$ or $\Phi(r) < \Phi(b)$. Note that $r = a - bq \in R$, and hence minimality of $\Phi(b)$ implies $r = 0$. Thus $a = bq$, so $I = (b)$. □

Example. $\mathbb{Z}$, $\mathbb{F}[X]$ for a field $\mathbb{F}$ and $\mathbb{Z}[i]$ are all PIDs. But $\mathbb{Z}[X]$ is not a PID.

To see this, consider the ideal $I = (2, X) = \left\{\sum a_i X^i : a_0 \text{ is even}\right\} \triangleleft \mathbb{Z}[X]$. Suppose that $I = (f(X))$ for some $f(X) \in \mathbb{Z}[X]$, in particular $2 = f(X)g(X)$ for some $g(X) \in \mathbb{Z}[X]$. We deduce that $f(X)$ has to be a constant polynomial, that is, of degree 0, and must be ± 1 or ± 2.

Also $X = f(X)h(X)$ for some $h(X) \in \mathbb{Z}[X]$. So $f(X)$ cannot be ± 2. So $f(X) = \pm 1$. Hence $(f(X)) = \mathbb{Z}[X]$. This is a contradiction, since $I \neq \mathbb{Z}[X]$, e.g. $1 \notin I$.

Example (Minimal polynomials of matrices with entries in a field $\mathbb{F}$)**.** Let $\Delta \in M_{n\times n}(\mathbb{F})$. Consider $I = \{f(X) \in \mathbb{F}[X] : f(\Delta) = 0\}$.

We can check that $I \triangleleft \mathbb{F}[X]$ is an ideal. But $\mathbb{F}[X]$ is a PID, so $I = (m(X))$, an by multiplying $m(X)$ by a unit in $\mathbb{F}$ we may assume $m(X)$ is monic. This is the *minimal polynomial* of Δ. (See the course *Linear Algebra* where Euclid's algorithm is explicitly used.)

Definition (Unique factorisation domain)**.** An integral domain is a *unique factorisation domain* (UFD) if

(i) every non-zero element that is not a unit may be expressed as a product of finitely many irreducibles,

(ii) whenever $p_1 \cdots p_m = q_1 \cdots q_n$ for products of irreducibles then $m = n$ and we can reorder so that p_i is an associate of q_i. (So factorisation is unique up to ordering and associates).

Our goal is the following proposition.

Proposition 2.18. If an integral domain R is a PID then R is a UFD.

Corollary 2.19. $\mathbb{Z}$, $\mathbb{F}[X]$ for a field $\mathbb{F}$ and $\mathbb{Z}[i]$ are all UFDs.

Example. $\mathbb{Z}[\sqrt{-5}]$ is not a UFD. Note that

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$$

where 2, 3, $1 + \sqrt{-5}$ and $1 - \sqrt{-5}$ are irreducibles. Units are ± 1 in $\mathbb{Z}[\sqrt{-5}]$, so these irreducibles are not associates of each other.

We need two lemmas about PIDs to prove Proposition 2.18.

Lemma 2.20. Let R be a PID. If p is an irreducible then it is prime. (This is the converse of Lemma 2.16 for PIDs.)

Proof. Let p be an irreducible and suppose $p \mid ab$. Suppose $p \nmid a$. The ideal (p, a) must be principle in a PID, $(p, a) = (d)$, say.

Thus $p = q_1 d$ and $a = q_2 d$ for some $q_1, q_2 \in R$. p is irreducible so either q_1 or d is a unit. But if q_1 is a unit, then $d = q_1^{-1} p$ and $a = q_2 d = q_2 q_1^{-1} p$, so $p \mid a$, a contradiction. Hence d is a unit. Thus $(d) = R$ and there exist $r, s \in R$ such that

$$1 = rp + sa$$

and so

$$b = rpb + sab.$$

Therefore, $p \mid b$. □

Remark. In fact, d is a highest common factor of p and a. Highest common factors are unique up to units, see Example Sheet 3.

Lemma 2.21 (Ascending chain condition, ACC)**.** Let R be a PID and $I_j \triangleleft R$ with $I_1 \subset I_2 \subset I_3 \subset \cdots$. Then, for some $n \in \mathbb{N}$, $I_n = I_{n+i}$ for all $i \geq 0$.

Proof. The union $I = \bigcup_j I_j \triangleleft R$ is an ideal, check. (For example, if $a \in I_j$ and $b \in I_k$, $j \leq k$ then $a \in I_k$, $B \in I_k$, so $a + b \in I_k \subset I$.) R is a PID, so $I = (a)$ for some $a \in I$. We must have $a \in I_n$ for some $n \in \mathbb{N}$. For all $i \geq 0$,

$$(a) \supset I \supset I_{n+i} \supset I_n \supset (a),$$

so we have equality throughout and in particular $I_{n+i} = I_n$. □

Lecture 13

Remark. Rings satisfying the ACC are called *Noetherian*.

Proof of (2.18). (i) Let $a \in R$, non-zero and not a unit. Assume it cannot be factorised as a product of finitely many irreducibles. So in particular a itself is not irreducible, hence $a = a_1 b_1$ with a_1, b_1 non-zero and not units. We may assume that a_1 cannot be factorised as a product of finitely many irreducibles, so write $a_1 = a_2 b_2$ and continue. (If both a_1, b_1 can be factorised into finitely many irreducibles, then so can a.)

We obtain $(a_1) \subsetneq (a_2) \subsetneq (a_3) \subsetneq \cdots$ with inequality at each stage, for if $(a_i) = (a_{i+1})$ then a_i and a_{i+1} would be associates and b_{i+1} a unit.

(ii) Suppose we have $p_1 \cdots p_m = q_1 \cdots q_n$ with products of irreducibles. Then p_1 is prime by Lemma 2.20 and $p_1 \mid q_1 \cdots q_n$, so p_1 divides one of the q_i, $p_1 \mid q_1$, say, and so $q_1 = p_1 u$. But q_1 is irreducible and p_1 is not a unit, hence u is a unit. This is to say p_1 and q_1 are associates. We have

$$p_1 \cdots p_m = u p_1 q_2 \cdots q_n.$$

Cancellation in R gives

$$p_2 \cdots p_m = (u q_2) q_3 \cdots q_n.$$

Note that (uq_2) is irreducible. Continuing in this way gives the result. □

Remark. The proof of (ii) depends just on Lemma 2.20, that irreducibles are prime, and (i) depends on the ACC.

There are several properties which follow quickly from the definition of a UFD.

(i) An element $p \in R$ is irreducible if and only if it is prime.

We have already shown the 'only if' direction for integral domains.

Suppose p is an irreducible and $p \mid ab$. We can express a and b as

$$a = p_1 \cdots p_m, \quad b = q_1 \cdots q_n$$

with p_i, q_i irreducible. Now write $ab = pc$ with $c = r_1 \cdots r_s$ where r_i is irreducible. Then

$$(p_1 \cdots p_m)(q_1 \cdots q_n) = p r_1 \cdots r_s$$

and uniqueness of factorisation implies that p is an associate of either some p_i or some q_i. Hence $p \mid a$ or $p \mid b$.

(ii) Highest common factors exist.

(iii) Lowest common multiples exist.

Definition (Highest common factor)**.** d is a *highest common factor* of $a_1, \ldots, a_n$, written $\mathrm{hcf}(a_1, \ldots, a_n)$, if $d \mid a_i$ for each i and whenever d' also divides each a_i then $d' \mid d$.

Note that one often refers to *the* highest common factor, although certainly multiplying by a unit will give another highest common factor.

Expressing each a_i as a product of irreducibles, each a_i is of the form

$$u_i \prod_j p_j^{n_{ij}}$$

where p_j is irreducible and p_j is not associate to p_k whenever $j \neq k$. u_i is a unit in R and $n_{ij} \geq 0$.

The claim is that $\prod_j p_j^{m_j}$ is a highest common factor of $a_1, \ldots, a_n$ where $m_j = \min_i\{n_{ij}\}$. It is clear that this is a factor of each a_i, and if $d' \mid a_i$ for each i then each irreducible dividing d' is also an irreducible dividing a_i and so must be one of the p_j. Thus

$$d' = u \prod_j p_j^{t_j}$$

where u is a unit. But the t_j can be at most $\min_i\{n_{ij}\}$ if $d' \mid a_i$ for each i. Thus $d' \mid d$.

2.5 Factorisations in polynomial rings, Gauss' lemma and Eisenstein's criterion

If $\mathbb{F}$ is a field then $\mathbb{F}[X]$ is a ED, PID and UFD. Every ideal $I \lhd \mathbb{F}[X]$ is principal, $I = (f(X))$ for some $f(X) \in \mathbb{F}[X]$. An element is irreducible if and only if it is prime. $\mathbb{F}[X]/I$ is a field, with elements of the form $r(X) + I$ with $r(X) = 0$ or $\deg r < \deg f$, if and only if I is maximal if and only if $f(X)$ is irreducible in $\mathbb{F}[X]$.

Gauss' lemma helps to determine when polynomials are irreducible in $\mathbb{F}[X]$.

Definition (Content)**.** Assume that the coefficient ring R is a UFD. Let $f(X) = a_0 + \cdots + a_n X^n$ with $a_n \neq 0$, $\deg f = n$. The *content* $c(f(X))$ is the highest common factor of $a_0, \ldots, a_n$. $f(X)$ is *primitive* if $c(f(X))$ is a unit, i.e. if the a_i are coprime.

Our aim is to prove the following lemma.

Lemma 2.22 (Gauss' lemma)**.** Let R be a UFD with $\mathbb{F}$ its field of fractions. Suppose $f(X) \in R[X]$ is primitive. Then $f(X)$ is irreducible in $R[X]$ if and only if $f(X)$ is irreducible in $\mathbb{F}[X]$.

In particular, when $R = \mathbb{Z}$ then $f(X)$ is irreducible in $\mathbb{Z}[X]$ if and only if it is irreducible in $\mathbb{Q}[X]$.

Example. $X^3 + X + 1$ is irreducible in $\mathbb{Z}[X]$, and is primitive in $\mathbb{Z}[X]$ and so Gauss' lemma implies $X^3 + X + 1$ is irreducible in $\mathbb{Q}[X]$, and so $\mathbb{Q}[X]/(X^3 + X + 1)$ is a field.

Suppose $X^3 + X + 1$ is reducible in $\mathbb{Z}[X]$, so $X^3 + X + 1 = g(X)h(X)$ for $g(X), h(X) \in \mathbb{Z}[X]$ not units. Primitivity implies that $g(X)$ and $h(X)$ are not constant polynomials, so $\deg g, \deg h \geq 1$. So one of $g(X)$ and $h(X)$ is of degree 1,

$$g(X) = b_0 + b_1 X,$$
$$h(X) = c_0 + c_1 X + c_2 X^2,$$

say. Considering coefficients of X^0 and X^3,

$$b_0 c_0 = 1,$$
$$b_1 c_2 = 1.$$

So $b_0 = \pm 1$, $b_1 = \pm 1$. This is a contradiction since ± 1 is not a root of $X^3 + X + 1$.

Lecture 14

Lemma 2.23. If $f(X)$ and $g(X)$ are primitive in $R[X]$ then so is $f(X)g(X)$.

Proof. Let

$$f(X) = a_0 + a_1X + \cdots + a_mX^m,$$
$$g(X) = b_0 + b_1X + \cdots + b_nX^n$$

be primitive. Suppose the product $f(X)g(X)$ is not primitive. So there is a prime $p \in R$ dividing $c(f(X)g(X))$ but $p \nmid c(f(X))$ and $p \nmid c(f(X))$.

Let k and l be such that

$$p \mid a_0, p \mid a_1, \ldots, p \mid a_{k-1}, p \nmid a_k,$$
$$p \mid b_0, p \mid b_1, \ldots, p \mid b_{l-1}, p \nmid b_l.$$

The coefficient c_{k+l} of X^{k+l} in $f(X)g(X)$ is

$$\cdots + a_{k+1}b_{l-1} + a_kb_l + a_{k-1}b_{l+1} + \cdots .$$

p divides all of these terms apart from a_kb_l which it does not divide. So $p \nmid c_{k+l}$, contradicting that $p \mid c(f(X)g(X))$. Thus $f(X)g(X)$ is primitive. □

Consider the polynomial ring $R[X]$ and let $\mathbb{F}$ be the field of fractions of R.

Corollary 2.24. For $f(X), g(X) \in R[X]$, the content $c(f(X)g(X))$ is an associate of $c(f(X))c(g(X))$. (Recall that highest common factors and hence contents are only defined up to associates.)

Proof. As R is a UFD we may write

$$f(X) = c(f(X))f_1(X),$$
$$g(X) = c(g(X))g_1(X)$$

where $f_1(X), g_1(X)$ are primitive. Then

$$f(X)g(X) = c(f(X))c(g(X))f_1(X)g_1(X)$$

where $f_1(X)g_1(X)$ is primitive by Lemma 2.23. So a highest common factor of the coefficients of $f(X)g(X)$ is $c(f(X))c(g(X))$. □

Remark. An irreducible $f(X)$ in $R[X]$ is either in R (and irreducible in R) or it is primitive in $R[X]$.

To see this, assume that $f(X)$ is irreducible and write $f(X) = c(f(X))f_1(X)$ with $f_1(X)$ primitive. If $f_1(X) \notin R$ then, since a polynomial of degree at least 1 cannot be a unit, $c(f(X))$ must be a unit.

Proof of Lemma 2.22, Gauss' lemma. Take $f(X)$ primitive in $R[X]$. Suppose it factorises in $R[X]$ as a product of non-units in $R[X]$. Since $f(X)$ is primitive, neither of these non-units is in R, the factors have degrees greater than 0. So $f(X)$ factorises as a product of non-units in $\mathbb{F}[X]$.

Conversely, suppose that $f(X) = g(X)h(X)$ with $g(X), h(X) \in \mathbb{F}[X]$ non-units and hence not constant. Multiply by $a \in R$ and $b \in R$ repectively such that $ag(X), bh(X) \in R[X]$. Thus

$$abf(X) = (ag(X))(bh(X)).$$

Write

$$ag(X) = c(ag(X))g_1(X),$$
$$bh(X) = c(bh(X))h_1(X)$$

where $g_1(X), h_1(X)$ are primitive in $R[X]$. Note that $g_1(X), h_1(X)$ are not constant and hence not units. Now Lemma 2.24 implies that ab is an associate of $c(ag(X))c(bh(X))$ by considering contents. So

$$abf(X) = uc(ag(X))c(bh(X))u^{-1}g_1(X)h_1(X)$$

where $ab = uc(ag(X))c(bh(X))$ and u a unit in R. Cancellation gives

$$f(X) = (u^{-1}g_1(X))h_1(X)$$

and thus $f(X)$ factorises as a product of non-units in $R[X]$. □

Remark. A similar argument shows that if $f(X) \in R[X]$, not necessarily primitive, and $f(X) = g_1(X)h(X)$ with $g_1(X)$ primitive in $R[X]$, $h(X) \in \mathbb{F}[X]$ then we can deduce that in fact

$$f(X) = g_1(X)h_0(X)$$

with $h_0(X) \in R[X]$.

We can see this as follows. There exists $b \in R$ with $bh(X) \in R[X]$. Consider $bf(X) = g_1(X)(bh(X))$. Write $bh(X) = c(bh(X))h_1(X)$ with $h_1(X)$ primitive. Consider contents, $b \mid c(g_1(X)bh(X))$ and so $b \mid c(bh(X))$ since $g_1(X)$ is primitive. Cancellation as in the previous result gives $f(X) = g_1(X)h_0(X)$ for some $h_0(X) \in R[X]$.

Thus if

$$I = g_1(X)\mathbb{F}[X] \lhd \mathbb{F}[X]$$

and

$$J = g_1(X)R[X] \lhd R[X]$$

then $I \cap R[X] = J$, with $g_1(X)$ primitive in $R[X]$.

Theorem 2.25. If R is a UFD then $R[X]$ is a UFD.

Proof. Let $f(X) \in R[X]$ be non-zero and a non-unit. Write $f(X) = c(f(X))f_1(X)$ with $f_1(X)$ primitive. (The first step is to show that we need only look at primitives.)

Observe that since R is a UFD, $c(f(X))$ is expressible as a product of irreducibles in R in an essentially unique way. These irreducibles are irreducible in $R[X]$.

If $f(X)$ factorises as a product of irreducibles in $R[X]$, then we can collect together the irreducibles in R and the primitive irreducibles. The product of these primitive

irreducibles is primitive by Lemma 2.23 and so the content of $f(X)$ is an associate of the product of the irreducibles in R. The essentially unique factorisation of $c(f(X))$ means that this product is essentially the same as the previous one.

Cancellation implies that the product of the primitives is an associate of $f_1(X)$. Thus we may assume that $f(X)$ is primitive.

$\mathbb{F}[X]$ is a UFD. So we can factorise $f(X)$ in $\mathbb{F}[X]$,

$$f(X) = p_1(X) \cdots p_k(X),$$

where $p_i(X)$ is irreducible in $\mathbb{F}[X]$.

There exists $a_i \in R$ with $a_i p_i(X) \in R[X]$. We have

$$a_i p_i(X) = c_i q_i(X)$$

with $c_i = c(a_i p_i(X))$, $q_i(X)$ primitive in $R[X]$ (and irreducible in $R[X]$ by Gauss' lemma). So

$$a_1 \cdots a_k f(X) = c_1 \cdots c_k q_1(X) \cdots q_k(X).$$

Considering contents and using the assumption that $f(X)$ is primitive, we have

$$a_1 \cdots a_k = u c_1 \cdots c_k$$

for some unit $u \in R$ since $q_i(X)$ is primitive by Lemma 2.23, i.e. $q_1(X) \cdots q_k(X)$ is primitive.

$$a_1 \cdots a_k f(X) = (u c_1 \cdots c_k)(u^{-1} q_1(X)) q_2(X) \cdots q_k(X).$$

Cancellation gives

$$f(X) = (u^{-1} q_1(X)) q_2(X) \cdots q_k(X),$$

a product of irreducibles in $R[X]$. This shows the existence part. For the uniqueness of the factorisation assume that

$$\begin{aligned} f(X) &= q_1(X) \cdots q_k(X) \\ &= r_1(X) \cdots r_l(X) \end{aligned}$$

with $q_i(X), r_i(X)$ irreducible in $R[X]$. We are assuming that $f(X)$ is primitive and so each $r_i(X)$ must be primitive.

By Gauss' lemma, $r_i(X)$ is irreducible in $\mathbb{F}[X]$. Uniqueness of factorisation in $\mathbb{F}[X]$ implies that $k = l$, and after reordering we have $q_i(X) = u_i r_i(X)$ with u_i a unit in $\mathbb{F}[X]$, hence in $\mathbb{F}$.

We can write $u_i = a_i / b_i$ for some $a_i, b_i \in R$ with $b_i \neq 0$. Then $b_i q_i(X) = a_i r_i(X)$. But $q_i(X)$ and $r_i(X)$ are primitive and hence b_i and a_i must be associates being the content of $b_i q_i(X) = a_i r_i(X)$. Cancelling gives that $q_i(X)$ and $r_i(X)$ are associates in $R[X]$. Thus factorisation is essentially unique in $R[X]$. □

Corollary 2.26. Let R be a UFD. Then $R[X_1, \ldots, X_n]$ is a UFD.

Proof. We assume that R is a UFD and repeated application of Lemma 2.23 gives that $R[X_1]$, $R[X_1, X_2] = (R[X_1])[X_2], \ldots, R[X_1, \ldots, X_n]$ is a UFD. □

Proposition 2.27 (Eisenstein's criterion)**.** Let R be a UFD and let $f(X) = a_0 + a_1 X + \cdots + a_n X^n \in R[X]$, $a_n \neq 0$, be primitive. Assume that for some irreducible p we have $p \nmid a_n$, $p \mid a_i$ for $0 \leq i < n$ and $p^2 \nmid a_0$. Then $f(X)$ is irreducible in $R[X]$, and hence in $\mathbb{F}[X]$ by Gauss' lemma.

Proof. Suppose that $f(X) = g(X)h(X)$ with

$$g(X) = r_k X^k + \cdots + r_0,$$
$$h(X) = s_l X^l + \cdots + s_0$$

where $r_k, s_l \neq 0$. Note that $k + l = n$. Since $p \mid a_0 = r_0 s_0$ and $p^2 \nmid a_0$ we may assume that $p \mid r_0$ and $p \nmid s_0$. $p \nmid a_n = r_k s_l$ implies $p \nmid r_k$, $p \nmid s_l$.

Set j to be such that $p \mid r_0, \ldots, p \mid r_{j-1}, p \nmid r_j$. Consider the term

$$a_j = r_0 s_j + \cdots + r_j s_0.$$

Note $p \nmid a_j$. Thus $j = n$, and hence $k = n$, $l = 0$. As $f(X)$ is primitive, the constant polynomial $h(X)$ must be a unit. Thus $f(X)$ cannot be factorised as a product of non-units. □

Example. Consider $R = \mathbb{Z}$ and let $f(X) = X^n - p \in \mathbb{Z}[X]$, where p is prime.

By Eisenstein's criterion, $f(X)$ is irreducible in $\mathbb{Z}[X]$ and hence in $\mathbb{Q}[X]$. So $f(X)$ does not have any roots in $\mathbb{Q}$, as a root would yield a linear factor of $f(X)$ in $\mathbb{Q}[X]$. Thus p has no nth roots in $\mathbb{Q}$.

Example. Consider the cyclotomic polynomial

$$f(X) = X^{p-1} + X^{p-2} + \cdots + 1$$

where p is prime. The claim is that this is irreducible in $\mathbb{Z}[X]$, and hence in $\mathbb{Q}[X]$.

Note that $(X - 1)f(X) = X^p - 1$. We make the substitution $X = 1 + Y$ and get

$$Y f(1 + Y) = (1 + Y)^p - 1,$$
$$f(1 + Y) = Y^{p-1} + \binom{p}{1} Y^{p-2} + \cdots + \binom{p}{p-1}$$

where $\binom{p}{i}$ are binomial coefficients. Eisenstein's criterion does now apply, we have $p \mid \binom{p}{i}$, $p^2 \nmid \binom{p}{p-1} = p$. We deduce that $f(1 + Y)$ is irreducible and hence $f(X)$ is irreducible in $\mathbb{Z}[X]$, and hence in $\mathbb{Q}[X]$.

2.6 Gaussian integers

We consider the set of *Gaussian integers* as introduced before,

$$\mathbb{Z}[i] = \{a + ib : a, b \in \mathbb{Z}\}.$$

Lecture 16

The norm

$$N(a+ib) = a^2 + b^2 = (a+ib)(a-ib) = z\bar{z},$$

where $z = a + ib$, is multiplicative, that is,

$$N(z_1 z_2) = N(z_1)N(z_2).$$

The units have to be of norm 1, hence the only units are $\pm 1, \pm i$.

Recall from an earlier example that $\mathbb{Z}[i]$ is a Euclidean domain and hence a PID and a UFD. Thus the irreducibles and primes are the same, by Lemma 2.16 and Lemma 2.20. What are they?

- $2 = (1+i)(1-i)$.
- 3, $N(3) = 9$. If it were to factorise as a product of two non-units then they would have norm 3. But there are no elements of norm 3.
- $5 = (1+2i)(1-2i)$.

Lemma 2.28. A prime number p in $\mathbb{Z}$ is irreducible (and prime) in $\mathbb{Z}[i]$ if and only if p is not of the form $x^2 + y^2$ with $x, y \in \mathbb{Z} - \{0\}$.

Proof. If $p = x^2 + y^2$ then $p = (x+iy)(x-iy)$ and so p is not irreducible in $\mathbb{Z}[i]$. Otherwise consider $N(p) = p^2$, p factorises into two non-units, necessarily of norm p, only if there is $x + iy$ with $N(x+iy) = p$. So $x^2 + y^2 = p$. □

Proposition 2.29. The irreducibles (and primes) in $\mathbb{Z}[i]$ are up to associates

(i) $p \in \mathbb{Z}$ prime with $p \equiv 3 \pmod 4$,

(ii) z with $z\bar{z} = p$, for p prime in $\mathbb{Z}$ with $p = 2$ or $p \equiv 1 \pmod 4$.

Proof. Let us first show that these are actually irreducibles (and primes).

(i) If $p \equiv 3 \pmod 4$ then it is not of the form $x^2 + y^2$ since squares modulo 4 are 0 or 1 and so $x^2 + y^2 \equiv 0, 1$ or $2 \pmod 4$. By Lemma 2.28, p is irreducible (and prime) in $\mathbb{Z}[i]$.

(ii) For example, $p = 2 = (1+i)(1-i)$ is of the correct form, $1 \pm i$ are irreducibles of norm 2.

Consider the finite field of p elements $\{0, 1, \ldots, p-1\} = \mathbb{F}_p \cong \mathbb{Z}/p\mathbb{Z}$. Its multiplicative group $\mathbb{F}_p \setminus \{0\}$ is *cyclic*. If not then the structure theorem for abelian groups implies we can find a subgroup of form $C_m \times C_m$ for some $m > 1$.

Thus there are at least m^2 elements satisfying $a^m = 1$. These are all roots of $X^m - 1$ in $\mathbb{F}_p[X]$ which is a UFD. $X^m - 1$ has a unique expression up to ordering and associates as a product of at most m irreducibles in $\mathbb{F}_p[X]$.

In particular, there are at most m linear factors $X - a$ and so at most m roots, contradicting that there are at least m^2 roots.

In this case, $\mathbb{F}_p$ has a cyclic multiplicative group. A cyclic group $\langle g \rangle$ of order $4n$ has a subgroup of order 4, namely $< g^n >$, and a unique element of order 2, namely g^{2n}. $p-1$ is the element of order 2 in $\mathbb{F}_p - \{0\}$, the multiplicative group.

An element of order 4 corresponds to $a \in \mathbb{Z}$ with $a^2 \equiv -1 \pmod{p}$. Thus $p \mid a^2+1 = (a+i)(a-i)$. But $p \nmid a \pm i$ and so p is *not* an irreducible (and prime) in $\mathbb{Z}[i]$.

Thus p must factorise, $p = z_1 z_2$ with z_i non-units in $\mathbb{Z}[i]$. We have $N(z_i) = p$. Write $z_i = x \pm iy$. We get $p = x^2 + y^2$.

Now assume α is irreducible in $\mathbb{Z}[i]$, so $\bar{\alpha}$ is irreducible. Take $p \mid N(\alpha)$. We use that $N(\alpha) = \alpha\bar{\alpha}$, a product of irreducibles.

If $p \equiv 3 \pmod{4}$ then p prime and unique factorisation implies that p is an associate of either α or $\bar{\alpha}$. So p is an associate of α.

If $p = 2$, or $p \equiv 1 \pmod{4}$ then $p = z\bar{z} \mid \alpha\bar{\alpha}$ and unique factorisation, so z is an associate of α or $\bar{\alpha}$ and so α is an associate of z or $\bar{z}$.

Thus our list is complete. □

Corollary 2.30. Let $n = p_1^{n_1} \cdots p_k^{n_k} \in \mathbb{Z}$ be the prime factorisation of n, with $p_1, \ldots, p_k$ distinct primes in $\mathbb{Z}$. Then n is of the form $x^2 + y^2$ if and only if whenever $p_i \equiv 3 \pmod{4}$ then n_i is even.

Proof. Suppose $n = x^2 + y^2 = (x+iy)(x-iy) = z\bar{z}$ with $z = x + iy$. Thus $N(z) = n$. Express z as a product of irreducibles in $\mathbb{Z}[i]$,

$$z = \alpha_1 \cdots \alpha_s,$$

say. But we know what the irreducibles are from Proposition 2.29. Either $N(\alpha_j) = p_j^2$ with $p_j \equiv 3 \pmod{4}$ or $N(\alpha_j) = p_j$ with $p_j = 2$ or $p_j \equiv 1 \pmod{4}$. Thus $n = N(z) = \prod N(\alpha_j)$ is of the required form.

Conversely, if $n = p_1^{n_1} \cdots p_k^{n_k}$ with n_j even for $p_j \equiv 3 \pmod{4}$ then we can replace any $p_j = 2$ or $p_j \equiv 1 \pmod{4}$ by $p_j = \alpha_j\bar{\alpha}_j$ and any other primes $p_j = \alpha_j = \bar{\alpha}_j$ and so $p_j^2 = \alpha_j\bar{\alpha}_j$ for some irreducible α_j. Thus n is of the form $z\bar{z}$ for some $z = x + iy$ and so $n = x^2 + y^2$. □

Example. Let $n = 65 = 5 \times 13$ and note that $5 = (2+i)(2-i)$, $13 = (2+3i)(2-3i)$. The unique factorisation up to reordering and associates is

$$n = (2+i)(2-i)(2+3i)(2-3i).$$

We use this to express n as $x^2 + y^2 = z\bar{z}$ with $z = x + iy$

$$n = [(2+i)(2+3i)][(2-i)(2-3i)] = (1+8i)(1-8i) = 1^2 + 8^2$$

or

$$n = [(2+i)(2-3i)][(2-i)(2+3i)] = (7-4i)(7+4i) = 4^2 + 7^2.$$

Lecture 17

2.7 $\mathbb{Z}[\alpha]$ with α an algebraic integer

Definition. A complex number $\alpha \in \mathbb{C}$ is an *algebraic integer* if there is a monic polynomial $f(X) \in \mathbb{Z}[X]$ with $f(\alpha) = 0$.

For example,

$\alpha \in \mathbb{C}$	Minimal polynomial of α
i	$f(X) = X^2 + 1$
$\sqrt{2}$	$f(X) = X^2 - 2$
$\frac{1}{2}(1 + \sqrt{-3})$	$f(X) = X^2 - X + 1$

$\mathbb{Z}[\alpha]$ is the smallest subring of $\mathbb{C}$ containing $\mathbb{Z}$ and α. $\mathbb{Z}[\alpha] \cong \mathbb{Z}[X]/I$ where I is the kernel of the map $\theta\colon \mathbb{Z}[X] \to \mathbb{Z}[\alpha], g(X) \mapsto g(\alpha)$.

Proposition 2.31. For an algebraic integer α this kernel is a principal ideal generated by a monic irreducible polynomial in $\mathbb{Z}[X]$.

Definition. This generator is called the *minimal polynomial* of α, written $f_\alpha(X)$.

Proof. By definition, $f(\alpha) = 0$ for some monic $f(X) \in \mathbb{Z}[X]$. So we can pick $f_\alpha(X) \in I$ of minimal degree ≥ 0, and we may assume $f_\alpha(X)$ is primitive in $\mathbb{Z}[X]$. The claim is that this is the required polynomial.

We wish to show that $I = (f_\alpha(X))$.

Let $h(X) \in I$. Then, as $\mathbb{Q}[X]$ is a Euclidean domain, we can write

$$h(X) = q(X)f_\alpha(X) + r(X)$$

with $r(X) = 0$ or $\deg r(X) < \deg f_\alpha(X)$. Clearing denominators, there exists $a \in \mathbb{Z}\setminus\{0\}$ such that

$$ah(X) = aq(X)f_\alpha(X) + ar(X)$$

with $aq(X) \in \mathbb{Z}[X]$ and $ar(X) \in \mathbb{Z}[X]$. But $ar(\alpha) = 0$ so $ar(X) \in I$. Minimality of degree of $f_\alpha(X)$ implies that $ar(X) = 0$ so $r(X) = 0$. So $ah(X) = aq(X)f_\alpha(X)$. Consider the contents of both sides,

$$a \mid c(ah(X)),$$
$$c(aq(X)f_\alpha(X)) = c(aq(X))$$

since $f_\alpha(X)$ is primitive. So a divides all the coefficients of $aq(X)$. So $q(X) \in \mathbb{Z}[X]$. Thus $h(X) \in (f_\alpha(X)) \lhd \mathbb{Z}[X]$.

We claim that $f_\alpha(X)$ is prime (and irreducible).

If $f_\alpha(X) = f_1(X)f_2(X)$ then $0 = f_1(\alpha)f_2(\alpha)$ and so $f_i(X) \in I$ for some i. $f_\alpha(X) \mid f_i(X)$ for some i.

Finally, we show that $f_\alpha(X)$ is monic.

Since $f_\alpha(X) \mid f(X)$ where $f(X)$ as in the first line. □

Definition (Rational integers)**.** The elements of $\mathbb{Z}$ are called *rational integers*.

Lemma 2.32. If α is an algebraic integer and $\alpha \in \mathbb{Q}$ then $\alpha \in \mathbb{Z}$.

Proof. $f_\alpha(X)$ is irreducible in $\mathbb{Z}[X]$ and monic. By Gauss' lemma, Lemma 2.22, $f_\alpha(X)$ is irreducible in $\mathbb{Q}[X]$. α is a root. So $f_\alpha(X) = X - \alpha$. Thus $\alpha \in \mathbb{Z}$. □

Example. Let $p \in \mathbb{Z}$ be prime and consider $\mathbb{Z}[X]/(p, f_\alpha(X))$. We have $\mathbb{F}_p \cong \mathbb{Z}/p\mathbb{Z}$. By the isomorphism theorems,

$$\mathbb{F}_p[X]/(\bar{f}_\alpha(X)) \cong \mathbb{Z}[X]/(p, f_\alpha(X)) \cong \mathbb{Z}[\alpha]/(p)$$

where $\bar{f}_\alpha(X)$ is the polynomial in $\mathbb{F}_p[X]$ obtained from $f_\alpha(X)$ by taking coefficients modulo p.

For example, $\alpha = i$, $f_\alpha(X) = X^2 + 1$. Then

$$\mathbb{F}_p/(X^2 + 1) \cong \mathbb{Z}[i]/(p).$$

For $p = 2$ or $p \equiv 1 \pmod 4$ this is not an integral domain, for $p \equiv 3 \pmod 4$ it is an integral domain.

Remark. There is quite a lot on algebraic integers in the Part II course *Number Fields*. For example, quadratic fields which are of the form $\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} : a, b \in \mathbb{Q}\} \leq \mathbb{C}$. In $\mathbb{Q}(\sqrt{d})$ the algebraic integers form a ring R. However, $(1 + \sqrt{-3})/2$ is an algebraic integer so R is not necessarily $\mathbb{Z}(\sqrt{d})$.

R is Euclidean if and only if d is -11, -7, -3, -2, -1, 2, 3, 5, 6, 7, 11, 13, 17, 19, 21, 29, 33, 37, 41, 57 or 73 (21 possibilities).

R is a UFD for $d < 0$ if and only if d is -1, -2, -3, -7, -11, -19, -43, -67, -163 (9 possibilities, cf. H.M. Stark *An introduction to number theory*). For $d > 0$ there exist 38 possibilities with $2 \leq d < 100$ so that R is a UFD. This question is still open.

2.8 Hilbert's basis theorem

Recall that we showed that a PID satisfies the ascending chain condition (ACC) on ideals.

Lemma 2.33. A ring R satisfies the ACC if and only if all ideals in R are finitely generated.

Proof. Suppose that all ideals are finitely generated and consider a chain

$$I_1 \subset I_2 \subset \cdots$$

with $I_j \triangleleft R$. Then the union $\bigcup_j I_j \triangleleft R$ is an ideal too. So by supposition $\bigcup_j I_j$ is finitely generated. There exists N such that all these finitely many generators lie in I_N.

If $m \geq N$ then $I_m \subset \bigcup_j I_j \subset I_N$. Also $I_N \subset I_m$ and so $I_m = I_N$.

Conversely, assume the ACC and let $J \triangleleft R$. Take $a_1 \in J$ with $a_1 \neq 0$. If $J \neq (a_1)$ pick $a_2 \in J \setminus (a_1)$. If $J \neq (a_1, a_2)$ pick $a_3 \in J \setminus (a_1, a_2)$ etc. We are producing a chain

$$(a_1) \subsetneq (a_1, a_2) \subsetneq (a_1, a_3, a_3) \subsetneq \cdots .$$

By the ACC this process stops, so $J = (a_1, \ldots, a_N)$ for some N. □

Definition. A ring with these properties is called *Noetherian*.

Theorem 2.34 (Hilbert's basis theorem)**.** If R is Noetherian then $R[X]$ is Noetherian.

Proof. Let $J \lhd R[X]$ be an ideal. We aim to show that it is finitely generated.

Consider

$$I_j = \left\{ a_j \in R : \exists f(X) \in J \; f(X) = \sum_{i=0}^{j} a_i X^i \in J \right\} \cup \{0\},$$

the set of leading coefficients of polynomials of degree j in J.

$I_j \lhd R$ is an ideal since if $\sum_{i=0}^{j} a_i X^i \in J$, $\sum_{i=0}^{j} b_i X^i \in J$ then $\sum_{i=0}^{j} (a_i + b_i) X^i \in J$ and if $a \in R$ then $\sum_{i=0}^{j} a a_i X^i \in J$.

$I_j \subset I_{j+1}$ since if $\sum_{i=0}^{j} a_i X^i \in J$ then $X\left(\sum_{i=0}^{j} a_i X^i\right) \in J$.

The ACC for R implies that there exists N with $I_m = I_N$ for all $m \geq N$ and I_N is finitely generated by the leading coefficients of $f_1(X), \ldots, f_k(X)$, say.

Now take any $f(X) \in J$ of degree $m \geq N$. The leading coefficient of $f(X)$ lies in $I_m = I_N$ and so there exists $r_1, \ldots, r_k \in R$ so that $r_1 f_1(X) + \cdots + r_k f_k(X)$ has the same leading coefficient. So $f(X) - (r_1 f_1(X) + \cdots + r_k f_k(X)) X^{m-N} \in J$ and is of degree less than m. Repeating this process yields $q_1(X), \ldots, q_k(X) \in R[X]$ such that $f(X) - (q_1(X) f_1(X) + \cdots + q_k(X) f_k(X)) \in J$ of degree less than N.

Now consider polynomials in J of degree less than N. For $j < N$ there is a finite set S_j of polynomials in J of degree j whose leading coefficients generate I_j. Let $S = \bigcup_{j<N} S_j$, a finite set.

A similar argument to the one before shows that any polynomial in J of degree less than N is of the form $\sum g_i(X) h_i(X)$ for $g_i(X) \in S$, $h_i(X) \in R[X]$.

Thus J is generated by $S \cup \{f_1(X), \ldots, f_k(X)\}$. □

Corollary 2.35. If $\mathbb{F}$ is a field then $\mathbb{F}[X_1, \ldots, X_n]$ is Noetherian and $\mathbb{Z}[X_1, \ldots, X_n]$ is also Noetherian.

Corollary 2.36. Any ring image of $\mathbb{Z}[X_1, \ldots, X_n]$ is Noetherian.

Proof. Let $\theta \colon \mathbb{Z}[X_1, \ldots, X_n] \to S$ be a homomorphism. Then θ is surjective. If $I \lhd S$ is an ideal then

$$J = \{ f(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n] : \theta(f(X_1, \ldots, X_n)) \in I \}$$

is finitely generated. Then I is generated by the images under θ of a finite generating set of J. □

Chapter 3

Modules

3.1 Introduction

Definition. Let R be a commutative ring. A set M is an *R-module* with operation $+$, with $(M, +)$ forming an abelian group, and also has a map

$$R \times M \to M, (r, m) \mapsto m$$

satisfying

$$\begin{aligned}(r_1 + r_2)m &= r_1 m + r_2 m \\ r(m_1 + m_2) &= rm_1 + rm_2 \\ r_1(r_2 m) &= (r_1 r_2)m \\ 1m &= m\end{aligned}$$

for all $r_1, r_2, r \in R$ and $m_1, m_2, m \in M$.

Example. (i) Let $R = \mathbb{F}$ be a field. Then all vector spaces over $\mathbb{F}$ are $\mathbb{F}$-modules.

(ii) For any R, R^n forms an R-module,

$$r(r_1, \ldots, r_n) = (rr_1, \ldots, rr_n).$$

In particular, when $n = 1$, R itself is an R-module.

(iii) If $I \lhd R$ is an ideal then I is an R-module and R/I is an R-module.

(iv) Let $R = \mathbb{Z}$. $\mathbb{Z}$-modules are precisely the abelian groups. Given an abelian group A with operation written as $+$ then the module map is given by

$$\begin{aligned}na &= \underbrace{a + \cdots + a}_{n \text{ times}} && \text{for } n \geq 1 \\ 0a &= 0 \\ (-n)a &= -(an) && \text{for } n \geq 1\end{aligned}$$

(v) $R = \mathbb{F}[X]$, $\mathbb{F}$ a field. If V is an $\mathbb{F}$-vector space and $\alpha \colon V \to V$ a linear map (vector space endomorphism) then V may be regarded as a $\mathbb{F}[X]$-module via

$$f(X).v = f(\alpha)(v)$$

for $v \in V$. Different maps α yield different $\mathbb{F}[X]$-modules.

(vi) If $R \leq S$ are rings then S can be regarded as an R-module via

$$rs = (rs)$$

for $r \in R, s \in S$. For example, let $\mathbb{Z} \leq \mathbb{Z}[\alpha]$, α an algebraic integer. We can regard $\mathbb{Z}[\alpha]$ as a $\mathbb{Z}$-module.

(vii) Let R be the prime subring of a finite field $\mathbb{F}$. Here $R \cong \mathbb{F}_p$ for some prime p. (Question 8 on Example Sheet 3 shows that $\mathbb{F}$ has cardinality p^n for some n.) We can view $\mathbb{F}$ as an $\mathbb{F}_p$-vector space.

Remark. There exists exactly one field of cardinality p^n for every prime p and $n \geq 1$, up to isomorphism. For example,

Cardinality	Field
4	$\mathbb{F}_2[X]/(X^2 + X + 1)$,
8	$\mathbb{F}_2[X]/(X^3 + X + 1)$

noting that e.g. $X^2 + X + 1$ is irreducible modulo 2. See also Question 9 on Example Sheet 2.

Definition. A subset N of an R-module M is an *R-submodule*, written $N \leq M$, if it is an additive subgroup of M and $rn \in N$ for all $r \in R, n \in N$.

Example. Ideals $I \lhd R$ are the R-submodules of the R-module R. In a $\mathbb{F}$-vector space the vector subspaces are the $\mathbb{F}$-submodules.

Definition. If $N \leq M$ then the *quotient module* M/N has elements $m + N$ with the operation $+$ defined as for quotients of the additive groups,

$$r(m + N) = rm + N$$

for $r \in R, m \in M$. Check that this turns M/N into an R-module.

Definition. A map $\theta\colon M \to N$ is an R-module homomorphism if

$$\theta(m_1 + m_2) = \theta(m_1) + \theta(m_2),$$
$$\theta(rm) = r\theta(m).$$

Example. If $R = \mathbb{F}$, $\mathbb{F}$ a field, then an R-module homomorphism is a linear map between $\mathbb{F}$-vector spaces.

Theorem 3.1 (First isomorphism theorem)**.** If $\theta\colon M \to N$ is an R-module homorphism then $\ker\theta$ is an R-submodule of M, the image of θ is a submodule of N and with

$$\ker\theta = \{a \in M : \theta(m) = 0\},$$
$$\operatorname{Im}\theta = \{\theta(m) : m \in M\}$$

we have that $M/\ker\theta \cong \operatorname{Im}\theta$.

Proof. Left to the reader. The isomorphism theorem for additive subgroups shows that $M/\ker\theta \cong \operatorname{Im}\theta$ for the additive groups. So we only need to check the multiplicative properties. □

Lecture 19

As usual there is a $1-1$ correspondence,

$$\left\{\begin{matrix}\text{submodules of} \\ M/N\end{matrix}\right\} \longleftrightarrow \left\{\begin{matrix}\text{submodules of } M \\ \text{containing } N\end{matrix}\right\}.$$

If $N \leq L \leq M$ are R-modules then

$$M/L \cong (M/N)/(L/N).$$

Compare this with the third isomorphism theorem.

Example. In the special case of $R = \mathbb{F}$ a field and vector spaces, we have quotient spaces V/W where $W \leq V$ and there are isomorphism theorems for linear maps ($\mathbb{F}$-module homomorphisms).

Definition. Let $m \in M$ then the *annihilator* of m is

$$\operatorname{Ann}(m) = \{r \in R : rm = 0\}.$$

The annihilator of M is

$$\begin{aligned}\operatorname{Ann}(M) &= \{r \in R : \forall m \in M \quad rm = 0\} \\ &= \bigcap_{m \in M} \operatorname{Ann}(m).\end{aligned}$$

These annihilators are ideals in R since

$$\begin{aligned}r_1 m = 0, r_2 m = 0 &\implies 0 = r_1 m + r_2 m = (r_1 + r_2)m, \\ r_1 m = 0 &\implies (r r_1)m = 0.\end{aligned}$$

Lemma 3.2. If M is an R-module and $m \in M$ then

$$Rm \cong R/\operatorname{Ann}(m)$$

where $Rm = \{rm : r \in R\}$.

Proof. Apply Theorem 3.1 to the R-module homomorphism $\theta\colon R \to M, r \mapsto rm$ with $\ker\theta = \operatorname{Ann}(m)$ and $\operatorname{Im}\theta = Rm$. □

Modules of the form Rm are *cyclic modules*. More generally, if $m_1, \ldots, m_k \in M$ and

$$M = Rm_1 + \cdots + Rm_k = \{r_1 m_1 + \cdots + r_k m_k : r_1, \ldots, r_k \in R\}$$

then M is generated by $m_1, \ldots, m_k$, and M is finitely generated.

Example. If $R = \mathbb{Z}$ then $\mathbb{Z}[\alpha]$ is a $\mathbb{Z}$-module where α is an algebraic integer. In fact, it is a finitely generated $\mathbb{Z}$-module. (Exercise.)

Lemma 3.3. Let $N \leq M$ be R-modules. If M is finitely generated then M/N is finitely generated.

Proof. Suppose $M = Rm_1 + \cdots + Rm_k$ then M/N is generated by $m_1 + N, \ldots, m_k + N$ since

$$\begin{aligned} m + N &= r_1 m_1 + \cdots + r_k m_k + N \\ &= r_1(m_1 + N) + \cdots + r_k(m_k + N). \end{aligned}$$

for some $r_1, \ldots, r_k \in R$. □

Warning. N need not be finitely generated.

Example. The polynomial ring $\mathbb{C}[X_1, X_2, \ldots]$ in countably infinitely many variables $X_1, X_2, \ldots$. Consider the ideal I of polynomials with zero constant term. This is not finitely generated since $I = \bigcup I_j$ where $I_j = (X_1, \ldots, X_j)$ and hence $I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \cdots$. Thus $\mathbb{C}[X_1, X_2, \ldots]$ is not Noetherian.

If $R = \mathbb{C}[X_1, X_2, \ldots]$ then the submodules are the ideals, and we have shown that there is a submodule which is not finitely generated inside the cyclic module R.

3.2 Direct sums, free modules

Definition. If $M_1, \ldots, M_k$ are R-modules then the *direct sum* $M_1 \oplus \cdots \oplus M_k$ has elements $(m_1, \ldots, m_k)$ with $m_i \in M_i$, addition

$$(m_1, \ldots, m_k) + (m_1', \ldots, m_k') = (m_1 + m_1', \ldots, m_k + m_k')$$

and scalar multiplication

$$r(m_1, \ldots, m_k) = (rm_1, \ldots, rm_k).$$

Definition. Let $m_1, \ldots, m_k \in M$. Then the set $\{m_1, \ldots, m_k\}$ is *independent* if

$$r_1 m_1 + \cdots + r_k m_k = 0 \implies r_1 = \cdots = r_k = 0.$$

Definition. The subset $S \subset M$ *generates M freely* if

(i) S generates M,

(ii) every map $\psi\colon S \to N$, where N is an R-module, extends to an R-module homomorphism $\theta\colon M \to N$.

Remark. If such a θ exists then it is unique since if we have two such θ_1 and θ_2 then $S \subset \ker(\theta_1 - \theta_2)$ and so $M \leq \ker(\theta_1 - \theta_2)$ since S generates M. Thus $\theta_1 = \theta_2$.

Definition. A module freely generated by some subset S is *free* and S is a *basis*.

Proposition 3.4. For $S = \{m_1, \ldots, m_k\} \subset M$ the following are equivalent:

(i) S generates M freely.

(ii) S generates M and is an independent set.

(iii) Every element of M is *uniquely* expressible in the form $r_1 m_1 + \cdots + r_k m_k$ for some $r_i \in R$.

Lecture 20

Proof. [(i) $\implies$ (ii)] Suppose that S generates M freely. We need to show independence. Let $N = R^k$ and define

$$\psi\colon S \to N, m_j \mapsto (0, \dots, 1, \dots),$$

i.e. a 1 in the jth place. So there is an R-module homomorphism $\theta\colon M \to N$. Then

$$\theta(r_1 m_1 + \cdots + r_k m_k) = (r_1, \dots, r_k).$$

So if $r_1 m_1 + \cdots + r_k m_k = 0$ then $r_1 = \cdots = r_k = 0$.

The two implications (ii) $\implies$ (iii) and (iii) $\implies$ (i) are left as an exercise. □

Example. For $R = \mathbb{Z}$, $M = \mathbb{Z}$, the set $\{2, 3\}$ is a generating set, but contains no basis for M. For example, $\{2\}$ is independent, but is not a generating set.

Proposition 3.5. Let M be a R-module, generated by $\{m_1, \dots, m_k\}$. Then there is a free module R^k and a surjective homomorphism $\theta\colon R^k \to M$. Thus $M \cong R^k / \ker\theta$.

Proof. Define an R-module homomorphism

$$\theta\colon R^k \to M, (r_1, \dots, r_k) \mapsto r_1 m_1 + \cdots + r_k m_k.$$ □

The kernel in Proposition 3.5 is the *relation module*.

Definition. An R-module M is *finitely presented* if there exists a finite generating set $S = \{m_1, \dots, m_k\} \subset M$ and the relation module is also finitely generated, by $\{n_1, \dots, n_l\}$. We say that M is *generated by S subject to relations* $r_1 m_1 + \cdots + r_k m_k = 0$ for each $n_i = (r_{i1}, \dots, r_{ik})$.

Proposition 3.6. Let R be a non-zero ring, M be an R-module freely generated by $\{u_1, \dots, u_m\}$ and $\{v_1, \dots, v_n\}$. Then $m = n$.

Proof. One can define determinants for square matrices with coefficients just as in *Linear Algebra* by

$$\begin{aligned} \det A &= \sum_{\pi \in S_n} \operatorname{sgn}(\pi) A_{1,\pi(1)} \cdots A_{n,\pi(n)}, \\ \det AB &= \det A \det B, \\ A \operatorname{adj}(A) &= (\det A) I. \end{aligned}$$

Thus A is invertible if and only if $\det A$ is a unit in R. Assume that $n \geq m$ with unique expressions

$$v_j = \sum_i A_{ij} u_i, \qquad u_k = \sum_j B_{jk} v_j.$$

So

$$\begin{aligned} u_k &= \sum_j \sum_i B_{jk} A_{ij} u_i \\ &= \sum_i \sum_j A_{ij} B_{jk} u_i \end{aligned}$$

$$= \sum_i (AB)_{ik} u_i.$$

As $\{u_1, \ldots, u_m\}$ is independent we have $AB = I$. If $n > m$ then

$$\begin{pmatrix} A & 0 \end{pmatrix} \begin{pmatrix} B \\ 0 \end{pmatrix} = I_m$$

so $\det \begin{pmatrix} A & 0 \end{pmatrix} \neq 0$. Contradiction. So $m = n$. □

3.3 Matrices over Euclidean domains, equivalence of matrices, Smith normal form

In this section we will assume that R is a Euclidean domain with Euclidean function $\phi\colon R - \{0\} \to \mathbb{Z}_{\geq 0}$.

If $a, b \in R$ then there exists a $\mathrm{hcf}(a, b)$, defined up to associates, obtained by Euclid's algorithm (cf. the Part IA course *Numbers and Sets*, Example Sheet Question 2). Moreover, $\mathrm{hcf}(a, b) = ax + by$ for some $x, y \in R$.

Definition. The elementary row operations on a $m \times n$ matrix A are:

(ER1) Add c times ith row to jth row. (Achieved by multiplying A on the left by $I + C$ where $C_{kl} = 0$ except for $C_{ij} = c$.)

(ER2) Interchange rows i and j, $i \neq j$. (Multiplying by $I + C$ where $C_{kl} = 0$, $C_{ii} = C_{jj} = -1$, $C_{ij} = C_{ji} = 1$.)

(ER3) Multiplying row i by a unit c. (Multiplying by C diagonal, all diagonal entries 1 apart from $C_{ii} = c$.)

All of these are achieved by multiplication on the left by suitable matrices. These operations are reversible as the corresponding matrices are invertible.

We could similarly consider elementary column operations (EC1), (EC2), (EC3), which are achieved by multiplying on the right by suitable invertible matrices.

Definition. Two $m \times n$ matrices are *equivalent* if one can get from one to the other via a sequence of elementary operations.

If A, B are equivalent then there exists an invertible $m \times m$ matrix Q and an invertible $n \times n$ matrix P with $B = QAP^{-1}$. We can see this by doing the bookkeeping on the accumulation of row and column operations.

Theorem 3.7 (Smith normal form)**.** Let A be an $m \times n$ matrix over a Euclidean domain R. Then by a sequence of elementary operations, we can put it into the form

$$\begin{pmatrix} d_1 & & & & & 0 \\ & \ddots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix}$$

Lecture 21

with $d_1, \ldots, d_r$ non-zero and $d_1 \mid d_2 \mid \cdots \mid d_r$. These d_k, the *invariant factors*, are unique up to associates.

In fact, the product $d_1 \cdots d_k$ is a highest common factor of the $k \times k$ minors of A.

Recall the following definition. A $k \times k$ minor of A is the determinant of a $k \times k$ submatrix of A. In particular, d_1 is a highest common factor of the entries of A, $d_1 d_2$ is a highest common factor of the 2×2 minors.

Lemma 3.8. The ideal in R generated by $k \times k$ minors of A is unchanged under elementary operations.

Proof. Consider the 1×1 minors, i.e. the entries of A. Under an elementary operation, A_{ij} either stays the same or is replaced by $A_{ij} + cA_{ik}$ or $A_{ij} + cA_{kj}$ or multiplied by a unit or replaced by some other entry A_{kl}, and so we deduce that the ideal generated by the entries of the new matrix is contained in the ideal generated by the entries of the old matrix. Now the other direction follows from the fact that the operations are revertible. We thus obtain equality.

More generally, have a similar but messier argument. □

Proof of Theorem 3.7. If $A = 0$ there is nothing to do. Suppose $A \neq 0$ and we may assume after switching rows and columns that $A_{11} \neq 0$. The idea is to use sequences of elementary operations to reduce $\phi(A_{11})$. The process must stop since $\phi(A_{11}) \in \mathbb{Z}_{\geq 0}$.

Case 1. If A_{11} does not divide some entry A_{1j} of the first row then use Euclid's algorithm, write $A_{1j} = qA_{11} + r$ with $r \neq 0$, so $\phi(r) < \phi(A_{11})$. Subtract q times the first column from the jth column to give entry r in the $(1\ j)$ position. Switch columns 1 and j so that r is now in the top left hand corner.

Case 2. If A_{11} does not divide some entry of the first column, use a similar process to get a new matrix with $r \neq 0$ and $\phi(r) < \phi(A_{11})$ in the top left hand corner.

Keep using these two cases until we cannot do so any longer, i.e. when A_k divides all entries in first row and column. Then subtract suitable multiples of the first column from the others and substract suitable multiples of the first row from the others to get matrix of the form

$$\begin{pmatrix} d & 0 \\ 0 & C \end{pmatrix}$$

with $d \neq 0$ for some $(m-1) \times (n-1)$ matrix C.

Case 3. If we have a matrix of the form

$$\begin{pmatrix} d & 0 \\ 0 & C \end{pmatrix}$$

but d does not divide some entry of C then $d \nmid A_{ij}$, say. Use Euclid's algorithm so that $A_{ij} = qd + r$ with $r \neq 0$ and $\phi(r) < \phi(d)$. Add column 1 to column j. Subtract q times row 1 from row i so that we have replaced A_{ij} by r. Switch columns j and 1, rows i and 1, so r appears in top left hand corner.

We are now in a position to apply Case 1 and Case 2 again to reduce the ϕ-value of the top left hand corner. Repeat the whole process (Cases 1, 2 and 3) until we can't anymore, i.e. when the matrix is of the form

$$\begin{pmatrix} d & 0 \\ 0 & C \end{pmatrix}$$

with d dividing all entries of C.

Keep on going with elementary operations on C etc. The result follows. □

Observe that for a matrix of the form

$$\begin{pmatrix} d_1 & & & & & 0 \\ & \ddots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix}$$

with $d_1 \mid d_2 \mid \cdots \mid d_r$ and $d_k \neq 0$ for all k the ideal generated by the $k \times k$ minors is $(d_1 \cdots d_k)$, thus by Lemma 3.8 the ideal generated by the $k \times k$ minors of the original matrix is $(d_1 \cdots d_k)$.

Example. We consider the following sequence of transformations,

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 \\ 3 & 5 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}.$$

(i) Add column 2 to column 1. (Multiply on the right by $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$.)

(ii) Add column 1 to column 2. (Multiply on the right by $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.)

(iii) Subtract 3× row 1 from row 2. (Multiply on the left by $\left(\begin{smallmatrix} 1 & 0 \\ -3 & 1 \end{smallmatrix}\right)$.)

Thus

$$\begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} A \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Lemma 3.9. Let M be a free R-module of rank m, i.e. there is a basis of cardinality m, and R be an Euclidean domain and suppose that $N \leq M$ is a submodule. Then N is finitely generated.

Proof. Pick a basis for M and so $M \cong R^m$. Identify M with R^m. Consider the ideal

$$I = \{r \in R : (r_1, r_2, \ldots, r_m) \in N\} \lhd R.$$

Since R is an ED and hence a PID, I is generated by $a \in R$, say. Fix an element $n \in N$ of form $(a, a_2, \ldots, a_m)$. Then for any $(r_1, r_2, \ldots, r_m) \in N$ we have $r_1 = ra$ for some $r \in R$. Then

$$\begin{aligned}(r_1, r_2, \ldots, r_m) - rn &= (r_1, r_2, \ldots, r_m) - r(a, a_2, \ldots, a_m) \in N \\ &= (0, r_2 - ra_2, \ldots, r_m - ra_m).\end{aligned}$$

Consider $\{(0, s_2, \ldots, s_m) \in N\} \leq \{(0, r_2, \ldots, r_m) \in R^m\}$ and apply induction to see that the module on the LHS is of rank $m - 1$, generated by $n_2, \ldots, n_m$, say. Then $n, n_2, \ldots, n_m$ generate N. □

Lecture 22

Theorem 3.10. Let R be a Euclidean domain, $N \leq R^m$. Then there is a basis $\{v_1, \ldots, v_m\}$ of R^m such that N is generated by $\{d_1 v_1, d_2 v_2, \ldots, d_r v_r\}$ for some $1 \leq r \leq m$ and $d_1 \mid d_2 \mid \cdots \mid d_r$.

Proof. Use Lemma 3.9 to see that N is finitely generated by $x_1, \ldots, x_n$, say. So write these as columns of a matrix A, an $m \times n$ matrix.

By Theorem 3.7 on Smith normal forms, we know that we can put A into the form

$$\begin{pmatrix} d_1 & & & & & 0 \\ & \ddots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix}, \quad d_i \neq 0, \quad d_1 \mid d_2 \mid \cdots \mid d_r$$

by elementary row and columns operations.

Observe that performing an elementary row operation is associated with a change of basis of R^m, e.g. adding c times row j to row i replaces the basis $\mathbf{e}_1, \ldots, \mathbf{e}_m$ of R^m by $\mathbf{e}_1, \ldots, \mathbf{e}_i, \ldots, \mathbf{e}_j - c\mathbf{e}_i, \ldots, \mathbf{e}_m$ (replacing the jth element $\mathbf{e}_j$ by $\mathbf{e}_j - c\mathbf{e}_i$) since

$$a_1\mathbf{e}_1 + \ldots + a_m\mathbf{e}_m = a_1\mathbf{e}_1 + \cdots + (a_i + ca_j)\mathbf{e}_i \cdots + a_j(\mathbf{e}_j - c\mathbf{e}_i) + \cdots + a_m\mathbf{e}_m$$

Thus the ith coefficient has been replaced by $a_i + ca_j$.

Similary, column operations arise in connection with the change of the generating set of N. Thus the elementary operations arise when changing basis for R^m and generating set for N.

At the end of the process the matrix is in Smith normal form and the basis $\{v_1, \ldots, v_m\}$ of R^m is such that $\{d_1 v_1, d_2 v_2, \ldots, d_r v_r, 0, \ldots, 0\}$ is a generating set for N, as required. We can throw away the 0s. □

Corollary 3.11. A submodule N of R^m when R is a ED is free of rank at most m.

Proof. The set $\{d_1 v_1, \ldots, d_r v_r\}$ *freely* generates N since $\{v_1, \ldots, v_m\}$ are free generators of R^m. □

Theorem 3.12. Let M be a finitely generated R-module, where R is an ED. Then

$$M \cong \frac{R}{(d_1)} \oplus \frac{R}{(d_2)} \oplus \cdots \oplus \frac{R}{(d_r)} \oplus R \cdots \oplus R$$

for some $d_k \neq 0$ with $d_1 \mid d_2 \mid \cdots \mid d_r$.

Remark. (i) We may assume that all the d_k are non-units, for if d_k is a unit then $R/(d_k) \cong \{0\}$ which may be thrown away.

(ii) A *finitely* generated R-module is a direct sum of cyclic R-modules. (Assuming that R an ED.)

Proof. Suppose the module M is finitely generated by $\{m_1, \dots, m_m\}$. Then $M \cong R^m/N$ by Proposition 3.5. This is because there is a R-module homomorphism

$$\theta : R^m \to M, (r_1, \dots, r_m) \mapsto r_1 m_1 + \cdots + r_m m_m$$

and $\operatorname{Im}\theta = M$ since $\{m_1, \dots, m_m\}$ generates M and $N = \ker\theta \leq R^m$. By the isomorphism theorem, $M \cong R^m/N$.

There exists a basis $\{v_1, \dots, v_m\}$ of R^m such that $\{d_1 v_1, \dots, d_r v_r\}$ is a generating set for N. So

$$R^m/N \cong R/(d_1) \oplus \cdots \oplus R/(d_r) \oplus R \oplus \cdots R,$$

as required. □

Example. Take $R = \mathbb{Z}$. The theorem tells us about finitely generated $\mathbb{Z}$-modules, that is, finitely generated abelian groups. Consider an abelian group A, written additively, generated by a, b, c subject to relations

$$\begin{aligned} 2a + 3b + c &= 0, \\ a + 4b &= 0, \\ 5a + 6b + 7c &= 0. \end{aligned}$$

Then A is a $\mathbb{Z}$-module and $A \cong \mathbb{Z}^3/N$, where N is generated by $(2,3,1), (1,4,0), (5,6,7)$. Write these as columns of a matrix

$$\begin{pmatrix} 2 & 1 & 5 \\ 3 & 4 & 6 \\ 1 & 0 & 7 \end{pmatrix}$$

and put this into Smith normal form,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 21 \end{pmatrix}.$$

(Check.) So $A \cong \mathbb{Z}^3/N \cong \mathbb{Z}/(1\mathbb{Z}) \oplus \mathbb{Z}/(1\mathbb{Z}) \oplus \mathbb{Z}/(21\mathbb{Z}) \cong \mathbb{Z}/(21\mathbb{Z})$, a cyclic abelian group of order 21.

Theorem 3.13 (Structure theorem for finitely generated abelian groups)**.** A finitely generated abelian group is isomorphic to

$$C_{d_1} \times C_{d_2} \times \cdots \times C_{d_r} \times C_\infty \times \cdots \times C_\infty,$$

where C_∞ is the infinite cyclic group.

Proof. Set $R = \mathbb{Z}$ in Theorem 3.12 and write the group operation multiplicatively. □

Remark. For a finite group there are no copies of C_∞, see Theorem 1.19.

Proposition 3.14 (Primary decomposition)**.** Let R be a ED. Then

$$R/(d) \cong R/(p_1^{n_1}) \oplus \cdots \oplus R/(p_s^{n_s}),$$

where $d = p_1^{n_1} \cdots p_s^{n_s}$ is the factorisation into irreducibles (and primes).

Lecture 23

Proof (cf. (Lemma 1.20). Split off one summand $R/\left(p_j^{n_j}\right)$ at a time using the following lemma. □

Lemma 3.15. If $d = r_1 r_2$ with $\mathrm{hcf}(r_1, r_2) = 1$ then $M = R/(d) \cong R/(r_1) \oplus R/(r_2)$.

Proof. Let m be a generator of M with $\mathrm{Ann}(m) = (d)$. If $\mathrm{hcf}(r_1, r_2) = 1$ then we can write $1 = xr_1 + yr_2$ for some $x, y \in R$ by Euclid's algorithm. Then

$$m = 1.m = x(r_1 m) + y(r_2 m).$$

We have $\mathrm{Ann}(r_1 m) = (r_2)$, $\mathrm{Ann}(r_2 m) = (r_1)$ using that we have good factorisation in R. Set

$$M_1 \cong R/(r_2), \qquad M_2 \cong R/(r_1).$$

$M_1 \cap M_2 = \{0\}$ since if $sm \in M_1 \cap M_2$ then s is a multiple of both r_1 and r_2, and hence of $r_1 r_2$ since $\mathrm{hcf}(r_1, r_2) = 1$ and so $sm = 0$.

Consider the R-module homomorphism

$$M_1 \oplus M_2 \to M, (m_1, m_2) \mapsto m_1 + m_2.$$

It is onto since $M = M_1 + M_2$, $M_1 \cap M_2 = \{0\}$. Hence $M \cong M_1 \oplus M_2$. □

Theorem 3.16. Let M be a finitely generated R-module, where R is a Euclidean domain. Then

$$M \cong N_1 \oplus N_2 \oplus \cdots \oplus N_t$$

where each $N_j \cong R/(p_j^{n_j})$ and p_j is prime (and irreducible) and $n_j \geq 1$, or $N_j \cong R$.

Proof. Use primary decomposition from Lemma 3.14 to split the components in Proposition 3.12 as direct sums of modules of this form.

Note that the p_j are not necessarily distinct. The $p_j^{n_j}$ are the *elementary divisors* and they are unique up to ordering. (Proof omitted.) □

3.4 Modules over $\mathbb{F}[X]$ for a field $\mathbb{F}$ — normal forms for matrices

Let $\alpha\colon V \to V$ be a linear map and V a finite dimensional vector space over a field $\mathbb{F}$. We regard V as a $\mathbb{F}[X]$-module via

$$g(X).v = g(\alpha)(v)$$

for $v \in V$, dependent on the choice of α.

Example. Cyclic modules over $\mathbb{F}[X]$, e.g. $V = M \cong \mathbb{F}[X]/(f(X))$. Here $f(X)$ is a polynomial of least degree such that $f(\alpha) = 0$. We may assume that $f(X)$ is monic, it is the minimal polynomial for α.

(i) $f(X) = X^r$. Take generator m of the $\mathbb{F}[X]$-module M, $\mathrm{Ann}(m) = (f(X))$. Then $m, Xm, X^2m, \ldots, X^{r-1}m$ is a vector space basis of $M = V$. Note that

$$(m, Xm, \ldots, X^{r-1}m) = (m, \alpha(m), \ldots, \alpha^{r-1}(m)).$$

α is represented by the matrix (with coefficients in $\mathbb{F}$)

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

(ii) $f(X) = (X - \lambda)^r$. Then $(\alpha - \lambda)^r = 0$. Consider $\beta = \alpha - \lambda.\iota$ then the minimal polynomial of β is X^r. So there exists a vector space basis of $M = V$ such that β is represented by

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

and so α is represented by

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

(iii) For a general $f(X)$, for a generator m with $\mathrm{Ann}(m) = (f(X))$ as in (i), we can pick a vector space basis $m, Xm, \ldots, X^{r-1}m$ where

$$f(X) = a_0 + a_1X + \cdots + a_{r-1}X^{r-1} + X^r$$

and

$$m, Xm, \ldots, X^{r-1}m = m, \alpha(m), \ldots, \alpha^{r-1}(m).$$

α is represented with respect to this basis by

$$\begin{pmatrix} 0 & 0 & & & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & \ddots & & \vdots \\ & & \ddots & 0 & \\ 0 & & & 1 & -a_{r-1} \end{pmatrix}.$$

This is called the *companion matrix* of $f(X)$, written $C(f)$.

Lecture 23

Theorem 3.17 (Rational canonical form)**.** Let $\alpha : V \to V$ be an endomorphism of a finite dimensional $\mathbb{F}$-vector space and $\mathbb{F}$ be a field. Then, regarding V as an $\mathbb{F}[X]$-module M, $M \cong M_1 \oplus \cdots \oplus M_s$ with each M_j cyclic and $M_j \cong \mathbb{F}[X]/(f_j(X))$ where $f_1(X) \mid f_2(X) \mid \cdots \mid f_s(X)$, and on choosing an vector space basis for each M_j as in Example (iii), α is represented by a matrix (with coefficients in $\mathbb{F}$)

$$\begin{pmatrix} C(f_1) & & & 0 \\ & C(f_2) & & \\ & & \ddots & \\ 0 & & & C(f_s) \end{pmatrix}.$$

Proof. Theorem 3.12 splits M as a direct sum of cyclic modules of the right form, and since M is finite dimensional as a $\mathbb{F}$-vector space there are no components isomorphic to $\mathbb{F}[X]$. Now use Example (iii) to represent α. □

Remark. The name is due to the special case where $\mathbb{F} = \mathbb{Q}$.

Remark. (i) The *invariant factors* $f_i(X)$ are unique up to associates. (This is not quite proved here).

(ii) If A is a square $n \times n$ matrix with coefficients in $\mathbb{F}$, then A represents a linear map $\alpha\colon V \to V$. So the theorem says we can pick a new basis with respect to which α is represented in rational canonical form. Thus A is conjugate to a matrix in rational canonical form.

(iii) Minimal polynomials: The minimal polynomial of α is a generator of the annihilator $\mathrm{Ann}(M)$, and this is equal to $f_s(X)$ after adjusting to make sure it is monic.

(iv) The characteristic polynomial of α is the product of the characteristic polynomials of the $C(f_i)$, that is, the product $f_1(X) \cdots f_s(X)$.

Now we can use Proposition 3.14 on primary decomposition to split the M_i as direct sums of cyclic modules with annihilators generated by powers of irreducibles.

Assume that $\mathbb{F} = \mathbb{C}$, so that the irreducibles are linear.

Theorem 3.18 (Jordan normal form for $\mathbb{C}$)**.** Let $\alpha\colon V \to V$ be an endomorphism of a finite dimensional $\mathbb{C}$-vector space and regard V as a $\mathbb{C}[X]$-module M. Then

$$M \cong M_1 \oplus \cdots \oplus M_s$$

where $M_j \cong \mathbb{C}[X]/((X - \lambda_j)^{a_j})$ for some $\lambda_j \in \mathbb{C}$. Here, $\lambda_1, \ldots, \lambda_s$ are not necessarily distinct.

Taking a $\mathbb{C}$-vector space basis for each M_j as in Example (ii), α is represented by a

matrix of the form

$$\begin{pmatrix}
\lambda_1 & & & 0 & & & & 0 & 0 \\
1 & \ddots & & & & & & & \\
 & \ddots & \lambda_1 & & & & & & \\
0 & & 1 & \lambda_1 & 0 & & & & \\
 & & & 0 & \lambda_2 & & & & \\
 & & & & 1 & \ddots & & & \\
 & & & & & \ddots & \lambda_2 & & \\
 & & & & 0 & & 1 & \lambda_2 & \\
0 & & & & & & & 0 & \ddots
\end{pmatrix}.$$

Remark. (i) A submatrix of the form

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

is called a *Jordan λ-block*.

(ii) The Jordan blocks for α are unique up to reordering. (This is not proved here.)

(iii) Minimal polynomials of α: Observe that each M_i is in rational canonical form, so the theorem yields a λ-block for each λ with $X - \lambda$ dividing $f_i(X)$. Since

$$f_1(X) \mid f_2(X) \mid \cdots \mid f_s(X),$$

the powers of $X - \lambda$ increase as we pass from $f_1(X)$ to $f_s(X)$. Thus the size of λ-blocks increases as we pass from M_1 to M_s, so the largest λ-block arises from M_s.

Recall that the minimal polynomial of α is

$$f_s(X) = \prod_{\lambda \text{ distinct}} (X - \lambda)^{a_\lambda}.$$

Then a_λ is the size of the largest λ-block.

(iv) The characteristic polynomial of α factorises into irreducibles as follows,

$$f_1(X) \cdots f_s(X) = \prod_{\lambda \text{ distinct}} (X - \lambda)^{b_\lambda}$$

where b_λ is the sum of the sizes of λ-blocks. Observe that the λ are the eigenvalues of α.

(v) The *geometric multiplicity* of λ is defined to be the dimension of the λ-eigenspace and equal to the number of λ-blocks.

(vi) Given a square complex matrix A, it is conjugate to a matrix in Jordan normal form.

Lecture 24

Example (Solutions of linear difference equations and differential equations)**.** Consider the space V of complex sequences $(z_k) \in \mathbb{C}^\infty$ that are solutions of

$$z_{i+k} + c_{k-1}z_{i+(k-1)} + \cdots + c_0 z_i = 0$$

for $i \geq 1$ with $c_0, \ldots, c_{k-1} \in \mathbb{C}$.

Note that V is a finite dimensional $\mathbb{C}$-vector space. Let $\alpha\colon V \to V$ be the left-shift,

$$(z_1, z_2, \ldots) \mapsto (z_2, z_3, \ldots).$$

The minimal polynomial of α is $X^k + c_{k-1}X^{k-1} + \cdots + c_0 = f(X)$, the auxiliary polynomial. Factorise this is as

$$f(X) = \prod_{\lambda \text{ distinct}} (X - \lambda)^{a_\lambda}.$$

Write down the $\mathbb{C}$-vector space associated with the Jordan normal form, as sequences for each λ,

$$\left(\binom{k}{a}\lambda^{k-a}\right), \qquad \text{for } 0 \leq a \leq a_\lambda - 1$$

e.g. for $a = 0$ we have the sequence $(\lambda, \lambda^2, \lambda^3, \ldots)$ and for $a = 1$ etc.

For differential equations, the linear map is differentiation.