

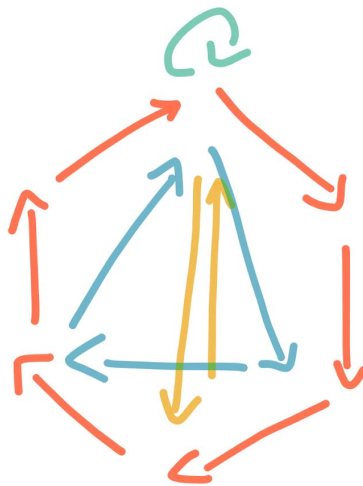
1st Year Mathematics
Imperial College London

Spring 2015

Algebra I

Lectured by:
Prof. J. BRITNELL

Humbly Typed by:
Karim BACCHUS



Caveat Lector: unofficial notes, *not* endorsed by Imperial College.
Comments and corrections should be sent to kb514@ic.ac.uk

Syllabus

Introductions to three topics in abstract algebra: The theory of vector spaces and linear transformations, the theory of groups and the theory of rings.

- Vector spaces: Linear maps, rank-nullity theorem, connections with linear equations and matrices.
- Groups: Axioms, examples. Cyclic groups, dihedral groups, symmetric groups. Lagranges theorem and applications.
- Rings: Polynomial rings, rings of the form $\mathbb{Z}[\sqrt{d}]$. Euclids algorithm for certain rings. Uniqueness of factorisation for these rings. Applications to Diophantine Equations.

Appropriate books

J. Fraleigh and R. Beaugregard, *Linear Algebra*
S. Lipschutz and M. Lipson, *Linear Algebra*
J. B. Fraleigh, *A First Course in Abstract Algebra*
R. Allenby, *Rings, Fields and Groups*
I. N. Herstein, *Topics in Algebra*

Contents

Linear Algebra	4
1 Vector Spaces	4
Spanning Sets and Bases (M1GLA Review)	4
More on Subspaces	4
Rank of a Matrix	8
Linear Transformations	12
Kernels and Images	15
Matrix of Linear Transformation	20
Eigenvalues and Eigenvectors	23
Diagonalisation of Linear Transformations	24
Change of Basis	25
Abstract Algebra	27
2 Group Theory	28
Groups	28
Subgroups	33
Cyclic Subgroups	35
The Order of a Group Element	36
Cycles	37
Order of a Permutation	40
Lagrange's Theorem	42
Modular Arithmetic	44
Dihedral Groups	48
3 Ring Theory	51
Arithmetic of Rings	54
The Rings $\mathbb{Z}[\sqrt{d}]$	56
Highest Common Factor / Greatest Common Divisor	58
Unique Factorisation	62

1 Vector Spaces

Spanning Sets and Bases (M1GLA Review)

Lecture 1

Let V be a vector space, and let $S = \{v_1, \dots, v_k\}$ be a finite subset of V . Then the *span* of S is the set $\{\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k, \lambda_1, \dots, \lambda_k \text{ scalar}\}$.

- * Span S is a subspace of V .
- * If Span $S = V$, then we say that S is a *spanning set* for V .
- * If V has a finite spanning set, then we say that V is finite dimensional.

Assume from now on that V is finite dimensional. The set S is *linearly independent* if the only solution to the equation $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = \mathbf{0}$ is $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$.

A *basis* for V is a linearly independent spanning set.

- * V has a basis
- * Every basis of V has the same size. This is the *dimension* of V , written $\dim V$.

Suppose that $\dim V = n$.

- * Any spanning set of size n is a basis.
- * Any linearly independent set of size n is a basis.
- * Every spanning set contains a basis as a subset.
- * Every linearly independent set is contained as a subset in a basis.
- * Any subset of V of size $< n$ is **not** a spanning set.
- * Any subset of V of size $> n$ is **not** linearly independent.
- * If W is a subspace of V then $\dim W \leq \dim V$ with equality only when $W = V$.

Every vector space has associated with it a set of *scalars*. E.g. \mathbb{R}^n has the scalar set \mathbb{R} . The scalars come as a structure called a *field* (To be defined in the Ring Theory Section). I'll write F for the field of scalars. It will usually be safe to assume that $F = \mathbb{R}$. Other fields include \mathbb{C} , \mathbb{Q} , integers and $(\mathbb{Z}/p\mathbb{Z})^\times$.

More on Subspaces

Definition 1. Let V be a vector space, and let U and W be subspaces of V . The *intersection* of U and W is $U \cap W = \{v : v \in U \text{ and } v \in W\}$. The *subspace sum* (or just *sum*) of U and W is $U + W = \{u + w : u \in U, w \in W\}$

Remark 2. Note that $\mathbf{0} \in U$ and $\mathbf{0} \in W$, so if $u \in U$, then $u = \mathbf{0} \in U + W$, and similarly if $w \in W$, then $w = \mathbf{0} \in U + W$. So $U \subseteq U + W$ and $W \subseteq U + W$. ($U + W$ usually contains many other vectors)

Example 3. Let $V \in \mathbb{R}^2$. Let $U = \text{Span}\{(1, 0)\}$, and $W = \text{Span}\{(0, 1)\}$. So $U = \{(\lambda, 0) : \lambda \in \mathbb{R}\}$, $W = \{(0, \lambda) : \lambda \in \mathbb{R}\}$.

We see that $U + W$ contains (λ, μ) for all $\lambda, \mu \in \mathbb{R}$ and so $U + W = V$.

Proposition 4. $U \cap W$ and $U + W$ are both subspaces of V .

PROOF. Do $U + W$ first. Checking the subspace axioms:

- (i) $\mathbf{0} \in U$ and $\mathbf{0} \in W \implies \mathbf{0} = \mathbf{0} + \mathbf{0} \in U + W$. So $U + W \neq \emptyset$.
- (ii) Suppose $v_1, v_2 \in U + W$. Then $v_1 = u_1 + w_1$ and $v_2 = u_2 + w_2$, where $u_1, u_2 \in U$ and $w_1, w_2 \in W$.

Now $v_1 + v_2 = u_1 + w_1 + u_2 + w_2 = (u_1 + u_2) + (w_1 + w_2) \in U + W \implies$ closed under addition.

- (iii) Let $v \in U + W$. Then $v = u + w$ where $u \in U$ and $w \in W$. Let $\lambda \in F$. Then $\lambda v = \lambda(u + w) = \lambda u + \lambda w \in U + W$.

Now do $U \cap W$:

- (i) $\mathbf{0} \in U$ and $\mathbf{0} \in W$, so $\mathbf{0} \in U \cap W$ by definition.
- (ii) Suppose $v_1, v_2 \in U \cap W$. Then $v_1, v_2 \in U$ and so $v_1 + v_2 \in U$ since U is closed. Similarly for $v_1 + v_2 \in W$. Hence $v_1 + v_2 \in U \cap W$.
- (iii) Suppose $v \in U \cap W$, and $\lambda \in F$. Then $v \in U$ and so $\lambda v \in U$. $v \in W \implies \lambda v \in W$. Hence $\lambda v \in U \cap W$. ■

Proposition 5. Let V be a vector space and let U & W be subspaces. Suppose that $U = \text{Span}\{u_1, \dots, u_r\}$ and $W = \text{Span}\{w_1, \dots, w_s\}$. Then $U + W = \text{Span}\{u_1, \dots, u_r, w_1, \dots, w_s\}$. Lecture 2

PROOF. By inclusion both ways:

Notice that $u_i \in U \subseteq U + W$ (by Remark 2) and similarly $w_i \in W \subseteq U + W$, $\forall i$. Since $U + W$ is a subspace of V , so $U + W$ is closed under linear combinations, so $\text{Span}\{u_1, \dots, u_r, w_1, \dots, w_s\} \subseteq U + W$.

For the reverse inclusion, let $v \in U + W$.

Then $v = u + w$ for some $u \in U, w \in W$. Since $U = \text{Span}\{u_1, \dots, u_r\}$, we have $u = \lambda_1 u_1 + \dots + \lambda_r u_r$, for some $\lambda_1, \dots, \lambda_r \in F$. Similarly $w = \mu_1 w_1 + \dots + \mu_s w_s$, for some $\mu_1, \dots, \mu_s \in F$.

Now $v = \lambda_1 u_1 + \dots + \lambda_r u_r + \mu_1 w_1 + \dots + \mu_s w_s \in \text{Span}\{u_1, \dots, u_r, w_1, \dots, w_s\}$. Hence $U + W \subseteq \text{Span}\{u_1, \dots, u_r, w_1, \dots, w_s\}$. ■

Examples 6.

Question: Let $V = \mathbb{R}^4$, $U = \text{Span}\{(1, 1, 2, -3), (1, 2, 0, -3)\}$ and $W = \text{Span}\{(1, 0, 5, -4), (-1 - 3, 0, 5)\}$. Find a basis for $U + W$.

Answer: By Proposition 5, we have:

$$U + W = \text{Span}\{(1, 1, 2, -3), (1, 2, 0, -3), (1, 0, 5, -4), (-1, -3, 0, 5)\}.$$

We then just row reduce the matrix:

$$\begin{pmatrix} 1 & 1 & 2 & -3 \\ 1 & 2 & 0 & -3 \\ 1 & 0 & 5 & -4 \\ -1 & -3 & 0 & 5 \end{pmatrix} \rightarrow \text{Echelon Stuff} \rightarrow \begin{pmatrix} 1 & 1 & 2 & -3 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The three non-zero rows are linearly independent, and have the same span as the original four vectors, so a basis for $U + W$ is:

$$\{(1, 1, 2, -3), (0, 1, -2, 0), (0, 0, 1, -1)\} \text{ (or just the first 3 vectors).}$$

Question: What about a basis for $U \cap W$?

Answer: If $v \in U \cap W$, then $v \in U = \text{Span}\{(1, 1, 2, -3), (1, 2, 0, -3)\}$. So $v = a(1, 1, 2, -3) + b(1, 2, 0, -3)$ for $a, b \in \mathbb{R}$.

And $v \in W = \text{Span}\{(1, 0, 5, -4), (-1, -3, 0, 5)\}$. So $v = c(1, 0, 5, -4) + d(-1, -3, 0, 5)$. So we have:

$$a(1, 1, 2, -3) + b(1, 2, 0, -3) - c(1, 0, 5, -4) - d(-1, 3, 0, 5) = \mathbf{0} \quad (*)$$

(*) gives us 4 simultaneous equations, which we can encode as a matrix equation:

$$\begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & 2 & 0 & -3 \\ 2 & 0 & -5 & 0 \\ -3 & -3 & 4 & -5 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \mathbf{0}$$

We find the solution space by row reducing:

$$\begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & 2 & 0 & -3 \\ 2 & 0 & -5 & 0 \\ -3 & -3 & 4 & -5 \end{pmatrix} \rightarrow \text{Echelon Stuff} \rightarrow \begin{pmatrix} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

There is one line of solutions, given by $a = 5d, b = -4d, c = 2d$. So pick $d = 1$, then $a = 5, b = -4, c = 2, d = 1$.

$$\text{So } v = a(1, 1, 2, -3) + b(1, 2, 0, -3) = (5, 5, 10, -15) - (4, 8, 0, -12) = (1, -3, 10, 3).$$

We can check our solutions with c, d : $v = c(1, 0, 5, -4) + d(-1, -3, 0, 5) = (1, -3, 10, 3)$ as expected. So $U \cap W$ is 1-dimensional, and a basis is $\{(1, -3, 10, 3)\}$.

Theorem 7

Let V be a vector space, and let U and W be subspaces. Then $\dim U + W = \dim U + \dim W - \dim U \cap W$.

PROOF. Let $\dim U = r$, $\dim W = s$, $\dim U \cap W = m$. Let $\{v_1, \dots, v_m\}$ be a basis for $U \cap W$.

Then $\{v_1, \dots, v_m\}$ is a linearly independent subset of U . So it is contained in some basis for U . So there exists u_1, \dots, u_{r-m} in U such that $\{v_1, \dots, v_m, u_1, \dots, u_{r-m}\}$ is a basis for U . Similarly, there exists w_1, \dots, w_{s-m} such that $\{v_1, \dots, v_m, w_1, \dots, w_{s-m}\}$.

Now let $B = \{v_1, \dots, v_m, u_1, \dots, u_{r-m}, w_1, \dots, w_{s-m}\}$. Then $\text{Span } B = U + W$ by Proposition 5.

Claim. B is linearly independent.

Proof of claim:

Suppose that $\alpha_1 v_1 + \dots + \alpha_m v_m + \beta_1 u_1 + \dots + \beta_{r-m} u_{r-m} + \gamma_1 w_1 + \dots + \gamma_{s-m} w_{s-m} = \mathbf{0}$. For $\alpha_i, \beta_i, \gamma_i \in F$. Then:

$$\begin{aligned} \sum_{i=1}^{s-m} \gamma_i w_i &= -\sum_{i=1}^m \alpha_i v_i - \sum_{i=1}^{r-m} \beta_i u_i \in U \\ \implies \sum_{i=1}^{s-m} \gamma_i w_i &\in U \cap W \implies \sum_{i=1}^{s-m} \gamma_i w_i = \sum_{i=1}^m \delta_i v_i, \text{ for some } \delta_i \in F \end{aligned}$$

$$\text{Hence: } \sum_{i=1}^m \delta_i v_i - \sum_{i=1}^m \alpha_i v_i = \mathbf{0}.$$

But $\{v_1, \dots, v_m, w_1, \dots, w_{s-m}\}$ is linearly independent (being a basis for W). So $\gamma_i = \delta_i = 0, \forall i$. Since $\gamma_i = 0 \forall i$, we have:

$$\sum_{i=1}^m \alpha_i v_i + \sum_{i=1}^{r-m} \beta_i u_i = \mathbf{0}.$$

But $\{v_1, \dots, v_m, u_1, \dots, u_{r-m}\}$ is linearly independent (being a basis for U). So $\alpha_i = \beta_i = 0, \forall i$. Hence B is linearly independent. So B is a basis for $U + W$. So $\dim U + W = |B| = m + (r - m) + (s - m) = r + s - m$. ■

Example 8. Question: Let $V = \mathbb{R}^3$, and let $U = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$. Similarly let $W = \{(x, y, z) \in \mathbb{R}^3 : -x + 2y + z = 0\}$. Find bases for $U, W, U \cap W, U + W$.

Lecture 3

Answer: A general element of U is $(x, y, -x - y) = x(1, 0, -1) + y(0, 1, -1) \implies$

$U = \text{Span}\{(1, 0, -1), (0, 1, -1)\}$. Since this set is clearly linearly independent, it's a basis for U .

A general element of W is $(x, y, x - 2y) = x(1, 0, 1) + y(0, 1, -2) \implies W = \text{Span}\{(1, 0, 1), (0, 1, -2)\}$. Again, clearly linearly independent, so a basis for W .

Suppose that $v = (x, y, z)$ lies in $U \cap W$. Then $v \in U$, and so $z = -x - y$. But also $v \in W$, and so $z = x - 2y$. Hence $-x - y = x - 2y \implies y = 2x$.

So a general element of $U \cap W$ is $(x, 2x, -3x) = x(1, 2, -3) \implies U \cap W = \{(1, 2, -3)\}$ is a basis for $U \cap W$.

As in the proof of theorem 7, we find a basis for U , and W , each of which contain a basis for $U \cap W$. So any linearly independent subset of U of size 2 is a basis for U , i.e. $U = \{(1, 2, -3), (1, 0, -1)\}$. Similarly $W = \{(1, 2, -3), (1, 0, 1)\}$. So a spanning set for $U + W = \{(1, 2, -3), (1, 0, -1), (1, 0, 1)\}$.

By Theorem 7, we know $\dim U + W = \dim U + \dim W - \dim U \cap W = 2 + 2 - 1 = 3$. Hence our spanning set is a basis for $U + W$.

Rank of a Matrix

Definition 9. Let A be an $m \times n$ matrix with entries from F . Define the *row-span* of A by $\text{RSp}(A) = \text{Span}\{\text{rows of } A\}$. This is a subspace of F^n . The *column-span*, $\text{CSp}(A)$ of A is $\text{Span}\{\text{columns of } A\}$, again a subspace of F^n .

The *row-rank* of A is $\dim \text{RSp}(A)$. The *column-rank* of A is $\dim \text{CSp}(A)$ i.e. the number of linearly independent rows / columns.

Example 10. $A = \begin{pmatrix} 3 & 1 & 2 \\ 0 & -1 & 1 \end{pmatrix}$, $F = \mathbb{R}$

$\text{RSp}(A) = \text{Span}\{(3, 1, 2), (0, -1, 1)\}$. Since the two vectors are linearly independent, we have that the row-rank $= \dim \text{RSp}(A) = 2$.

$$\text{CSp}(A) = \text{Span}\left\{\begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}.$$

This is linearly dependent, since $\begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

$$\text{So } \text{CSp}(A) = \text{Span}\left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\} \implies \text{CSp}(A) = 2.$$

How do we calculate the row-rank of a matrix A ?

Procedure 11.

Step 1: Use row operations to reduce the matrix A , to row echelon form.

$$A_{ech} = \begin{pmatrix} 1 & & & * \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{pmatrix}$$

Step 2: The row-rank of A is the number of non-zero rows in A_{ech} , and the non-zero rows of A_{ech} form a basis for $RSp(A)$. [i.e. you don't need to go back to original vectors get the basis.]

Justification We need to show:

- (i) $RSp(A) = RSp(A_{ech})$
- (ii) The non-zero rows of A_{ech} are linearly independent.

To show (1), recall that A_{ech} is obtained from A by a series of row-operations:

$$\begin{cases} r_i := r_i + \lambda r_j & (i \neq j) \\ r_i := \lambda r_i & (\lambda \neq 0) \\ r_i \longleftrightarrow r_j & (i \neq j) \end{cases}$$

Suppose that A' is obtained from A by one row-operation. Then it is clear than every row of A' lies in $RSp(A)$. So $RSp(A') \subseteq RSp(A)$. But every row operation is invertible by another row operation. i.e.

$$\begin{cases} r_i := r_i + \lambda r_j & \text{has inverse } r_i := r_i - \lambda r_j \\ r_i := \lambda r_i & \text{has inverse } r_i := r_i - 1/\lambda r_j \\ r_i \longleftrightarrow r_j & \text{has inverse } r_i \longleftrightarrow r_j \end{cases}$$

So we have $RSp(A) \subseteq RSp(A')$, and so the row spaces are equal. It follows that $RSp(A_{ech}) = RSp(A)$.

For (2), consider the form of A_{ech} , and denote r_1, \dots, r_k as the non-zero rows. Say r_i has it's leading entry in column c_i . Suppose that $\lambda_1 r_1 + \dots + \lambda_k r_k = \mathbf{0}$ (*).

Look at the co-ordinate corresponding to column c_1 . The only contribution is $\lambda_1 r_1$ since all of the other rows have 0 in that co-ordinate. Since r_1 has 1 in this co-ordinate $\implies \lambda_1 = 0$.

Since $\lambda_1 = 0$, the only contribution to c_2 is $\lambda_2 r_2$. So $\lambda_2 = 0$. We can continue this argument for $c_2, \dots, c_i, \dots, c_k \implies$ the only solution to (*) is $\lambda_1, \dots, \lambda_k = 0 \implies$ linearly independence of vectors.

Lecture 4

Example 12. Find the row-rank of $A = \begin{pmatrix} 1 & 2 & 5 \\ 2 & 1 & 0 \\ -1 & 4 & 15 \end{pmatrix}$

Answer: Reduce A to echelon form:

$$A = \begin{pmatrix} 1 & 2 & 5 \\ 2 & 1 & 0 \\ -1 & 4 & 15 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 5 \\ 0 & 1 & 10/3 \\ 0 & 0 & 0 \end{pmatrix} = A_{ech}$$

Since A_{ech} has two non-zero rows, the row-rank of A is 2. (Note: Scaling the second row to make the leading entry 1 was not necessary)

Example 13. Find the dimension of $W = \text{Span}\{(-1, 1, 0, 1), (2, 3, 1, 0), (0, 1, 2, 3)\} \subseteq \mathbb{R}^4$

Answer: Notice that $W = \text{RSp} \begin{pmatrix} -1 & 1 & 0 & 1 \\ 2 & 3 & 1 & 0 \\ 0 & 1 & 2 & 3 \end{pmatrix} = A$

$$\Rightarrow A_{ech} = \begin{pmatrix} -1 & 1 & 0 & 1 \\ 0 & 5 & 1 & 2 \\ 0 & 0 & 9 & 12 \end{pmatrix}$$

There are 3 non-zero rows in A_{ech} , so the row rank is 3.
Hence $\dim W = 3$.

Procedure 14. The columns of A are the rows of A^T . So apply Procedure 11 to the matrix A^T . (Alternatively, use column operations to reduce A to “column echelon form”, and then count the no. of non-zero columns).

Example 15. Let $A = \begin{pmatrix} 1 & 2 & 5 \\ 2 & 1 & 0 \\ -1 & 4 & 15 \end{pmatrix}$

The column rank of A is the row-rank of A^T .

$$A^T = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 4 \\ 5 & 0 & 15 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & -1 \\ 0 & -3 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

There are two non-zero rows, so $\text{row-rank}(A^T) = 2$, and so $\text{column-rank}(A) = 2$. A basis for $\text{RSp}(A^T)$ is $\{(1, 2, -1), (0, -3, 6)\}$, so a basis for $\text{CSp}(A)$ is the transpose of these vectors.

Theorem 16

For any matrix A , the row-rank of A is equal to the column-rank of A .

PROOF. Let the rows of A be r_1, \dots, r_m , so $r_i = (a_{i1}, a_{i2}, \dots, a_{in})$.

Let the columns of A be c_1, \dots, c_n , so $c_j = (a_{1j}, a_{2j}, \dots, a_{mj})^T$.

Let k be row-rank of A . Then $\text{RSp}(A)$ has basis $\{v_1, \dots, v_k\}$, $v_i \in F^n$. Every row r_i is a linear combination of v_1, \dots, v_k . Say that:

$$r_i = \lambda_{i1}v_1 + \lambda_{i2}v_2 + \dots + \lambda_{ik}v_k \quad (*)$$

Let $v_i = (b_{i1}, \dots, b_{in})$. Looking at the j th co-ordinate in $(*)$, we have $a_{ij} = \lambda_{i1}b_{1j} + \lambda_{i2}b_{2j} + \dots + \lambda_{ik}b_{kj}$

$$c_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{pmatrix} = \begin{pmatrix} \lambda_{11}b_{1j} + \lambda_{12}b_{2j} + \dots + \lambda_{1k}b_{kj} \\ \lambda_{21}b_{1j} + \lambda_{22}b_{2j} + \dots + \lambda_{2k}b_{kj} \\ \dots \\ \lambda_{m1}b_{1j} + \lambda_{m2}b_{2j} + \dots + \lambda_{mk}b_{kj} \end{pmatrix}$$

So every column of A is a linear combination of $(\lambda_{1i}, \lambda_{2i}, \dots, \lambda_{mi})^T$ for $1 \leq i \leq k$. So the column space of A is spanned by k vectors and so $\text{column-rank}(A) \leq k = \text{row-rank}(A)$.

But $\text{row-rank}(A) = \text{column-rank}(A^T)$, $\text{column-rank}(A) = \text{row-rank}(A^T)$.

By the argument above, $\text{column-rank}(A^T) \leq \text{row-rank}(A)$. So $\text{row-rank}(A) \leq \text{column-rank}(A)$. Hence $\text{row-rank}(A) = \text{column-rank}(A)$. ■

Definition 17. Let A be matrix. The *rank* of A , written $rk(A)$, is the row-rank of A . (which is also the column-rank of A .)

Example 18. Let $A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ -1 & 1 & 0 & 1 \\ 0 & 3 & -1 & 1 \end{pmatrix}$

Notice that $r_3 = r_1 + r_2$. So a basis for $\text{RSp}(A)$ is $\{(1, 2, -1, 0), (-1, 1, 0, 1)\}$

Write the rows of A as linear combinations of $\{v_1, v_2\}$:

$$r_1 = 1v_1 + 0v_2 \quad r_2 = 0v_1 + 1v_2 \quad r_3 = 1v_1 + 1v_2$$

(The scalars here are the λ_{ij} from the proof of Theorem 16)

According to the proof, a spanning set of for $\text{CSp}(A)$ is given by:

$$\{(1, 0, 1)^T, (0, 1, 1)^T\}$$

We verify this - We have $c_1 = (1, -1, 0)^T = w_1 - w_2$. $c_2 = (2, 1, 3)^T = 2w_1 + w_2$, $c_3 = (-1, 0, -1)^T = -w_1$, and $c_4 = (0, 1, 1)^T = w_2$.

Lecture 5 Proposition 19. Let A be an $n \times n$ (square) matrix. Then the following statements are equivalent:

- (i) $\text{rk}(A) = n$ (A has “full rank”)
- (ii) The rows of A form a basis for F^n
- (iii) The columns of A form a basis of F^n
- (iv) A is invertible (so $\det A \neq 0$, row reduced to I etc.)

PROOF. (1) \iff (2):

$$\begin{aligned}
 \text{rk}(A) = n &\iff \dim \text{RSp}(A) = n \\
 &\iff \text{RSp}(A) = F^n \\
 &\iff \text{Rows of } A \text{ are spanning a set for } F^n \text{ of size } n \\
 &\iff \text{The rows of } A \text{ form a basis for } F^n
 \end{aligned}$$

(1) \iff (3): The same, using columns instead.

(1) \iff (4):

$$\begin{aligned}
 \text{rk}(A) = n &\iff A_{ech} = I \\
 &\iff A \text{ can be row-reduced to } I \\
 &\iff A \text{ is invertible.}
 \end{aligned}$$

■

Linear Transformations

Suppose that V and W are vector spaces over a field F . Let $T : V \rightarrow W$ be a function.

- * Say that T “preserves addition”, if whenever $T : V \mapsto W$ and $T : v_1 \mapsto w_1$ and $T : v_2 \mapsto w_2$, we also have $T : v_1 + v_2 \mapsto w_1 + w_2$. (Briefly: $T(v_1 + v_2) = Tv_1 + Tv_2$.)
- * Say that T “preserves scalar multiplication”, if whenever $T : v \mapsto w$ and $\lambda \in F \implies T : \lambda v \mapsto \lambda w$. (Briefly: $T(\lambda v) = \lambda T(v)$)

Definition 20. The function $T : V \rightarrow W$ is a *linear transformation* (or *linear map*), if it preserves addition and scalar multiplication. So:

$$T(v_1 + v_2) = Tv_1 + Tv_2 \text{ and } T(\lambda v) = \lambda(T(v)), \forall v_1, v_2, v \in V \text{ \& } \lambda \in F$$

Examples 21.

- (a) $T : \mathbb{R}^2 \rightarrow \mathbb{R}, T(x, y) = x + y$. I claim this is a linear transformation. Check it preserves addition:

$$\begin{aligned}
 T((x_1, y_1) + (x_2, y_2)) &= T((x_1 + x_2, y_1 + y_2)) \\
 &= x_1 + x_2 + y_1 + y_2 \\
 &= x_1 + y_1 + x_2 + y_2 \\
 &= T((x_1, y_1)) + T((x_2, y_2))
 \end{aligned}$$

And T also preserves scalar multiplication, since if $l \in \mathbb{R}$, then:

$$T(l(x, y)) = T((lx, ly)) = lx + ly = l(x + y) = lT((x, y))$$

- (b) $T : \mathbb{R}^2 \rightarrow \mathbb{R}, T(x, y) = x + y + 1$. This is not linear.
For example, $2T((1, 0)) = 4$, but $T((2, 0)) = 3$, so it doesn't preserve scalar multiplication \implies not a linear map.
- (c) $T : \mathbb{R} \rightarrow \mathbb{R} T(x) = \sin(x)$, is not linear.
 $2T(\pi/2) = 2$, but $T(2 \times \pi/2) = T(\pi) = 0$. Again it doesn't preserve scalar multiplication, so not a linear map.
- (d) Let V be the space of all polynomials in a single variable x with co-efficients from \mathbb{R} . Define $T : v \rightarrow V$ by $T(f(x)) = \frac{d}{dx}f(x)$. Then T is a linear transformation.
 - (I) $\frac{d}{dx}(f(x) + g(x)) = \frac{d}{dx}f(x) + \frac{d}{dx}g(x) \implies$ preserves addition
 - (II) $\frac{d}{dx}(lf(x)) = l\frac{d}{dx}f(x) \implies$ preserves scalar multiplication

Proposition 22. Let A be an $m \times n$ matrix over F . Define $T : F^n \rightarrow F^m$ by $T(b) = Av$. This is a linear transformation. (We say that T is a matrix transformation.)

PROOF.

1. $T(v_1 + v_2) = A(v_1 + v_2)$
 $= Av_1 + Av_2$
 $= Tv_1 + Tv_2 \quad \forall v_1, v_2 \in F^n \implies T$ preserves addition.
2. $T(lv) = A(lv) = lAv$
 $= lTv \quad \forall v \in F^n, l \in F \implies T$ preserves scalar multiplication. ■

Examples 23.

- (a) Define a map $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ by:

$$T \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 - 2a_2 + a_3 \\ a_1 + a_2 - 2a_3 \end{pmatrix}$$

Then T is linear because:

$$T \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

and so T is a matrix transformation. So Proposition 22 applies.

- (b) Define $\rho_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ to be a rotation through an angle of θ (anticlockwise). Then ρ_θ is linear since it is given by the matrix:

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

Proposition 24. (Basic properties of linear transformations)

Let $T : V \rightarrow W$ be a linear transformation.

- (i) If $\mathbf{0}_v$ is the zero vector in V and $\mathbf{0}_w$ is the zero vector in w , then $T(\mathbf{0}_v) = \mathbf{0}_w$
- (ii) Suppose that $v_1, \dots, v_k \in V$ and that $v = l_1 v_1 + \dots + l_k v_k$ ($l_i \in F$). Then $Tv = l_1 T v_1 + l_2 T v_2 + \dots + l_k T v_k$.

PROOF. (i). Since T preserves scalar multiplication, for any $v \in V$, we have $T(0v) = 0Tv$, so $T(\mathbf{0}_v) = \mathbf{0}_w$. ■

Lecture 6 PROOF. (ii) We observe:

$$\begin{aligned} T(v) &= T(\lambda_1 v_1 + \dots + \lambda_k v_k) \\ &= T(\lambda_1 v_1 + \dots + \lambda_{k-1} v_{k-1}) + T(\lambda_k v_k) \\ &= T(\lambda_1 v_1 + \dots + \lambda_{k-1} v_{k-1}) + \lambda_k v_k \end{aligned}$$

Now a straightforward argument by induction tells us that:

$$T(v) = \lambda_1 T(v_1) + \lambda_2 T(v_2) + \dots + \lambda_k T(v_k) \quad \blacksquare$$

Example 25. Question: Find a linear transformation, $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, which sends $(1, 0) \rightarrow (1, -1, 2)$ and sends $(0, 1) \rightarrow (0, 1, 3)$.

Answer: Notice that $\{(1, 0), (0, 1)\}$ is a basis for \mathbb{R}^2 . So a general element of \mathbb{R}^2 is $a(1, 0) + b(0, 1)$, for $a, b \in \mathbb{R}$.

So if T is a solution to the question, then we must have $T(a(1, 0) + b(0, 1)) = a(1, -1, 2) + b(0, 1, 3)$, by Proposition 24(ii).

So we are forced to take $T(a, b) = (a, b - a, 2a + 3b)$. This is indeed a linear transformation, since it is a matrix transformation. And we do have $T(0, 1) = (1, -1, 2)$ and $T(0, 1) = (0, 1, 3)$ as required. So this is a solution and it is the unique solution.

Proposition 26. Let V and W be vector spaces over F . Let $\{v_1, \dots, v_n\}$ be a basis for V . Let $\{w_1, \dots, w_n\}$ be any n vectors in W . Then there is a unique linear transformation $T : V \rightarrow W$ such that $T(v_i) = w_i \forall i$.

Remark. The vectors w_1, \dots, w_n don't have to be linearly independent, or even distinct.

PROOF. Suppose that $v \in V$. Then there exists unique scalars $\lambda_1, \dots, \lambda_n$ such that $v = \lambda_1 v_1 + \dots + \lambda_n v_n$.

Define $T : V \rightarrow W$ by $T(v) = l_1 w_1 + \dots + l_n w_n$. (This makes sense, since the scalars l_i are uniquely determined by v .)

Show that T is linear:

Take $u, v \in V$. Write $u = \mu_1 v_1 + \dots + \mu_n v_n$, $v = l_1 v_1 + \dots + l_n v_n$. Then $u + v = (\mu_1 + l_1)v_1 + \dots + (\mu_n + l_n)v_n$. Now by the definition of T , we have:

$$T(u) = \mu_1 w_1 + \dots + \mu_n w_n \text{ and } T(v) = l_1 w_1 + \dots + l_n w_n$$

$$\text{Also } T(u + v) = (\mu_1 + l_1)w_1 + \dots + (\mu_n + l_n)w_n$$

So $T(u + v) = T(u) + T(v)$, so T preserves addition.

Now let $\pi \in F$. We have $\pi = \pi l_1 v_1 + \dots + \pi l_n v_n$ ■

Remark 27. Once we know what a linear transformation does on a basis, we know all about it. This gives a convenient shorter way of defining a linear transformation.

Example 28. Let V be the vector space of polynomials in a variable x over \mathbb{R} of degree ≤ 2 . A basis for V is $\{1, x, x^2\}$.

Pick three “vectors” in V ; $w_1 = 1 + x$, $w_2 = x - x^2$, $w_3 = 1 + x^2$. By Proposition 26, there should be a unique linear transformation, $T : V \rightarrow V$ such that $T(1) = w_1, T(x) = w_2, T(x^2) = w_3$.

Let’s work out what T does to an arbitrary polynomial, $c + bx + ax^2$ from V . We must have $c + bx + ax^2 \mapsto c(1 + x) + b(x - x^2) + a(1 + x^2) = (a + c) + (b + c)x + (a - b)x^2$.

Kernels and Images

Definition 29. Let $T : V \rightarrow W$ be a linear transformation. The *image* of T is the set $\{Tv : v \in V\} \subseteq W$. The *kernel* of T is the set of $\{v \in V : Tv = \mathbf{0}\}$. We write $\text{Im}(T)$ for the image, and $\text{Ker}(T)$ for the kernel.

Lecture 7

Example 30. Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be defined by:

$$T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

So:

$$T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3x_1 + x_2 + 2x_3 \\ -x_1 + x_3 \end{pmatrix}$$

So $\text{Im}(T)$ is the set:

$$\begin{aligned} & \left\{ \begin{pmatrix} 3x_1 + x_2 + 2x_3 \\ -x_1 + x_3 \end{pmatrix} : x_1, x_2, x_3 \in \mathbb{R} \right\} \\ &= \left\{ x_1 \begin{pmatrix} 3 \\ -1 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ 1 \end{pmatrix} : x_1, x_2, x_3 \in \mathbb{R} \right\} \\ &= \text{CSp}(A) \end{aligned}$$

The kernel of T is the set:

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : 3x_1 + x_2 + 2x_3 = 0, \quad -x_1 + x_3 = 0 \right\}$$

This is $\{v \in \mathbb{R}^3 : Av = \mathbf{0}\}$, the solution space of $Av = \mathbf{0}$. (Solved this in M1GLA).
(In this case the kernel is the span $\{(1, -5, 1)^T\}$)

Proposition 31. *Let $T : V \rightarrow W$ be a linear transformation. Then:*

- (i) $\text{Im}(T)$ is a subspace of W
- (ii) $\text{Ker}(T)$ is a subspace of V

(In general we write $U \leq V$ to mean that U is a subspace of V . So Proposition 31 says $\text{Im}(T) \leq W$, $\text{Ker}(T) \leq V$).

PROOF. (i) $\text{Im}(T) \leq W$:

Certainly $T(\mathbf{0}) \in \text{Im}(T)$, so $\text{Im}(T) \neq \emptyset$.

Suppose that $w_1, w_2 \in \text{Im}(T)$. Then there exists $v_1, v_2 \in V$ such that $Tv_1 = w_1$ and $Tv_2 = w_2$. Now $T(v_1 + v_2) = Tv_1 + Tv_2 = w_1 + w_2$. (Since T preserves addition). So $w_1 + w_2 \in \text{Im } T$, so $\text{Im}(T)$ is closed under addition.

Now suppose that $w \in \text{Im } T$ and $\lambda \in F$. Then there exists $v \in V$ such that $Tv = w$. Now $T(\lambda v) = \lambda T(v) = \lambda w$. (Since T preserves scalar multiplication). So $\lambda w \in \text{Im}(T)$, so $\text{Im}(T)$ is closed under scalar multiplication.

(ii) $\text{Ker}(T) \leq V$:

We know that $T(\mathbf{0}) = \mathbf{0}w$. So $\mathbf{0}v \in \text{Ker } T \implies \text{Ker } T \neq \emptyset$.

Suppose that $v_1, v_2 \in \text{Ker}(T)$. So: $Tv_1 = Tv_2 = \mathbf{0}$. Now $T(v_1 + v_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}$ (Since T preserves addition). So $v_1 + v_2 \in \text{Ker}(T)$, and so $\text{Ker}(T)$ is closed under addition.

Now suppose we have $v \in \text{Ker}(T)$, $\lambda \in F$. Then $Tv = \mathbf{0}$. Now $T(\lambda v) = \lambda Tv = \lambda \mathbf{0} = \mathbf{0}$. So $\lambda v \in \text{Ker}(T)$. So $\text{Ker}(T)$ is closed under scalar multiplication. ■

Example 32. Let V_n be the vector space of polynomials in a variable x over \mathbb{R} of degree $\leq n$.

We have $V_0 \leq V_1 \leq V_2 \leq \dots$

Define $T : V_n \rightarrow V_{n-1}$ by $T(f(x)) = \frac{d}{dx}f(x)$.

We have $\text{Ker}(T) = \{\text{constant polynomials}\} = V_0$.

If $g(x) \in V_{n-1}$, Let $f(x)$ be the antiderivative (integral) of $g(x)$. Since $\deg g(x) \leq n-1$, we have $\deg f(x) \leq n$. And $T(f(x)) = \frac{d}{dx}f(x) = g(x)$.

So $g(x) \in \text{Im}(T)$, and so $\text{Im}(T) = V_{n-1}$.

Proposition 33. Let $T : V \rightarrow W$ be a linear transformation. Let $v_1, v_2 \in V$. Then $Tv_1 = Tv_2 \iff T(v_1 - v_2) = \mathbf{0}$. $v_1 - v_2 \in \text{Ker}(T)$.

PROOF. $Tv_1 = Tv_2 \iff Tv_1 - Tv_2 = \mathbf{0} \iff T(v_1 - v_2) = \mathbf{0}$
(Since T preserves addition and multiplication by -1). ■

Proposition 34. Let $T : V \rightarrow W$ be a linear transformation. Suppose that $\{v_1, \dots, v_n\}$ is a basis for V . Then $\text{Im}(T) = \text{Span}\{Tv_1, \dots, Tv_n\}$.

PROOF. Let $w \in \text{Im}(T)$. Then there exists $v \in V$ such that $T(v) = w$. We can write v as a linear combination of basis vectors:

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n.$$

Now $Tv_1 = \lambda_1 Tv_1 + \dots + \lambda_n Tv_n$ by Proposition 24(ii). So $w = \lambda_1 Tv_1 + \dots + \lambda_n Tv_n \in \text{Span}\{Tv_1, \dots, Tv_n\} \subseteq \text{Im}(T)$

Since $Tv_i \in \text{Im}(T)$ for all i , $\text{Span}\{Tv_1, \dots, Tv_n\} \subseteq \text{Im}(T)$. So $\text{Im}(T) = \text{Span}\{Tv_1, \dots, Tv_n\}$. ■

Proposition 35. Let A be an $m \times n$ matrix. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be given by $Tx = Ax$. Lecture 8

(i) $\text{Ker}(T)$ is the solution space to the equation $Ax = \mathbf{0}$.

(ii) $\text{Im}(T)$ is the column space $\text{CSp}(A)$.

(iii) $\dim \text{Im}(T)$ is the rank $\text{rk}(A)$.

(Compare with Example 30).

PROOF. (i): This is immediate from the definitions.

(ii): Take the standard basis e_1, \dots, e_n for \mathbb{R}^n , that is $e_i = (0, 0, \dots, 1, 0, 0)^T$, where the 1 is in the i th position. $T(e_i) = A(0, 0, \dots, 1, 0, 0)^T = c_i$, the i th column of A . By Proposition 34, $\text{Im}(T) = \text{Span}\{T_{e_1}, \dots, T_{e_n}\} = \text{Span}\{c_1, \dots, c_n\} = \text{CSp}(A)$.

(iii): By (ii), $\dim \text{Im}(T) = \dim \text{CSp}(A) = \text{column-rank} = \text{rk}(A)$ ■

Example 36. Define $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ by:

$$Tx = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 1 & 4 & 7 \end{pmatrix} x \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Question: Find a basis for $\text{Ker}(T)$ and $\text{Im}(T)$.

Answer: To find $\text{Ker}(T)$, we solve $Ax = \mathbf{0}$.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 4 & 7 & 0 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$$\implies x_1 - x_3 = 0; x_2 + 2x_3 = 0. \text{ So we must have } x_1 = x_3, x_2 = -2x_3$$

So a basis for $\text{Ker}(T)$ is $\{(1, -2, 1)^T\}$

For $\text{Im}(T)$, we notice from the row-reduced matrix above, that $\text{rk}(A) = 2$. So $\dim \text{Im}(T)$ is 2. Since $\text{Im}(T)$ is $\text{CSp}(A)$, a basis can be obtained by taking any two lin. indep. columns. So $\{(1, -1, 1)^T, (2, 0, 4)^T\}$ is a basis.

Theorem 37: Rank Nullity Theorem.

Let $T : V \rightarrow W$ be a linear transformation. Then $\dim \text{Im}(T) + \dim \text{Ker}(T) = \dim(V)$. ($\text{rank } T = \dim \text{Im}(T)$, $\text{nullity} = \dim \text{Ker}(T)$)

PROOF. Let $\{u_1, \dots, u_s\}$ be a basis for $\text{Ker}(T)$, and $\{w_1, \dots, w_r\}$ be a basis for $\text{Im}(T)$. For each $i \in \{1, \dots, r\}$, there exists some $v_i \in V$ such that $T(v_i) = w_i$ (since $w_i \in \text{Im}(T)$).

I claim that $B = \{u_1, \dots, u_s\} \cup \{v_1, \dots, v_r\}$ is a basis for V . This needs to be proved:

(i) Show that $V = \text{Span}(B)$

Take $v \in V$. Then $Tv \in \text{Im}(T)$, and so $Tv = \mu_1 w_1 + \dots + \mu_r w_r$, for some $\mu_i \in F$. Define

$\bar{v} \in V$ by $\mu_1 v_1 + \dots + \mu_r v_r$. Then $T\bar{v} = \mu_1 w_1 + \dots + \mu_r w_r = Tv$.

So $v - \bar{v} \in \text{Ker}(T)$ by Proposition 33. So $v - \bar{v} = l_1 u_1 + \dots + l_s u_s$ for some $l_i \in F$. Now $v = \bar{v} + l_1 u_1 + \dots + l_s u_s = \mu_1 v_1 + \dots + \mu_r v_r + l_1 u_1 + \dots + l_s u_s \in \text{Span}(B)$.

(ii) Show that B is linearly independent. Suppose that:

$$l_1 u_1 + \dots + l_s u_s + \mu_1 v_1 + \dots + \mu_r v_r = \mathbf{0} \quad (*)$$

We want to show that $l_i = 0$ and $\mu_i = 0$ for all i . Apply T to both sides of $(*)$. Since $Tu_i = \mathbf{0}$ for all i , we get $\mu_1 w_1 + \dots + \mu_r w_r = \mathbf{0}$. But $\{w_1, \dots, w_r\}$ is linearly independent (a basis for $\text{Im}(T)$). So $\mu_i = 0 \quad \forall i$. Now from $(*)$, we have:

$$l_1 u_1 + \dots + l_s u_s = \mathbf{0}$$

But $\{u_1, \dots, u_s\}$ is linearly independent (a basis for $\text{Ker}(T)$). So $l_i = 0$ for all i . Hence B is a basis.

So $\dim V = |B| = r + s = \dim \text{Im}(T) + \dim \text{Ker}(T)$. ■

Corollary 38. *Let A be an $m \times n$ matrix. Let U be the solution space to $Ax = \mathbf{0}$. Then:*

$$\dim U = n - \text{rk}(A)$$

PROOF. Let T be the matrix transformation $\mathbb{R}^n \rightarrow \mathbb{R}^m$ given by $Tx = Ax$. Then $\text{rk}(A) = \dim \text{Im}(T)$, and $\text{Ker}(T) = U$. So the result follows by Theorem 37. ■

Matrix of Linear Transformation

Lecture 9

Definition 39. Let V be a vector space over a field F . Let $B = \{v_1, \dots, v_n\}$ be a basis for V . (Actually we should write (v_1, \dots, v_n) , because we care about the order of the basis vectors. But that notation could be confusing, so we'll continue to use set brackets).

Any $v \in V$ can be written uniquely as a linear combination $v = \lambda_1 v_1 + \dots + \lambda_n v_n$. Define the *vector of V with respect to B* to be $[V]_B = (l_1 \dots l_n)^T$.

Examples 40.

(a) Let $V = \mathbb{R}^n$, and let E be the standard basis $\{e_1, \dots, e_n\}$, where $e_i = (0, \dots, 1, \dots, 0)^T$ (1 in the i th position). Let $v = (a_1, \dots, a_n)^T$. Then $[V]_E = V$, since $v = a_1 e_1 + \dots + a_n e_n$.

(b) Let $V = \mathbb{R}^2$, and let $B = \{(1, 1)^T, (0, 1)^T\}$. Let $v = (1, 3)^T$. What is $[v]_B$?

We have to solve $V = l_1 v_1 + l_2 v_2$. We have the matrix equation:

$$\begin{pmatrix} | & | \\ v_1 & v_2 \\ | & | \end{pmatrix} \begin{pmatrix} l_1 \\ l_2 \end{pmatrix} = v. \quad \text{So:} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} l_1 \\ l_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

Solve by row reducing:

$$\left(\begin{array}{cc|c} 1 & 0 & 1 \\ 1 & 1 & 3 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array} \right)$$

from which we find $l_1 = 1, l_2 = 2$. So $[v]_B = (1, 2)^T$

Definition 41. Let V be a vector space of dimension n over F . Let $B = \{v_1, \dots, v_n\}$ be a basis. Let $T : V \rightarrow V$ be a linear transformation. For all $i \in \{1, \dots, n\}$, we have $Tv_i = a_{1i}v_1 + a_{2i}v_2 + \dots + a_{ni}v_n$.

The matrix of T with respect to the basis B , is:

$$[T]_B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Notice that the i th column of $[T]_B$ is $[Tv_i]_B$.

Examples 42.

(a) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is given by:

$$T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 \\ x_1 + 2x_2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\begin{aligned}
 \text{So } Tv &= \sum_{i=1}^n l_i \left(\sum_{j=1}^n a_{ji} v_j \right) \\
 &= \sum_{j=1}^n \left(\sum_{i=1}^n l_i a_{ji} \right) v_j
 \end{aligned}$$

(interchanging the order of summation). So

$$[Tv]_B = \begin{pmatrix} \sum_{i=1}^n l_i a_{1i} \\ \sum_{i=1}^n l_i a_{2i} \\ \vdots \\ \sum_{i=1}^n l_i a_{ni} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix}$$

$$= [T]_B [v]_B \text{ as required.}$$

■

Eigenvalues and Eigenvectors

Definition 44. Let $T : V \rightarrow V$ be a linear transformation. We say that v , is an *eigenvector* of T if $v \neq \mathbf{0}$, and $Tv = lv$ for some $l \in F$. We say that l is an *eigenvalue* of T .

Example 45. Let $V = \mathbb{R}^2$ and let $T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3x_1 + x_2 \\ -x_1 + x_2 \end{pmatrix}$.

We see that $T \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, and so $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is an eigenvector of T with eigenvalue 2.

Note that $T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ where $A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$.

We have $T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = l \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \iff A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = l \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

and so the eigenvectors and eigenvalues of T are the same as those of A . (We know how to find those from M1GLA). This will always work if T is a matrix transformation.

How do we find eigenvalues and eigenvectors in general?

Use the fact that we can represent any $T : V \rightarrow V$ as a matrix transformation.

Proposition 46. Let $T : V \rightarrow V$ and let $B = \{v_1, \dots, v_n\}$ be a basis for V . Then:

- (i) Eigenvalues of T are the same as eigenvalues of the matrix $[T]_B$
- (ii) The eigenvectors of T are those vectors v , such that $[v]_B$ is an eigenvector of $[T]_B$.
(So if $[v]_B = (l_1, \dots, l_n)^T$, then $v = l_1 v_1 + \dots + l_n v_n$.)

PROOF.

$$\begin{aligned} Tv = lv &\iff [Tv]_B = [lv]_B \\ &\iff [T]_B[v]_B = l[v]_B \text{ by Proposition 43} \\ &\iff [v]_B \text{ is an eigenvector for } [T]_B \text{ with eigenvalue } l \end{aligned}$$

■

Example 47. $V =$ Space of polynomials in x of degree ≤ 2 over \mathbb{R} . Let $T : V \rightarrow V$ be given by $T(f(x)) = f(x+1) - f(x)$ [Ex: Check T is linear]

Question: Calculate the eigenvalues and eigenvectors of T .

Answer: Let $B = \{1, x, x^2\}$, a basis for V . We have $T(1) = 0, T(x) = x + 1 - x =$

$$1, T(x^2) = (x+1)^2 - x^2 = 2x + 1.$$

$$[T]_B = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

The characteristic polynomial is then:

$$\begin{vmatrix} l & -1 & -1 \\ 0 & l & -2 \\ 0 & 0 & l \end{vmatrix} = l^3$$

So the only eigenvalue is 0. Find eigenvectors $[T]_B$ by solving:

$$\left(\begin{array}{ccc|c} 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

The solutions are vectors of the form $(a, 0, 0)^T$, so the eigenvectors of $[T]_B$ are $\{(a, 0, 0)^T \mid a \in F \setminus \{0\}\}$. So the eigenvectors of T are polynomials $f(x)$ such that $[f(x)]_B = (a, 0, 0)^T$ ($a \neq 0$).

So these are polynomials $a \cdot 1 + 0 \cdot x + 0 \cdot x^2$, which are the non-zero constant polynomials.

Diagonalisation of Linear Transformations

Proposition 48. Let $T : V \rightarrow V$ be a linear transformation. Let B be a basis for V . Then $[T]_B$ is a diagonal if and only if every basis vector in B is an eigenvector for T .

PROOF. Let e_1, \dots, e_n be the standard basis vectors, $(1, 0, \dots, 0)^T, (0, 1, \dots, 0)^T$ etc. in F^n , where $n = \dim V$. Let A be an $n \times n$ matrix. Then A is diagonal if and only if e_1 is an eigenvector of A for all i .

So $[T]_B$ is diagonal if and only if all the e_i are eigenvectors.

But $e_i = [v_i]_B$, where $B = \{v_1, \dots, v_n\}$, since $v_i = 0v_1 + 0v_2 + \dots + 1v_i + \dots + 0v_n$. So e_i is an eigenvector for $[T]_B$ if and only if v_i is an eigenvector for T , by Proposition 46. ■

Definition 49. A linear transformation $T : V \rightarrow V$ is *diagonalisable* if there is a basis of V such that every element of V is an eigenvector of T .

Examples 50.

- (a) V, T as in Example 47. $T(f(x)) = f(x+1) - f(x)$.
Is T diagonalisable?

We calculated earlier that the eigenvectors of T all lie in $\text{Span}\{1 + 0 \cdot x + 0 \cdot x^2\}$

which is a one-dimensional subspace of V . So the eigenvectors do not span V , and so there is no basis of V consisting of the eigenvector of T . So T is not diagonalisable.

(b) V as before (polynomial space of degree ≤ 2)

Define $T : V \rightarrow V$ by $T(f(x)) = f(1 - x)$, a basis for V .

(Exercise: check T is linear).

We have $T(1) = 1$, $T(x) = 1 - x$, $T(x^2) = (1 - x)^2 = 1 - 2x + x^2$. So:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

The characteristic polynomial is $(l - 1)^2(l + 1)$ [the diagonals are the roots].
So the eigenvalues are 1, -1.

We need to know whether there exists two linearly independent eigenvectors with eigenvalue 1. Using M1GLA techniques...

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

are both eigenvectors with eigenvalue 1. They are linearly independent. We have $\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$ as an eigenvector with eigenvalue -1.

$$\text{So } C = \{(1, 0, 0)^T, (0, 1, -1)^T, (1, -2, 0)^T\}$$

is a basis for V , whose elements are eigenvectors of T .
(Or $C = \{1, x - x^2, 1 - 2x\}$).

$$\text{So } T \text{ is diagonalisable and } [T]_C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Change of Basis

Let V be a vector space. Let $B = \{v_1, \dots, v_n\}$ and $C = \{w_1, \dots, w_n\}$.
For j in $\{1, \dots, n\}$, we can write $w_j = l_{1j}v_1 + l_{2j}v_2 + \dots + l_{nj}v_n$.

Lecture 11

Write:

$$P = \begin{pmatrix} l_{11} & l_{12} & \dots & l_{1n} \\ \vdots & & & \vdots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{pmatrix} = (l_{ij})$$

So the j th column of this matrix of P is $[w_j]_B$.

Proposition 51.

(i) $P = [T]_B$, where T is the unique linear transformation $V \rightarrow V$, such that $Tv_i = w_i$, for all i .

(ii) For all vectors $v \in V$, we have $P[v]_C = [v]_B$.

PROOF. (i) We know that $[T]_B[v_i]_B = [Tv_i]_B = [w_i]_B$. And $[T]_B$ is the only matrix with this property. So it is enough to show that $P[v_i]_B = [w_i]_B$.

But $[v_i]_B = e_i = (0, \dots, 1, \dots, 0)^T$ (1 in i th row), and so $P[v_i]_B = Pe_i = i$ th column of P is $[w_i]_B$, as above. ■

PROOF. (ii) First note that $P[w_i]_C = Pe_i = [w_i]_B$, as we saw in (i).

Now if $v \in V$ then we can write $v = a_1w_1 + \dots + a_nw_n$, $a_i \in F$. Now $[v]_C = (a_1, \dots, a_n)^T = a_1e_1 + \dots + a_ne_n$. So:

$$\begin{aligned} P[v]_C &= a_1Pe_1 + \dots + a_nPe_n \\ &= a_1[w_1]_B + \dots + a_n[w_n]_B \\ &= [a_1w_1 + \dots + a_nw_n]_B \quad \blacksquare \end{aligned}$$

Definition 52. The Matrix P as defined above is the *change of basis matrix* from B to C .

Proposition 53. Let V, B, C, P be all as above. Let $T : V \rightarrow V$ be a linear transformation.

- (i) P is invertible, and its inverse is the change of basis matrix from C to B .
- (ii) $[T]_C = P^{-1}[T]_BP$

PROOF. (i) Let Q be the change of basis matrix from C to B . Then $Q[v]_B = [v]_C$, for all $v \in V$. We calculate:

$$PQe_i = PQ[v_i]_B = P[v_i]_C = [v_i]_B = e_i$$

It follows that $PQ = I$, and so $Q = P^{-1}$. ■

PROOF. (ii)

$$\begin{aligned} P^{-1}[T]_BP e_i &= P^{-1}[T]_BP[w_i]_C \\ &= P^{-1}[T]_B[w_i]_B \\ &= P^{-1}[Tw_i]_B \\ &= [Tw_i]_C \\ &= [T]_C[w_i]_C \\ &= [T]_C e_i \end{aligned}$$

So the i th column of $P^{-1}[T]_BP$ is the same as the i th column of $[T]_C$ for all i , and so $P^{-1}[T]_BP = [T]_C$, as required. ■

Example 54. Let $V = \mathbb{R}^2$, and let $T : V \rightarrow V$ be given by:

$$T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -2x_1 + 3x_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Let $B = E = \{e_1, e_2\}$. Let $C = \{(1, 1)^T, (1, 2)^T\}$. The basis elements of C are eigenvectors of the matrix $\begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$.

The change of basis matrix from B to C is $P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$.

We have $[T]_B = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$

$$\begin{aligned} [T]_C &= P^{-1}[T]_B P \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \end{aligned}$$

so we have diagonalised T . (M1GLA - $D = P^{-1}AP$)

- End of Linear Algebra -

2 Group Theory

Groups

Lecture 12 Let S be a set. A *binary operation* on S is function $S \times S \rightarrow S$. S_{\times} takes pairs of elements, (S_1, S_2) , to some element $S_1 * S_2$.

Examples 55.

(i) $S = \mathbb{Z}$ (or $\mathbb{Q}, \mathbb{R}, \mathbb{C}$):

$$a * b = a + b, \text{ or } a * b = ab, \text{ or } a * b = a - b.$$

But *not* $a * b = a/b$. (Since, for example, b could be 0.)

(ii) $S = \setminus \{0\}$:

Now $a * b = a/b$ is a binary operation.

(iii) $S = \mathbb{Z}$ (or \mathbb{Q}, \mathbb{R})

$$a * b = \min(a, b)$$

(iv) S any set at all:

$$a * b = a$$

(v) $S = \{1, 2, 3\}$, $a * b$ defined by a table:

$a \backslash b$	1	2	3
1	1	2	1
2	2	1	2
3	1	2	3

Suppose that S is a set with binary operation $*$. Then the expression “ $a * b * c$ ” is ambiguous: it could mean $(a * b) * c$ or $a * (b * c)$.

Example: $S = \mathbb{Z}$, $a * b = a - b$. In general $(a - b) - c \neq a - (b - c)$. (except when $c = 0$).

Definition 56. If $(a * b) * c = a * (b * c)$ for all $a, b, c \in S$, then we say that $*$ is *associative*.

If an operation $*$ is associative, then $a * b * c$ is now unambiguous. So we can omit brackets in expressions of this sort.

Example 57 (For motivation).

Question: Solve $x + 1 = 2$, $x \in \mathbb{Z}$

Answer: (carefully explaining our reasoning)

(i) Use the fact that -1 in \mathbb{Z} . Add it to both sides.

$$(x + 1) + (-1) = 2 + (-1)$$

- (ii) Use associativity of $+$ to rewrite the left-hand side:

$$x + (1 + (-1)) = 2 + (-1) \implies x + 0 = 1$$

- (iii) Use the fact that $x + 0 = x$ for all x , so $x = 1$.

A group is a set with a binary operation $*$ in which equations $x * b = c$ can be solved for x in these three steps.

Definition 58. A *group* is a set G , with a binary operation $*$, satisfying the following axioms:

- (i) **Associativity:** $(a * b) * c = a * (b * c)$, for all $a, b, c \in G$
- (ii) **Identity Axiom:** There is an element $e \in G$ such that $x * e = e * x = x$ for all $x \in G$.
- (iii) **Inverses Axiom:** For every $x \in G$, there exists $y \in G$ such that $x * y = e$, where e is the element from the identity axiom.

Notes.

- (i) Most books also list another axiom:

0. **Closure:** If $x, y \in G$ then $x * y \in G$.

For us this is implied by the fact that $*$ is a binary operation on G .

- (ii) The element e from (2) is called the *identity element* of G .

- (iii) For $x \in G$, the element y from (3) is the *inverse* of x .

Examples 59.

$(G, *)$	Associativity	Identity	Inverse	Group
$(\mathbb{Z}, +)$	Yes	0	$-x$	Yes
$(\mathbb{Z}, -)$	No	No ($0 - n \neq n$)	No	No
(\mathbb{Z}, \times)	Yes	1	No	No
$(\mathbb{Q}, +)$	Yes	0	$-x$	Yes
(\mathbb{Q}, \times)	Yes	1	No, 0 has no inverse	No
$(\mathbb{Q} \setminus \{0\}, \times)$	Yes	1	$1/x$	Yes
$(\mathbb{R} \setminus \{0\}, \times)$	Yes	1	$1/x$	Yes
$(\mathbb{C} \setminus \{0\}, \times)$	Yes	1	$1/x$	Yes
$(\{1, -1, i, -i\} \subset \mathbb{C}, \times)$	Yes	1	$1/x$	Yes

Proposition 60. Let $(G, *)$ be a group.

- (i) G has exactly one identity element.
- (ii) Every element of G has exactly one inverse.
- (iii) (Left cancellation) If $x, y, z \in G$, and $x * y = x * z$, then $y = z$

(iv) (*Right cancellation*) If $x, y, z \in G$ and $x * z = y * z$, then $x = y$

Lecture 13 PROOF. (1) Let e, f be identity elements for G . Then $e * x = x$ for all $x \in G$. So $e * f = f$. But $x * f = x$ for all $x \in G$, so $e * f = e$. Hence $e = f$. ■

PROOF. (2) Suppose y, z be inverses for x . Look at $y * x * z$. Since $y * x = e$, we have $(y * x) * z = e * z = z$. But also $x * z = e$, and so $y * (x * z) = y * e = y$. So $y = z$. ■

PROOF. (3) [(4) similar] Let w be the inverse of x . Since $x * y = x * z$, we have $w * (x * y) = w * (x * z)$. By associativity $(w * x) * y = (w * x) * z$. But $w * x = e$, so $e * y = e * z$, and so $y = z$. ■

The two most common notations for groups are:

- (i) **Additive notation:** We write $+$ for $*$ and 0 for e . The inverse of x is $-x$. We write (for instance) $2x$ for $x + x$ etc. (This is normally used when the group operation really “is” addition in some set).
- (ii) **Multiplicative notation:** We write xy for $x * y$. We write e or 1 for the identity. Write x^{-1} for the inverse. Write (for instance) x^2 for $x * x$.

We will usually use multiplicative notation.

We say that a group G is *finite* if it has finitely many elements. In this case we say that $|G|$ is the *order* of G .

Examples 61.

- (i) Let F be a field of scalars (say $F = \mathbb{R}, \mathbb{C}$). Let S be the set of $n \times n$ matrices, with entries from F . Let $*$ be matrix multiplication. Is $(S, *)$ a group?

Certainly $*$ is a binary operation on S , since if $A, B \in S$, then $AB \in S$.

- ASSOCIATIVITY: Yes. $(AB)C = A(BC)$
- IDENTITY: Yes. I_n .
- INVERSES: No. Non-invertible matrices exist.

So S is not a group.

- (ii) Let G be the set of invertible $n \times n$ matrices over F . Let $*$ be matrix multiplication. Is $(G, *)$ a group?

Check that $*$ is a binary operation on G . If A has inverse A^{-1} and if B has inverse B^{-1} . Then AB has inverse $B^{-1}A^{-1}$. So if $A, B \in G$, then $AB \in G$.

- ASSOCIATIVITY: Yes, as above.
- IDENTITY: Yes, $I_n \in G$.
- INVERSE: Yes, by definition of G .

So G is a group - the general linear group of dimension n over F . We write $GL_n(F)$.

Notice $GL_1(F) = \{(x) : x \in F \setminus \{0\}\}$. So this is really just $(F \setminus \{0\}, \times)$ in disguise. $GL_2(F) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in F, ad - bc \neq 0 \right\}$

Definition 62.

- (i) Let G be a group (we write G multiplicatively.) Let $a, b \in G$. If $ab = ba$, then we say that a and b *commute*.
- (ii) If $xy = yx$ for all $x, y \in G$, then G is *abelian*.
[Niels Henrik Abel 1802-1829
- like many great mathematicians in his time,
he never met his 30th birthday]



Most groups are not abelian. Example: in $GL_2()$, we have:

$$\begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 2 & 0 \end{pmatrix}$$

But

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix}$$

So $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$ do not commute, so $GL_2(\mathbb{R})$ is not abelian.

But many of the groups we have seen so far are abelian.

Examples: $(\mathbb{Z}, +)$, $(F, +)$, $(F \setminus \{0\}, \times)$, $GL_1(F)$, $(\{1, -1, i, -i\}, \times)$ are all abelian.

Definition 63. Let X be a set. A function $f : X \rightarrow X$ is a *permutation* of X if it is a bijection. (Injective + surjective).

Examples 64.

- (i) $X = \{1, 2, 3, 4\}$. $f : 1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 4, 4 \mapsto 1$ is a permutation.
- (ii) $X = \mathbb{Z}$, $f : n \mapsto n + 3$ is a permutation.
- (iii) $X = \mathbb{Z}$, $f : n \mapsto 3n$ is *not* a permutation, (not surjective).

Notation for permutations (First attempt.)

Lecture 14

Assume that $X = \{1, \dots, n\}$. Let $f : X \rightarrow X$ be a permutation. We can write f as a matrix with two rows:

$$\begin{pmatrix} 1 & 2 & \dots & n \\ f(1) & f(2) & \dots & f(n) \end{pmatrix}$$

is called *two-row notation*.

Examples: if $f : 1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 4, 4 \mapsto 1$, we can write f as $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$. If g is $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$, then $g : 1 \mapsto 3, 2 \mapsto 1, 3 \mapsto 2, 4 \mapsto 4$.

Let f and g be permutations of a set X . We define the *composition*, $f \circ g$ by $f \circ g(x) = f(g(x))$ for all $x \in X$. For example, if f and g are as in the example above, we have $f \circ g(1) = 4, f \circ g(2) = 2, f \circ g(3) = 3, f \circ g(4) = 1$. So $f \circ g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$.

Proposition 65. *Let X be any set. Let S be the set of all permutations of X . Let \circ be the composition operation as above. Then (S, \circ) is a group, called the symmetric group on X , written $\text{Sym}(X)$.*

PROOF. We first check that \circ is a binary operation on S . Certainly if $f : X \rightarrow X$ and $g : X \rightarrow X$, then $f \circ g : X \rightarrow X$. The composition of two bijections is a bijection. So if $f, g \in S$, then $f \circ g \in S$.

Now the group axioms:

- (i) ASSOCIATIVITY: If $x \in X$, and $f, g, h \in S$, then $(f \circ g) \circ h(x) = (f \circ g)(h(x)) = f(g(h(x))) = f(g \circ h(x)) = f \circ (g \circ h(x))$. Since they agree on all $x \in X$, we have $(f \circ g) \circ h = f \circ (g \circ h)$.
- (ii) IDENTITY: Let e be the permutation defined by $e(x) = x$ for all $x \in X$. Then we have $e \circ f(x) = f(x) = f \circ e(x)$ for all $f \in S$. So $e \circ f = f \circ e = f$.
- (iii) INVERSES: Bijections have inverses. ■

Further notation: We almost always write symmetric groups multiplicatively. So write fg for $f \circ g$, and so on.

When $X = \{1, \dots, n\}$, we write S_n for the $\text{Sym}(X)$.

Examples 66. $S_1 = \text{Sym}(\{1\}) = \{e\} = (1, 1)^T$.

$$S_2 = \text{Sym}\{1, 2\} = \{e, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\}$$

$$S_3 = \text{Sym}\{1, 2, 3\} = \{e, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}\}.$$

We have $|S_1| = 1, |S_2| = 2, |S_3| = 6$

Proposition 67. *The group S_n has order $n!$*

PROOF. By induction. Inductive hypothesis. if X and Y are the two set of size n , then the number of bijections form X to Y is $n!$. (This gives us the result by taking $Y = X$.)

Base case $n = 1$ is obvious.

Inductive step: Suppose the result is true for n :

Let $|X| = |Y| = n + 1$. Take $x \in X, y \in Y$. The number of bijections $f : X \rightarrow Y$ such that $f(x) = y$ is equal to the number of bijections $X \setminus \{x\} \rightarrow Y \setminus \{y\}$, which is $n!$ by the inductive hypothesis. So the total number of bijections $X \rightarrow Y$ is $(n+1)n! = (n+1)!$. ■

The Group Table. Let G be a finite group. We can record the multiplication (binary operation) in G in a table called the *Group Table* or *Cayley Table* of G . If $G = \{a, b, c, \dots\}$, write:

	a	b	c	\dots
a	aa	ab	ac	
b	ba	bb	bc	\dots
c	ca	cb	cc	
\vdots		\vdots		

Example 68. Let $G = S_3$

Write $a = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$. Then $a^2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$, and $a^3 = e$.

Write $b = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$. Then $b^2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$, $a^2b = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$.

So $S_3 = \{e, a, a^2, b, ab, a^2b\}$. To work out the group table, it is helpful to check: $b^2 = e, ba = a^2b, ba^2 = ab$. Now it is easy to write down:

	e	a	a^2	b	ab	a^2b
e	e	a	a^2	b	ab	a^2b
a	a	a^2	e	ab	a^2b	b
a^2	a^2	e	a	a^2b	b	ab
b	b	a^2b	ab	e	a^2	a
ab	ab	b	a^2b	a	e	a^2
a^2b	a^2b	ab	b	a^2	a	e

Notice that every element of the group appears exactly once in each row and each column. (This follows from left and right cancellation laws.)

Subgroups

Definition 69. Let $(G, *)$ be a group. A *subgroup* of G is a subset of G which is itself a group under the operation $*$.

Examples 70.

- (i) $(\mathbb{Z}, +)$ is a subgroup of $(\mathbb{Q}, +)$. Both are subgroups of $(\mathbb{R}, +)$. All are subgroups of $(\mathbb{C}, +)$. But $(\mathbb{N}, +)$ is not a subgroup of any - it has no inverses.
- (ii) $(\mathbb{R} \setminus \{0\}, \times)$ is not a subgroup of $(\mathbb{R}, +)$. The group operation is different.
- (iii) $\{e\}$ is a subgroup of any group (where e is the identity element). The *trivial* subgroup.

- (iv) Every group is a subgroup of itself.

Lecture 15 *Recall:* a subgroup of a group G is a subset of G which is a group under the same operation as G .

Proposition 71. (*Subgroup Criteria*) Let G be a group. We write G multiplicatively. Let $H \subseteq G$ be a subset. Then H is a subgroup if and only if the following conditions hold:

- (i) $e \in H$, where e is the identity of G .
- (ii) If $a, b \in H$, then $ab \in H$, for all $a, b \in G$.
- (iii) If $a \in H$, then $a^{-1} \in H$, where a^{-1} is the inverse of a in G .

PROOF. “if”. Condition (2) says that the binary operation on G restricts to a binary operation on H . Since the operation is associative on G , it is also associative on H . Condition (1) gives us an identity, and Condition (3) gives inverses. So if (1),(2),(3) hold then H is a subgroup.

“only if”. Certainly (2) must hold if H is a subgroup, since we need the binary operation on G to restrict to a binary operation on H . If H is a subgroup, then H has an identity, e_H . Write e_G for the identity of G . Then $e_G e_H = e_H$, and also $e_H e_H = e_H$. Now $e_G = e_H$, by right cancellation. So $e_G \in H$, and so (1) holds.

Let $a \in H$. Let b be the inverse of a in G , and let c be the inverse of a in H . $ab = e_G = e_H = ac$. So $b = c$ by left cancellation. So $b \in H$, and so (3) holds. ■

Example 72. Let $G = GL_2(\mathbb{Z})$. For $n \in \mathbb{Z}$, define $u_n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$, so $u_n \in G$, for all $n \in \mathbb{Z}$. Define $U = \{u_n : n \in \mathbb{Z}\}$. Then U is a subgroup of G . Check the subgroup criteria:

- (i) $e_G = I = u_0 \in U$
- (ii) $u_m u_n = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & m+n \\ 0 & 1 \end{pmatrix} = u_{m+n} \in U$
- (iii) From (1) and (2), we see that $u_m^{-1} = u_{-m}$, so $u_m^{-1} \in U$. So U is a subgroup.

Notation: If H is a subgroup of G , we write $H \leq G$. (We can write $H < G$ if $H \neq G$).

Powers in groups: Let G be a group written multiplicatively. Let $g \in G$. We can write $g^1 = g$, $g^2 = gg$, $g^3 = ggg$, and so on. We also write $g^0 = e$, and $g^{-n} = (g^{-1})^n$. So now g^n is defined for all $n \in \mathbb{Z}$.

Proposition 73. (a) $(g^n)^{-1} = g^{-n}$.

(b) $g^m g^n = g^{m+n}$.

(c) $(g^m)^n = g^{mn}$.

Proof omitted.

Caution: It is not generally true that $a^n b^n = (ab)^n$. (Though true for Abelian Group)

Cyclic Subgroups

Proposition 74. Let G be a group, and let $g \in G$. Define $\langle g \rangle = \{g^n : n \in \mathbb{Z}\}$. Then $\langle g \rangle$ is a subgroup of G .

PROOF. Check subgroup criteria:

- (i) $g^0 = e$
- (ii) $g^m g^n = g^{m+n}$
- (iii) $(g^n)^{-1} = g^{-n}$ ■

Definition 75. The subgroup $\langle g \rangle$ is the *cyclic subgroup* generated by g .

Examples 76.

- (i) $U = \{u_n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} : n \in \mathbb{Z}\}$ is $\langle u_1 \rangle$. (Easy induction to show that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$ for all $n \in \mathbb{Z}$)
- (ii) In $(\mathbb{Z}, +)$, what is $\langle 3 \rangle$?
Note that we write this group additively, write ng instead of g^n . So $\langle 3 \rangle = \{n \cdot 3 : n \in \mathbb{Z}\}$. So $\langle 3 \rangle$ contains precisely all multiples of 3.
- (iii) $G = S_3$. What are the cyclic subgroups? From the group table we can easily calculate:
 - $\langle e \rangle = \{e\}$
 - $\langle a \rangle = \{e, a, a^2\}$, since $a^3 = e$.
 - $\langle a^2 \rangle = \{e, a, a^2\}$
 - $\langle b \rangle = \{e, b\}$ since $b^2 = e$
 - $\langle ab \rangle = \{e, ab\}$, since $(ab)^2 = e$
 - $\langle a^2 b \rangle = \{e, a^2 b\}$, since $(a^2 b)^2 = e$

Recall: $\langle g \rangle = \{g^n : n \in \mathbb{Z}\}$

Lecture 16

Definition 77. A group is *cyclic* if $G = \langle g \rangle$ for some $g \in G$. In this case g is a *generator* for G .

Examples 78.

- (i) $(\mathbb{Z}, +)$ is cyclic, since $\mathbb{Z} = \langle 1 \rangle$. Another generator is -1 . There are no other

generators.

- (ii) $\{1, -1, i, -i\}$ is cyclic, since it is equal to $\langle i \rangle$, and $\langle -i \rangle$. So i and $-i$ are generators. 1 and -1 are not.
- (iii) Let $\Omega_n = \{ \text{complex } n\text{th roots of unity} \}$, under complex multiplication. Let $w = e^{2\pi i/n}$. Then $\langle w \rangle = \{e^{2\pi i k/n} : k \in \mathbb{Z}\} = \Omega_n$. So Ω_n is a group - a cyclic subgroup of \mathbb{C} .

Since $|\Omega_n| = n$, it follows that there exists a group of order n for $n \in \mathbb{N}$.

- (iv) S_3 is *not* cyclic. We have calculated all of its cyclic subgroups (Example 76, 3), and none of them were equal to S_3 .

The Order of a Group Element

Definition 79. Let G be a group, and $g \in G$. The *order* of g is the least positive integer k , such that $g^k = e$, if such an integer exists. Otherwise g has infinite order. Write $\text{ord}(g)$ or $\text{o}(g)$ for the order of g . Write $\text{ord}(g) = \infty$ if g has infinite order.

Examples 80.

- (i) In any group G , the identity e has order 1. No other element has order 1. (If $g^1 = e$ then $g = e$.)
- (ii) Let $G = S_3$. Let $a = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$. Then $a^1 \neq e$, $a^2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \neq e$. But $a^3 = e$. So $\text{ord}(a) = 3$.
Let $b = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$. Then $b \neq e$, but $b^2 = e$. So $\text{ord}(b) = 2$. Can also check that $\text{ord}(a^2) = 3$, $\text{ord}(ab) = 2$, $\text{ord}(a^2b) = 2$.
- (iii) Let $G = (\mathbb{Z}, +)$. We know that $\text{ord}(0) = 1$. Suppose $n \neq 0$. Then since $\underbrace{n + n + \dots + n}_k = kn \neq 0$ for any positive integer k . We must have $\text{ord}(n) = \infty$.
- (iv) $G = GL_2(\mathbb{C})$. Let $A = \begin{pmatrix} i & 0 \\ 0 & e^{2\pi i/3} \end{pmatrix}$. What is the order of A ?

We see that for $k \in \mathbb{Z}$ we have $A^k = \begin{pmatrix} i^k & 0 \\ 0 & e^{2\pi i k/n} \end{pmatrix}$, which is equal to $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ if and only if $i^k = 1$ and $e^{2\pi i k/n} = 1$.

Now $i^k = 1$ if and only if $4 \mid k$, and $e^{2\pi i k/n} = 1$ if and only if $3 \mid k$. So $A^k = I$ if and only if $12 \mid k$. Since the smallest positive integer divisible by 12 is 12, we have $\text{ord}(A) = 12$.

Proposition 81. If G is a group and $g \in G$, then $|\langle g \rangle| = \text{ord}(g)$. ($\langle g \rangle$ is infinite $\iff \text{ord}(g) = \infty$).

PROOF. Suppose first that $\text{ord}(g) = \infty$. So $g^k \neq e$ for any $k \in \mathbb{N}$. We claim that if $m \neq n \in \mathbb{Z}$, then $g^m \neq g^n$. Suppose w.l.o.g. that $n > m$. Let $k = n - m$. Then $k \in \mathbb{N}$. Now $g^n = g^{m+k} = g^m g^k$. If $g^m = g^n$, then $g^m = g^m g^k$, and so $g^k = e$ by left cancellation. But this is a contradiction, and this proves the claim.

Now we have $g^0, g^1, g^2, g^3, \dots$ are all distinct elements of $\langle g \rangle$, and so $\langle g \rangle$ is infinite.

Now suppose that $\text{ord}(g) = k \in \mathbb{N}$. I claim that $\langle g \rangle = \{g^0, g^1, \dots, g^{k-1}\}$, and that the elements g^0, \dots, g^{k-1} , are distinct. (So $|\langle g \rangle| = k$). We use the fact that $n \in \mathbb{Z}$ can be written as $pk + q$, where $p, q \in \mathbb{Z}$ and $0 \leq q < k$. So $g^n = g^{pk+q} = (g^k)^p g^q = e^p g^q = g^q$. But g^q is one of the elements g^0, \dots, g^{k-1} , and so $\langle g \rangle = \{g^0, \dots, g^{k-1}\}$.

Suppose $g^i = g^j$, where $0 \leq i < j \leq k-1$. Let $l = j - i$. Then $g^i = g^j = g^{i+l}$, and so $g^l = e$ by left cancellation. But l is a positive integer less than k , so this contradicts the assumption that $\text{ord}(g) = k$. This proves the claim. ■

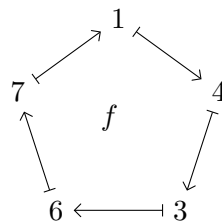
We have seen Proposition 81 illustrated in several examples already.

Examples 82.

- (i) Comparing the list of cyclic subgroups of S_3 with the list of the orders of elements, we saw that $\text{ord}(g) = |\langle g \rangle|$ in each case.
- (ii) In the case $G = (\mathbb{Z}, +)$, we saw that $\langle 3 \rangle = 3\mathbb{Z} = \{3k : k \in \mathbb{Z}\}$. So $|\langle 3 \rangle| = \infty$, and we have also seen that $\text{ord}(3) = \infty$.
- (iii) If $w = e^{2\pi i/n}$, then clearly $\text{ord}(w) = n$. And $\langle w \rangle = \Omega_n$ has order n too.

Cycles

Let $f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 5 & 6 & 3 & 2 & 7 & 1 & 8 \end{pmatrix}$. Looking at the successive images of 1 under the permutation f , we get back to 1 again after 5 steps via 4, 3, 6, 7. Lecture 17



We say that 1, 4, 3, 6, 7 forms a 5-cycle of f . Write $(1\ 4\ 3\ 6\ 7)$ for this cycle.

Similarly: $2 \xrightarrow{f} 5 \xrightarrow{f} 2$ is a 2-cycle, written $(2\ 5)$.

Finally 8 is a fixed point or 1-cycle of f . We can write (8) if we like – but usually we do not write out 1-cycles.

We can think of cycles as permutations in their own right. Take everything outside the cycle to be fixed.

So $(1\ 4\ 3\ 6\ 7)$ is the permutation: $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 2 & 6 & 3 & 5 & 7 & 1 & 8 \end{pmatrix}$.

And $(2\ 5)$ is the permutation: $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 5 & 3 & 4 & 2 & 6 & 7 & 8 \end{pmatrix}$.

And (8) represents the identity permutation.

Notice that $f = (1\ 4\ 3\ 6\ 7)(2\ 5)$ (where the cycles are multiplied by composition, as elements of S_8). So we have factorised f into cycles. These cycles have no common points – they are disjoint. This is the *disjoint cycle notation* for f .

Method 83. To calculate the disjoint cycle notation for a permutation $f \in S_n$.

Step 1. Pick the least element $i \in \{1, \dots, n\}$ which we haven't used yet. (Initially, we choose $i = 1$.) Open a new cycle with i .

Step 2. Continue the cycle with successive images of i under f until we get back to i again. Then close the cycle.

Step 3. If all $i \in \{1, \dots, n\}$ have appeared, then stop. Otherwise go back to Step 1.

Examples 84. $f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 6 & 1 & 7 & 5 & 8 & 2 & 4 & 10 & 9 \end{pmatrix}$.

Open a cycle with 1. Continue the cycle with $f(1) = 3$. Since $f(3) = 1$ again, close the cycle. $(1\ 3)$.

Now open a new cycle with 2. Continue it with 6, 8, 4, 7. Since $f(7) = 2$, we then close the cycle. $(2\ 6\ 8\ 4\ 7)$.

Start a cycle with 5. But 5 is a fixed-point of f , so close the cycle immediately. (5) .

Finally, start a cycle with 9. Continue it with $f(9) = 10$. But $f(10) = 9$, so close the cycle. $(9\ 10)$.

Now all points have appeared, so we stop. So the disjoint cycle notation for f is $f = (1\ 3)(2\ 6\ 8\ 4\ 7)(5)(9\ 10)$. We usually omit 1-cycles. So $f = (1\ 3)(2\ 6\ 8\ 4\ 7)(9\ 10)$.

Proposition 85. Method 83 always works. Every permutation in S_n can be written as a product of disjoint cycles.

PROOF. First we show that whenever we open a new cycle at Step 1, we are able to close it at Step 2. (We always get back to the starting point eventually.)

Suppose that x is the starting point. Since n is finite, the points $x, f(x), f^2(x), \dots$ cannot be all distinct. So there must exist some least k such that $f^k(x) = f^j(x)$, for some $j < k$. Suppose that $j \neq 0$. Then $f^{-1}f^k(x) = f^{-1}f^j(x)$, and so $f^{k-1}(x) = f^{j-1}(x)$.

But this contradicts the assumption that k was the *least* to give a repeat. So by contradiction, we must have $j = 0$. So $f^k(x) = f^0(x) = x$. So the first repeated term in

the cycle is x itself, and we close the cycle at that point.

Next we check that the cycles arising from Method 83. are disjoint. Let c and d be two cycles. Suppose that c starts with x and d starts with y . Suppose that c was constructed first. Then y cannot be in the cycle c (since otherwise we would not have used it to start a new cycle.)

Suppose z is in both c and d . So $z = f^i(x) = f^j(y)$ for some i, j . But now $f^{i-j}(x) = y$. This implies that y is in the cycle c , which is a contradiction. Hence c and d are disjoint. ■

Multiplication. To multiply permutations given in disjoint cycle notation, recall that $fg = f \circ g$, so $fg(x) = f(g(x))$. Now use Method 83 to get the disjoint cycle notation of fg .

Example: $f = (1\ 3\ 5)(2\ 4\ 6)$, $g = (1\ 2\ 3\ 4)(5\ 6) \in S_6$. Calculate $fg : 1 \mapsto 4 \mapsto 3 \mapsto 6 \mapsto 1$ $2 \mapsto 5 \mapsto 2$. Hence $fg = (1\ 4\ 3\ 6)(2\ 5)$.

Inverses. These are very easy. Just write the cycles backwards.

If $f = (1\ 3\ 5)(2\ 4\ 6)$, then $f^{-1} = (5\ 3\ 1)(6\ 4\ 2) = (1\ 5\ 3)(2\ 6\ 4)$.

Non-uniqueness.

- i. Order of the cycles doesn't matter.
- ii. The choice of starting point in each cycle doesn't matter.
- iii. We can include or exclude 1-cycles.

Example 86. In disjoint cycle notation, the group table for S_3 from Example 68 becomes:

Lecture 18

	e	(123)	(132)	(23)	(12)	(13)
e	e	(123)	(132)	(23)	(12)	(13)
(123)	(123)	(132)	e	(12)	(13)	(23)
(132)	(132)	e	(123)	(13)	(23)	(12)
(23)	(23)	(13)	(12)	e	(132)	(123)
(12)	(12)	(23)	(13)	(123)	e	(132)
(13)	(132)	(12)	(23)	(132)	(123)	e

Remark 87. Disjoint cycles commute. (If cycles c_1 and c_2 have no points in common, then $c_1c_2 = c_2c_1$). Example: $(12)(345) = (345)(12)$.

Cycles which are not disjoint don't usually commute. Example: $(12)(13) = (132)$. $(13)(12) = (123)$.

Definition 88. The *cycle shape* of a permutation is the sequence of cycle lengths in descending order of size when the permutation is written in disjoint cycle notation. We include 1-cycles.

Example: $f = (12)(3)(456)(7)(89) \in S_9$ then f has cycle shape $(3, 2, 2, 1, 1)$. We abbreviate this to $(3, 2^2, 1^2)$. The identity of S_9 has cycle shape $(1, 1, 1, 1, 1, 1, 1, 1, 1)$, or (1^9) .

Examples 89. What are the cycle shapes in S_4 , and how many elements are there of each shape?

Cycle shapes are given by weakly decreasing sequences of positive integers adding up to 4. There are (4) , $(3, 1)$, (2^2) , $(2, 1^2)$ and (1^4) . (These are the *partitions* of the integer 4). How many of each type?

(4) : May start the cycle at 1. There are 3 choices for $f(1)$. Then there are 2 choices for $f^2(1)$. This determines the cycle. So 6 possible 4-cycles. (These are (1234) , (1243) , (1324) , (1342) , (1423) , (1432) .)

$(3, 1)$: There are 4 choices for the fixed-point (or 1-cycle). Once the fixed point is chosen, there are 2 choices for the 3-cycle. So there are 8 possible permutations with this shape.

(2^2) : 1 is in a 2-cycle, and there are 3 choices for the other point in that cycle. This determines the permutation. So there are only 3 permutations with this shape. These are $(12)(34)$, $(13)(24)$, $(14)(23)$.

$(2, 1^2)$: There are $\binom{4}{2}$ ways of choosing the two fixed points. This determines the permutations. So there are 6 permutations.

(1^4) : Only the identity has this shape.

Check: we have $6 + 8 + 3 + 6 + 1 = 24 = |S_4|$.

Order of a Permutation

Proposition 90. Let G be a group, and let $g \in G$. Suppose $\text{ord}(g) = d$. Then for all $k \in \mathbb{Z}$, we have $g^k = e \iff d \mid k$.

PROOF. By Euclid's Lemma, we can write $k = xd + y$, where $x, y \in \mathbb{Z}$ and $0 \leq y < d$. Now $g^k = g^{xd+y} = (g^d)^x g^y = e^x g^y = g^y$. But $y < d$, and so we have $g^y = e \iff y = 0 \iff d \mid k$. ■

Proposition 91. Let G be a group, and let $a, b \in G$ be elements such that $ab = ba$. Then:

(i) $a^{-1}b = ba^{-1}$

(ii) $a^i b^j = b^j a^i$, for $i, j \in \mathbb{Z}$.

(iii) $(ab)^k = a^k b^k$, for $k \in \mathbb{Z}$.

PROOF.

1. We have $ab = ba$. So $a^{-1}(ab)a^{-1} = a^{-1}(ba)a^{-1}$. Hence $ba^{-1} = a^{-1}b$.
2. If $j < 0$, then replace b with b^{-1} . We know $ab^{-1} = b^{-1}a$ by 1. In this way, we may assume $j \geq 0$. Now by induction on j . [left as exercise.]
3. If $k < 0$, then write $d = a^{-1}$, $c = b^{-1}$. Then $(ab)^{-1} = cd$, and so $(ab)^k = (cd)^{-k}$. In this way we may assume $k \geq 0$. Now work by induction on k . [left as exercise.] ■

Proposition 92. Let f be a permutation with cycle shape (r_1, r_2, \dots, r_k) . Then $\text{ord}(f) = \text{lcm}(r_1, r_2, \dots, r_k)$.

PROOF. Write $f = c_1 c_2 \dots c_k$, where c_i has length r_i for all i , and the cycles c_i are disjoint from one another.

Lecture 19

Recall *Remark 87*: disjoint cycles commute. So $c_i c_j = c_j c_i$, for all i, j . Let $t \in \mathbb{Z}$.

Claim: $f^t = c_1^t c_2^t \dots c_k^t$.

PROOF OF CLAIM. First observe that $f = c_1, \dots, c_{k-1} c_k = c_k c_1, \dots, c_{k-1}$, since c_k commutes with all of the other cycles. So c_k commutes with $c_1 \dots c_{k-1}$. So by Prop 91.3, we have $f^t = ((c_1 \dots c_{k-1}) c_k)^t = (c_1 \dots c_{k-1})^t c_k^t$. Now an easy induction on k gives that $f^t = c_1^t c_2^t \dots c_k^t$, as required. ■

Continuing the proof of Prop 92, we see that $f^t = e \iff c_i^t = e$, for all i . But $c_i^t = e$ if and only if $r_i \mid t$. So $f^t = e \iff r_i \mid t$, for all $i \iff t$ is divisible by $\text{lcm}(r_1, r_2, \dots, r_k)$.

So $\text{ord}(f)$ is the last positive integer divisible by $\text{lcm}(r_1, \dots, r_k)$, which is $\text{lcm}(r_1, \dots, r_k)$ itself. ■

Examples 93.

- (i) $\text{ord}((12)(3456)) = \text{lcm}(2, 4) = 4$.
- (ii) $\text{ord}(13)(3456)) \neq 4$ — the cycles are not disjoint. We calculate $(13)(3456) = (13456)$, which has order 5.
- (iii) Clearly 1-cycles never affect the order of a permutation.
- (iv) Suppose we have a pack of 8 cards. We shuffle them by cutting the pack into 2 equal parts, and then interleaving the parts.
We get the permutation $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 5 & 2 & 6 & 3 & 7 & 4 & 8 \end{pmatrix}$. In disjoint cycle form, this is $(253)(467)$, which has order 3. So if we repeat our shuffle three times, the pack is back in its original order.

Exercise: How many shuffles do you need for a full pack of 52 cards to be shuffled back to its original order?

Lagrange's Theorem

Theorem 94: Lagrange's Theorem

Let G be a finite group. Let H be a subgroup of G . Then $|H|$ divides $|G|$.

Examples 95.

- (i) $|S_3| = 6$. We have seen that the cyclic subgroups of S_3 have orders 1, 2, 3. There are no other subgroups, except S_3 itself.
- (ii) $G = \{1, -1, i, -i\}$, order 4. The subgroups of G are $\{1\}$, $\{1, -1\}$, and G , with orders 1, 2, 4.
- (iii) Let $C_n = \langle w \rangle \leq \mathbb{C} \setminus \{0\}$, where $w = e^{2\pi i/n}$. If d divides n , then $w^{n/d} = e^{2\pi i/d}$, which has order d . So $\langle w^{n/d} \rangle$ is a subgroup of C_n with order d . In fact all of the subgroups of C_n have this form, for some d .

Remark. Although it is true for the elementary groups above, it is not true in general that a group G has a subgroup of order d for every divisor.

Example: S_5 has order 120, but has no subgroup of order 15, 30, 40.

Corollary 96. (to Lagrange's Theorem) If G is a finite group, and $g \in G$, then $\text{ord}(g)$ divides $|G|$.

PROOF. $\text{ord}(g) = |\langle g \rangle|$, and this divides $|G|$ by Lagrange's Theorem. ■

Corollary 97. Suppose $|G| = n$. Then $g^n = e$, for all $g \in G$.

PROOF. We know $\text{ord}(g) \mid n$ by Corollary 96. So $g^n = e$ by Prop 90. ■

Corollary 98. If $|G|$ is a prime number, then G is cyclic.

PROOF. Let $g \in G$ be such that $g \neq e$. Then $\text{ord}(g) \neq 1$, but $\text{ord}(g) \mid |G|$. So $\text{ord}(g) = |G|$ (Since $|G|$ is prime.) Hence $\langle g \rangle = G$, and so G is cyclic. ■

Some ideas for the proof of Lagrange's Theorem.

G is a finite group. $H \leq G$ is a subgroup. We shall divide the elements of G into disjoint subsets, so that every element is in exactly one subset, and each subset has the same size as H . Now if there are k subsets, then $|G| = k|H|$, and we are finished.

What are these subsets? One of these will be H itself. Now if $x \in G \setminus H$, we need a subset including x . We take $Hx = \{hx : h \in H\}$. Now if $y \in G \setminus (H \cup Hx)$, then we take Hy , and keep on going until we have used up all the elements of G .

Definition 99. Let G be a group, and let $H \leq G$. Let $x \in G$. The *right coset* Hx is the set $\{hx : h \in H\}$. [The *left coset* xH is the set $\{xh : h \in H\}$.]

Claim 100. $|Hx| = |H|$, for all $x \in G$.

PROOF. It is enough to show that $h_1x \neq h_2x$, if $h_1 \neq h_2$. But this is true by right cancellation, since $h_1x = h_2x \implies h_1 = h_2$. ■

Claim 101. Let $x, y \in G$. Then either $Hx = Hy$, or else $Hx \cap Hy = \emptyset$. (Two right cosets of H are equal or disjoint.)

PROOF. Suppose that Hx and Hy are not disjoint. So there exists $z \in Hx \cap Hy$. So there exists $h_1, h_2 \in H$ such that $z = h_1x = h_2y$. Now $y = h_2^{-1}h_1x$, and so $y \in Hx$. Now any hy in Hy is equal to $hh_2^{-1}h_1x$, which is in Hx . So $Hy \subseteq Hx$. By a similar argument, $Hx \subseteq Hy$ and so $Hx = Hy$. ■

Claim 102. $x \in Hx$ for all $x \in G$.

PROOF. $e \in H$, and so $x = ex \in Hx$. ■

Proof of Lagrange's Theorem.

Consider the set $S = \{Hx : x \in G\}$, a set of cosets.

- (i) Every $x \in G$ is in one member of S . (By Claim 102.)
- (ii) Every $x \in G$ is in no more than one member of S . (By Claim 101.)

It follows that $|G| = \sum_{X \in S} |X|$

- (iii) Every member of S has size $|H|$. (By Claim 100.)

Hence $|G| = k|H|$, where $|S| = k$. So $|H|$ divides $|G|$ as required. ■

Remarks 103. (Further Properties of cosets)

- (i) $H = Hx \iff x \in H$.

PROOF. Notice that $H = He$, so is a coset. If $x \in H$, then certainly $H \cap Hx \neq \emptyset$, so $H = Hx$, by Claim 101. If $x \notin H$, then $H \neq Hx$, since $x \in Hx$. ■

- (ii) If $y \in Hx$, then $Hx = Hy$.

PROOF. We have $y \in Hx \cap Hy$, and so $Hx \cap Hy \neq \emptyset$. So $Hx = Hy$, by Claim 101. ■

(iii) If $x \notin H$, then Hx is *not* a subgroup of G .

PROOF. $H \cap Hx = \emptyset$, by (1). Hence $e \notin Hx$. ■

(iv) $Hx = Hy$ if and only if $xy^{-1} \in H$.

PROOF. $Hx = Hy$ if and only if $x \in Hy$, and this occurs if and only if $x = hy$ for some $h \in H$. But this is equivalent to $xy^{-1} = h$. ■

Definition 104. Let G be a finite group, and $H \leq G$. The integer $\frac{|G|}{|H|}$ is the *index* of H in G , written $|G : H|$. So $|G : H|$ is the number of right cosets of H in G . (If is also the number of left cosets.)

Examples 105. $G = S_3$.

(i) $H = \langle (123) \rangle = \{e, (123), (132)\}$.

One coset is $He = H$. Since $\frac{|G|}{|H|} = \frac{6}{3} = 2$, there will be only one other coset. So this must be $\{(12), (13), (23)\}$.

(ii) $H = \langle (12) \rangle = \{e, (12)\}$. This time $\frac{|G|}{|H|} = \frac{6}{2} = 3$. So we are looking for two more cosets. We have:

$$\begin{aligned} H(123) &= \{e(123), (12)(123)\} = \{(123), (23)\}. \\ H(132) &= \{(132), (13)\}. \end{aligned}$$

Note that $H(23) = H(123)$, and $H(13) = H(132)$.

Remarks 106. (i) If $H = \{e\}$ then the right cosets of H in G are just singleton (1-element) sets $\{g\}$ for $g \in G$. (These are also the left cosets.)

(ii) If $H = G$ then the only right (or left) coset is H itself.

Modular Arithmetic

Lecture 21 **Recall:** Let $m \in \mathbb{N}$, and $a, b \in \mathbb{Z}$. We say that “ a is congruent to b modulo m ” if $a - b$ is divisible by m . (Equivalently: a and b give the same remainder when you divide by b .)

Write $a \equiv b \pmod{m}$.

Definition 107. Fix $m \in \mathbb{N}$. For $a \in \mathbb{Z}$, define the *residue class* of a modulo m , $[a]_m$, by $[a]_m = \{s \in \mathbb{Z} : s \equiv a \pmod{m}\} = \{km + a : k \in \mathbb{Z}\}$.

Examples 108. $[0]_5 = \{5k : k \in \mathbb{Z}\} = \{\dots, -15, -10, -5, 0, 5, 10, 15, \dots\}$.
 $[1]_5 = \{\dots, -14, -9, -4, 1, 6, 11, 16, \dots\}$

Recall that every integer $n \in \mathbb{Z}$ can be written as $n = qm + r$, where $q, r \in \mathbb{Z}$, with $0 \leq r < m$. Now $n \equiv r \pmod{m}$, so $[n]_m = [r]_m$. So $\mathbb{Z} = [0]_m \cup [1]_m \cup \cdots \cup [m-1]_m$. This is a *disjoint* union - every integer lies in exactly one of these residue classes.

Proposition 109. *For any $m \in \mathbb{N}$, the residue class $[0]_m$ is a subgroup of $(\mathbb{Z}, +)$. The other residue classes modulo m are cosets of $[0]_m$.*

PROOF. To check that $[0]_m$ is a subgroup, check the subgroup criteria:

- (i) We have $0 \in [0]_m$. So the identity is in $[0]_m$.
- (ii) Let $a, b \in [0]_m$. Then $a = km$ and $b = lm$ for some $k, l \in \mathbb{Z}$. Now $a + b = (k + l)m \in [0]_m$. So $[0]_m$ is closed under $+$.
- (iii) Let $a \in [0]_m$. Then $a = km$ for $k \in \mathbb{Z}$. Now the inverse of a is $-a = (-k)m$. So $-a \in [0]_m$. Hence $[0]_m$ is a subgroup.

Now $[r]_m = \{km + r : k \in \mathbb{Z}\} = \{x + r : x \in [0]_m\} = [0]_m + r$, which is the right coset of $[0]_m$ containing r . ■

Notation If m is fixed and understood, we then just write $[r]$ for $[r]_m$. We write \mathbb{Z}_m for the set of residue classes modulo m . So:

$\mathbb{Z}_m = \{[0]_m, [1]_m, \dots, [m-1]_m\}$, a set of size m .

Definition 110. (Binary operation on \mathbb{Z}_m .)

$$[a]_m + [b]_m = [a + b]_m.$$

$$[a]_m \times [b]_m = [ab]_m.$$

We need to make sure that these operations are well defined. Suppose $[a]_m = [a']_m$ and $[b]_m = [b']_m$. Then $a' = a + km$, and similarly $b' = b + lm$, for some $k, l \in \mathbb{Z}$. Now $a' + b' = a + km + b + lm = a + b + (k + l)m$. So $[a' + b']_m = [a + b]_m$.

And $a'b' = (a + km)(b + lm) = ab + (al + bk + klm)m$, so $[a'b']_m = [ab]_m$. So both of these operations are well defined.

Proposition 111. $(\mathbb{Z}_m, +)$ is a group.

PROOF. $([a]_m + [b]_m) + [c]_m = [(a + b) + c]_m = [a + (b + c)]_m = [a]_m + ([b]_m + [c]_m)$. (since $+$ is associative on \mathbb{Z} .)

$[0]_m$ is an identity. The inverse of $[a]_m$ is $[-a]_m$. ■

Note that (\mathbb{Z}_m, \times) is *not* a group, if $m > 1$. It is associative, and $[1]_m$ is an identity. But $[0]_m$ has no inverse, unless $[0]_m = [1]_m$. And $[0]_m \neq [1]_m$, unless $m = 1$.

How about $\mathbb{Z}_m^* = \mathbb{Z}_m \setminus \{[0]_m\}$?

Proposition 112. (\mathbb{Z}_m^*, \times) is a group if and only if m is a prime number.

PROOF. If $m = 1$, then $\mathbb{Z}_m^* = \emptyset$, which is not a group. So suppose $m > 1$. If m is *not* prime, we can write $m = ab$, where $1 < a \leq b < m$. Now m does not divide a or b , so $[a]_m \neq [0]_m$ and $[b]_m \neq [0]_m$. So $[a]_m, [b]_m \in \mathbb{Z}_m^*$. But $[a]_m \times [b]_m = [m]_m = [0]_m \notin \mathbb{Z}_m^*$. So \times is not a binary operation on \mathbb{Z}_m^* .

Now suppose m is prime. Property of prime numbers: If p is prime, and $p \mid ab$, then p divides at least one of a or b . Suppose $[a]_m, [b]_m \in \mathbb{Z}_m^*$. Then m does not divide a or b . So m does not divide ab , by the Property above. Hence $[ab]_m \in \mathbb{Z}_m^*$. So \times is a binary operation on \mathbb{Z}_m^* . Then the axioms:

ASSOCIATIVITY: same argument as for $+$ on \mathbb{Z}_m

IDENTITY: $[1]_m$.

INVERSES: If $[a]_m \in \mathbb{Z}_m^*$, then m does not divide a . So $\text{hcf}(m, a) = 1$. By Euclid's Algorithm, there exists $x, y \in \mathbb{Z}$, with $mx + ay = 1$. Now $ay \equiv 1 \pmod{m}$, and so $[a]_m \times [y]_m = [1]_m$. So $[a]_m^{-1} = [y]_m$. ■

Lecture 22

Examples 113.

- (i) $\mathbb{Z}_5^* = \{[1], [2], [3], [4]\}$. Is \mathbb{Z}_5^* cyclic? Check the powers of $[2]$:
 $[2]^2 = [4]$, $[2]^3 = [8] = [3]$, $[2]^4 = [16] = [1]$.

So $\text{ord}[2] = 4$, and so $\langle [2] \rangle = \mathbb{Z}_5^*$. So yes, \mathbb{Z}_5^* is cyclic.

Fact: \mathbb{Z}_p^* is cyclic for any prime p .

- (ii) In \mathbb{Z}_{31}^* , what is $[7]^{-1}$?
 We need to find $x, y \in \mathbb{Z}$ such that $31x + 7y = 1$. (Then $[y] = [7]^{-1}$)
 We use Euclid's algorithm:
 $31 = 4 \cdot 7 + 3$
 $7 = 2 \cdot 3 + 1$
 So $1 = 7 - 2 \cdot 3 = 7 - 2 \cdot (31 - 4 \cdot 7) = 9 \cdot 7 - 2 \cdot 31$
 So $x = -2, y = 9$. Hence $[7]^{-1} = [9]$

Theorem 114: Fermat's Little Theorem

Let p be a prime, and let $n \in \mathbb{Z}$. If p does not divide n , then $n^{p-1} \equiv 1 \pmod{p}$.

PROOF. If p does not divide n , then $[n]_p \in \mathbb{Z}_p^*$. So $\text{ord}[n]_p$ divides $|\mathbb{Z}_p^*| = p - 1$ by the Corollary to Lagrange's Theorem. Hence $[n]_p^{p-1} = [1]_p$. ■

Corollary 115. (Alternative statement of FLT). Let p be a prime, and let $n \in \mathbb{Z}$. Then $n^p \equiv n \pmod{p}$.

PROOF. If p does not divide n , then by FLT, $n^{p-1} \equiv n \pmod{p}$. So $n^p \equiv n \pmod{p}$. If p does divide n , then $[n]_p = [0]_p$. So $[n]_p^p = [0]_p^p = [0]_p = [n]_p$. ■

Example 116. Find the remainder when 6^{82} is divided by 17.

Answer: Note that $6^{82} = 6^{5 \cdot 16 + 2} = (6^{16})^5 (6^2) \equiv 1^5 6^2 \equiv 36 \equiv 2 \pmod{17}$, using FLT.

Definition 117. A *perfect number* is a natural number n which is the sum of its proper divisors. (i.e. divisors other than n itself.)

Examples: $6 = 1 + 2 + 3$. Also 28, 496, 8128, 35,550, 336, etc.

Theorem 118

If $2^m - 1$ is a prime number, then $n = 2^{m-1}(2^m - 1)$ is perfect. [Euclid] Every even perfect number is of this form. [Euler]

It is still unknown whether any odd perfect numbers exist. If they do, they are $> 10^{1500}$

Our interest is in the primes $2^m - 1$. These are *Mersenne primes*.

Examples: $3 = 2^2 - 1$, $7 = 2^3 - 1$, $31 = 2^5 - 1$, $127 = 2^7 - 1$. The largest known Mersenne prime is $2^{57,885,161} - 1$.

Proposition 119. *If $2^m - 1$ is prime, then m is prime. [The converse is not true]*

PROOF. Recall the polynomial identity:

$$(x^c - 1) = (x - 1)(x^{c-1} + x^{c-2} + \cdots + x + 1).$$

Suppose that m is not prime. So $m = ab$, where $a, b > 1$. Now

$$2^m - 1 = 2^{ab} - 1 = (2^a)^b - 1 = (2^a - 1)(2^{a(b-1)} + \cdots + 2^a + 1)$$

(using the polynomial identity with $x = 2^a$). So $2^a - 1$ divides $2^m - 1$. But since $1 < a < m$, we see that $2^a - 1 \neq 1, 2^m - 1$, so it is a proper divisor greater than 1. So $2^m - 1$ is not prime. ■

When is $2^p - 1$ prime? Use the group \mathbb{Z}_p^* :

Proposition 120. *Let $n = 2^p - 1$, where p is prime. Suppose that q is a prime divisor of n . Then $q \equiv 1 \pmod{p}$.*

PROOF. Since q divides n , we have q divides $2^p - 1$. So $2^p \equiv 1 \pmod{q}$.

So $[2]_q^p = [1]_q$. So $\text{ord}([2]_q)$ divides p . Since p is prime, $\text{ord}([2]_q)$ is 1 or p .

Now $2 \not\equiv 1 \pmod{q}$, so $[2]_q \neq [1]_q$. So $\text{ord}([2]_q) \neq 1$. Hence $\text{ord}([2]_q) = p$. So p divides

$|\mathbb{Z}_q^*| = q - 1$, by the Corollary to Lagrange's Theorem. Hence $q \equiv 1 \pmod{p}$. ■

Remark 121. What this tells us is that when we look for prime divisors of $n = 2^p - 1$ (to test primality), we need only test primes which are $\equiv 1 \pmod{p}$. (If $p = 57, 885, 161$, this is a serious saving!)

Exercise: Check that in fact, we only need check primes which are $1 \pmod{2p}$.

Dihedral Groups

Lecture 23

Definition 122. Let S be a subset of \mathbb{R}^n . Let $A \in GL_n(\mathbb{R})$. Write $AS = \{Av : v \in S\}$. If $AS = S$, we say A *preserves* S .

Proposition 123. For any $S \subseteq \mathbb{R}^n$, the set $H = \{A \in GL_n(\mathbb{R}) : A \text{ preserves } S\}$ is a subgroup of $GL_n(\mathbb{R})$. [The symmetry group of S .]

PROOF. Check the subgroup criteria:

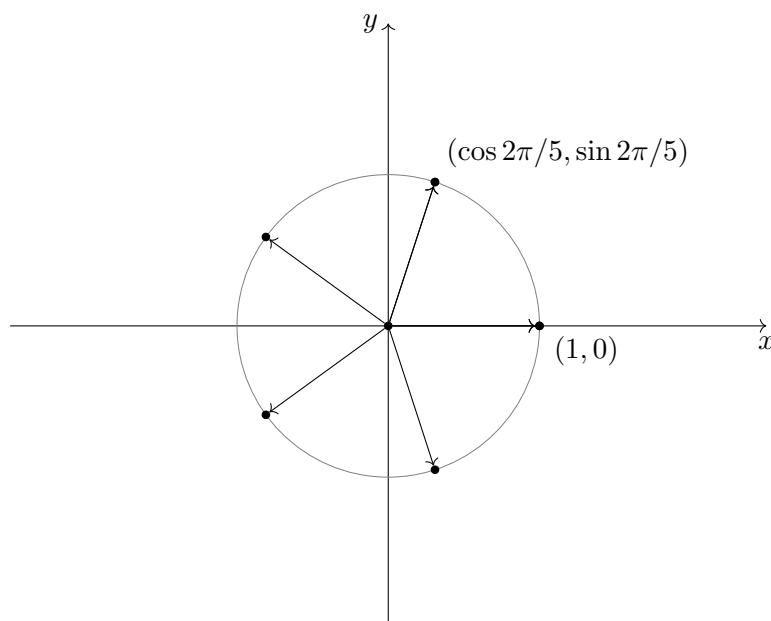
Certainly $IS = S$. So $I \in H$.

Suppose that $A, B \in H$. So $AS = S$, $BS = S$.

So $(AB)S = \{(AB)v : v \in S\} = \{(A(Bv)) : v \in S\}$. But $BS = S$, and so $Bv \in S \iff v \in S$. So $(AB)S = \{Av : v \in S\} = AS = S$. Hence $AB \in H$.

For inverses suppose $A \in H$. So $AS = S$. Now $A^{-1}S = A^{-1}(AS) = (A^{-1}A)S = IS = S$. Hence $H \leq GL_n(\mathbb{R})$. ■

Definition 124. Let $n > 2$. Let P_n be the regular n -gon in \mathbb{R}^2 with vertices $\begin{pmatrix} \cos 2\pi k/n \\ \sin 2\pi k/n \end{pmatrix}$, $k = 0, \dots, n-1$. The *dihedral group* D_{2n} is the symmetry group of P_n . So $D_{2n} = \{A \in GL_2(\mathbb{R}) : AP_n = P_n\}$.

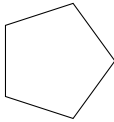


Remark We could instead take P_n to be the boundary of the n -gon, or just the set of vertices. We get the same symmetry group in each case. (Exercise: Check this)

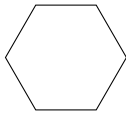
Proposition 125. *The group D_{2n} has order $2n$.*

PROOF. (1) List $2n$ elements:

Clearly the rotation of \mathbb{R}^2 through $2\pi/n$, certainly preserves P_n . So any power of $\begin{pmatrix} \cos 2\pi/n & -\sin 2\pi/n \\ \sin 2\pi/n & \cos 2\pi/n \end{pmatrix}$ preserves P_n . Now $|\langle A \rangle| = n$, so we have n rotations in D_{2n} . We also have n reflections. If n is odd:



Here is an axis of symmetry passing through any vertex, and the midpoint of the opposite edge. There are n vertices, so n reflections.



If n is even, there is an axis of symmetry passing through any pair of opposite vertices, or the midpoints of opposite edges. This gives n reflections in this case as well.

(2) Show D_{2n} has at most $2n$ elements:

If $A \in D_{2n}$ and if v is a vertex of P_n , then Av is a vertex of P_n (by the Remark above.) Now if w is a vertex of P_n which is adjacent to v , then $A\{v, w\}$ is also a pair of adjacent vertices. But $\{v, w\}$ is a basis for \mathbb{R}^2 . So once we know Av and Aw , we know what A is. Now the number of choices for Av and Aw is the number of pairs of adjacent vertices of P_n . There are $2n$ such pairs, so $|D_{2n}| \leq 2n$. Hence $|D_{2n}| = 2n$, as required. ■

Corollary 126. *Every element of D_{2n} is either a rotation or a reflection.*

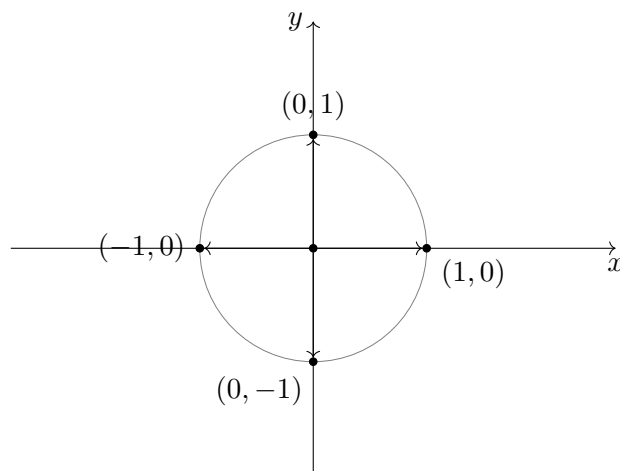
PROOF. The $2n$ elements, rotation and reflection, constructed in (1) is a complete list. ■

Remark. The rotations have determinant 1, and the reflections have determinant -1 .

Proposition 127. *The set of rotations in D_{2n} is a subgroup, of index 2. The reflections form a coset of this subgroup.*

PROOF. $\left\langle \begin{pmatrix} \cos 2\pi/n & -\sin 2\pi/n \\ \sin 2\pi/n & \cos 2\pi/n \end{pmatrix} \right\rangle$ is a cyclic subgroup of order n , containing all of the rotations in D_{2n} . There is only one other coset of this subgroup containing everything else - namely the n reflections. ■

Example 128. D_8 is the group of symmetries of a square.



Rotations: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

Angles: $0 \quad \pi/2 \quad \pi \quad 3\pi/2$

Reflections: $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$

3 Ring Theory

Many of the sets we've looked at so far have *two* binary operations, $+$ and \times . We may be interested in studying both at the same time. Lecture 24

For instance we might want to look at structures where equations like $x^2 + 1 = 2$, or $y^3 = x^2 + 2$ make sense, which involve both $+$ and \times .

Definition 129. A *ring* is a set R with two binary operations $+$ and \times such that the following axioms hold:

- (i) $(R, +)$ is an abelian group.
- (ii) $(x \times y) \times z = x \times (y \times z)$ for all $x, y, z \in R$. [ASSOCIATIVITY]
- (iii) $x \times (y + z) = (x \times y) + (x \times z)$, and $(x + y) \times z = (x \times z) + (y \times z)$ for all $x, y, z \in R$ [DISTRIBUTIVITY] “multiplication distributes over addition”

R is a *ring with unity* if it also satisfies:

- (iv) There exists $1 \in R$ such that $1 \times x = x \times 1 = x$.

R is a *commutative ring* if it also satisfies:

- (v) $x \times y = y \times x$ for all $x, y \in R$.

In this course, all rings are assumed to be commutative rings with unity.

Examples 130.

- (i) $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ with their usual $+$ and \times .
- (ii) \mathbb{Z}_n for any $n > 0$, with $+$ and \times defined on residue classes as before.
- (iii) A non-commutative ring (so Axiom 5 fails): Let $n > 1$ and let $R = \text{Mat}_n(\mathbb{R})$, the set of all $n \times n$ real matrices with the usual $+$ and \times .

Remarks 131.

- (i) We write 0 for the identity of $(R, +)$.
- (ii) We can write $x.y$ or xy for $x \times y$.
- (iii) Fact: $0x = 0$ for any $x \in R$.

PROOF. $0x = 0x + 0$. But $0x = (0 + 0)x = 0x + 0x$ by Distributivity. So $0x + 0 = 0x + 0x$. Hence $0x = 0$ by left cancellation in $(R, +)$. ■

- (iv) It is possible to have a ring R in which $0 = 1$. But in this case R only has one element. This ring is the *trivial ring*. We have seen it before - it is $\mathbb{Z}_1 = \{[0]_1\}$.
- (v) The axioms do not say that multiplicative inverses exist. In fact if R is non-trivial, then it has at least one element with no inverse, since 0 is such an element.

PROOF. $0x = 0$ for any x . So $0x \neq 1$ for any x unless $0 = 1$. But in this case R is trivial. ■

Definition 132.

Let F be a field (e.g. \mathbb{R}). The *polynomial ring* $F[x]$ is the set of all polynomials in x with coefficients from F , with the usual $+$ and \times for polynomials. Example: $f_1(x) = x^2 + 1$, $f_2(x) = x - 3$.

$$(f_1 + f_2)(x) = x^2 + x - 2 \quad (f_1 f_2)(x) = x^3 - 3x^2 + x - 3$$

Definition 133. A non-zero polynomial $f(x)$ has a *degree*, which is the largest power of x occurring in $f(x)$. Write $\deg f(x)$ for the degree. Then we have $\deg f(x) \geq 0$ for all $f(x)$. $\deg f(x) = 0$ if and only if $f(x)$ is a non-zero constant polynomial.

We do not define the degree of the zero polynomial $f(x) = 0$. (People who do define it to be $-\infty$)

Exercise: Show that $F[x]$ is indeed a ring.

Definition 134. Let R be a ring with unity. A *unit* in R is an element u such that there exists $w \in R$ with $uw = wu = 1$. So a unit in R is an element with a multiplicative inverse.

Examples 135.

- (i) If R is \mathbb{Q} , \mathbb{R} , \mathbb{C} , then all of its non-zero elements are units.
- (ii) If $R = \mathbb{Z}$, then the units are $\{\pm 1\}$.
- (iii) If p is prime and $R = \mathbb{Z}_p$, then all non-zero elements are units, since \mathbb{Z}_p^* is a group under \times .
- (iv) Let $R = \mathbb{Z}_m$, with $n \neq 1$ and n is not prime. The units in \mathbb{Z}_m are the residue classes $[x]_m$ for $x \in \mathbb{Z}$ such that $\text{hcf}(x, m) = 1$ (if $\text{hcf}(a, b) = 1$, we say that a, b are *coprime*)

PROOF. Suppose that $[x]_m$ is a unit in \mathbb{Z}_m . Then there exists $[y]_m \in \mathbb{Z}_m$ such that $[x]_m[y]_m = [y]_m[x]_m = [1]_m$. But $[x]_m[y]_m = [xy]_m$, and so $xy \equiv 1 \pmod{m}$. Hence $xy - 1 = km$ for some $k \in \mathbb{Z}$. Now $xy - km = 1$. But $xy - km$ is divisible by $\text{hcf}(x, m)$. Hence $\text{hcf}(x, m) = 1$.

Conversely, suppose $\text{hcf}(x, m) = 1$. Then Euclid's Algorithm tells us that there exists $y, z \in \mathbb{Z}$ such that $xy + mz = 1$. Now $xy \equiv 1 \pmod{m}$ so $[x]_m[y]_m = [xy]_m = [1]_m$. Hence $[x]_m$ is a unit in \mathbb{Z}_m . ■

- (v) What are the units in $F[x]$, where F is a field?

Observation: Let $f_1(x), f_2(x)$ be non-zero polynomials. Then $\deg(f_1(x) \times$

$f_2(x)) = \deg f_1(x) + \deg f_2(x)$. In particular, it follows that $\deg (f_1(x) \times f_2(x)) \geq \deg f_1(x)$.

Suppose that $f_1(x)$ is a unit in $F[x]$. Then there exists $f_2(x)$ such that $f_1(x) \times f_2(x) = 1$. So $\deg f_1(x) \leq \deg 1 = 0$. Hence $\deg f_1(x) = 0$. Hence $f_1(x)$ is a constant polynomial.

Conversely if $f_1(x) = c$, a non-zero constant polynomial, then $f_2(x) = 1/c$ is an inverse for $f_1(x)$. So $f_1(x)$ is a unit. We have shown that units in $F[x]$ are precisely the non-zero constant polynomials.

- (vi) A non-commutative example - the units in the ring $\text{Mat}_n(F)$. ($n \times n$ matrices over F) are the invertible matrices. There are the elements of $GL_n(F)$.

Definition 136. Given a ring R with unity, we write R^\times for the set of units in R .

Theorem 137

For any ring R with unity, (R^\times, \times) is a group. [Called the unit group of R .]

PROOF. First check that \times gives a binary operation on R^\times . Suppose that $u_1, u_2 \in R^\times$. Then there exists $w_1, w_2 \in R^\times$ such that $u_1 w_1 = w_1 u_1 = 1$ and $u_2 w_2 = w_2 u_2 = 1$. $(u_1 u_2)(w_1 w_2) = u_1(u_2 w_2)w_1 = u_1 1 w_1 = u_1 w_1 = 1$. Also $(w_2 w_1)(u_1 u_2) = w_2(w_1 u_1)u_2 = w_2 1 u_2 = w_2 u_2 = 1$. So $w_2 w_1$ is an inverse for $u_1 u_2$, and so $u_1 u_2 \in R^\times$.

Now check the group axioms:

ASSOCIATIVITY is given by the ring axioms.

IDENTITY: 1 is a unit since $1 \times 1 = 1$. So we have an identity in R^\times

INVERSES: Let $u \in R^\times$. Then u has an inverse w in R . So $uw = wu = 1$. Now u is an inverse for w , and so $w \in R^\times$. Hence u has an inverse in R^\times . ■

Remarks:

- (i) The proof of Theorem 136 did not assume that R is commutative.
- (ii) Since all units lie in R^\times , and since any inverse to a unit is itself a unit, it follows that every unit in R has a *unique* inverse. (since this is true in the unit group.) So we can write u^{-1} for the inverse of the unit u .

Examples 138.

- (i) If R is $\text{Mat}_n(F)$, then the unit group is $GL_n(F)$.
- (ii) If R is \mathbb{Q} , \mathbb{R} , \mathbb{C} , or \mathbb{Z}_p where p is prime, then $R^\times = R^* = R \setminus \{0\}$.

Definition 139. A *field* is a commutative ring, F , such that $F^* = F \setminus \{0\}$ (Every non-zero element is a unit.)

Arithmetic of Rings

Definition 140. Let R be a ring. Let $a, b \in R$. Say that a *divides* b if there exists $c \in R$ such that $ac = b$. We write $a \mid b$ to mean that a divides b .

Examples 141.

- (i) If $R = \mathbb{Z}$ then “ a divides b ” means what we expect it to.
- (ii) If $R = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ or any other field, then $a \mid b$ for any $a, b \in R$, except when $a = 0$ and $b \neq 0$.
- (iii) In a polynomial ring $F[x]$, “divides” means what it ought to. For example $x - 1$ divides $x^2 - 2x + 1$ in $\mathbb{Q}[x]$.

Lecture 26

Proposition 142.

- (i) If $a \mid b$ then $a \mid bc$ for all $c \in R$.
- (ii) If $a \mid b$ and $a \mid c$ then $a \mid b + c$.
- (iii) If $a \mid b$ and $b \mid c$ then $a \mid c$.
- (iv) If $u \in R$ is a unit, then $u \mid b$ for all $b \in R$.
- (v) If $a \mid b$ and if u is a unit, then $au \mid b$.
- (vi) If u is a unit and $w \mid u$ then w is also a unit.
- (vii) $a \mid 0$ for any $a \in R$.
- (viii) If $0 \mid b$ then $b = 0$.

(Proof. Left as exercise)

Definition 143. Let R be a ring, and $r \in R$. We say that r is a *zero-divisor* if $\exists s \in R$ such that $s \neq 0$, and $rs = 0$. A ring R with no zero-divisors, apart from 0, is called an *integral domain*.

Examples 144.

- (i) $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_p$ (p prime), $F[x]$ (where $F = \mathbb{Q}, \mathbb{R}, \mathbb{C}$) are all integral domains.
- (ii) If $m > 1$, and m is not prime, then \mathbb{Z}_m is *not* an integral domain, since if we have $m = ab$, $1 < a < m$, then $[a]_m$ and $[b]_m$ are non-zero, but their product is $[m]_m = [0]_m$, and so $[a]_m$ has a zero-divisor.

Proposition 145. A unit in a ring R is not a zero-divisor (unless R is trivial.)

PROOF. Let u be a unit. Suppose that $ua = 0$ for some $a \in R$. Then $u^{-1}(ua) = u^{-1}0 = 0$. So (since $u^{-1}(ua) = 1a = a$) we have $a = 0$. Hence u is not a zero-divisor ■

Corollary 145. Any field is an integral domain.

PROOF. All non-zero elements in a field are units, and so they are not zero-divisors. ■

Proposition 146. *Let R be an integral domain. Let $a, b \in R$ be such that $a \mid b$ and $b \mid a$. Then there exists a unit $u \in R$ such that $b = au$.*

PROOF. Since $a \mid b$, there exists $c \in R$ with $b = ac$. And since $b \mid a$, there exists $d \in R$ such that $a = bd$. $(ac)d = bd = a$, and by associativity it follows that $a(cd) = a$. Hence $a(1 - cd) = 0$. So one of a or $1 - cd = 0$. If $a = 0$, then $b = 0$ since $a \mid b$, so $a \cdot 1 = b$. If $1 - cd = 0$, then $cd = 1$, and so d is an inverse for c , so c is a unit, with $b = ac$. ■

Definition 147. Let R be an integral domain. Let r be a non-zero, non-unit in R . We say that r is *irreducible* if it can't be written as $r = st$, where $s, t \in R$, and neither s nor t is a unit. Otherwise r is *reducible*.

Examples 148.

- (i) In \mathbb{Z} the irreducible elements are the primes (positive and negative). Note that if p is prime then p has two factorisations, $p = 1 \times p = (-1) \times (-p)$. Here 1 and -1 are the units in \mathbb{Z} .
- (ii) In a field there are no irreducible (or reducible) elements, since every element is either 0 or a unit.
- (iii) $f(x) = x - 3$ is irreducible in $\mathbb{Q}[x]$ or $\mathbb{R}[x]$ or $\mathbb{C}[x]$.
- (iv) $f(x) = x^2 - 2$ is irreducible in $\mathbb{Q}[x]$. But in $\mathbb{R}[x]$ or $\mathbb{C}[x]$, $f(x)$ is reducible since $x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$, and neither of these factors are units.
- (v) $f(x) = x^2 + 1$ is irreducible in $\mathbb{Q}[x]$ and $\mathbb{R}[x]$, but is reducible in $\mathbb{C}[x]$ as $x^2 + 1 = (x + i)(x - i)$.

Remark. In an integral domain, every element is exactly one of: zero, a unit, irreducible or reducible.

Proposition 149. *Let F be \mathbb{Q} , \mathbb{R} , \mathbb{C} and let $f(x) \in F[x]$. Let $a \in F$. Then $f(a) = 0$ if and only if $x - a \mid f(x)$.*

PROOF.

Lecture 27

Suppose that $x - a \mid f(x)$. Then $f(x) = (x - a)g(x)$, for some $g(x) \in F[x]$. Now $f(a) = (a - a)g(a) = 0$.

For the converse, suppose that $f(a) = 0$. Let $f(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n$, where $\alpha_0, \dots, \alpha_n \in F$. Then $f(x) - f(a) = (\alpha_0 - \alpha_0) + (\alpha_1 x - \alpha_1 a) + \cdots + (\alpha_n x^n - \alpha_n a^n) = \alpha_1(x - a) + \alpha_2(x^2 - a^2) + \cdots + \alpha_n(x^n - a^n)$.

Note that $x^k - a^k = (x - a)(x^{k-1} + x^{k-2}a + \cdots + a^{k-1})$. So $x - a$ divides $x^k - a^k$ for all $k \in \mathbb{N}$. Hence $x - a$ divides $f(x) - f(a)$. But $f(a) = 0$ by assumption, and so $x - a$ divides $f(x)$. ■

Theorem 150: Fundamental Theorem of Algebra

Let $f(x) \in \mathbb{C}[x]$ be a polynomial of degree greater than 0. Then $f(x)$ has a root in \mathbb{C} . (So there exists $a \in \mathbb{C}$ such that $f(a) = 0$.)

Proof. Next year's Complex Analysis course.

Corollary 151. *The irreducible elements in $\mathbb{C}[x]$ are the linear polynomials, $\alpha_1 x + \alpha_0$, where $\alpha_1 \neq 0$.*

PROOF. First show that $\alpha_1 x + \alpha_0$ is irreducible. Suppose that $\alpha_1 x + \alpha_0 = f(x)g(x)$. Then $\deg f(x) + \deg g(x) = 1$. So one of $f(x)$ or $g(x)$ has degree 0, and so is a unit in $\mathbb{C}[x]$. Hence $\alpha_1 x + \alpha_0$ is irreducible in $\mathbb{C}[x]$.

Conversely, suppose that $r(x)$ has degree d greater than 1. Then $r(x)$ has a root a by the Fundamental Theorem. So $r(a) = 0$, and so $x - a$ divides $r(x)$ by Proposition 149. So $r(x) = (x - a)s(x)$ for some $s(x) \in \mathbb{C}[x]$. Now $\deg s(x) = \deg r(x) - 1 = d - 1 > 0$, so neither $s(x)$ nor $r(x)$ are units in $\mathbb{C}[x]$. So $r(x)$ is reducible. ■

Exercise: What are the irreducible elements in the ring $\mathbb{R}[x]$?

The Rings $\mathbb{Z}[\sqrt{d}]$

Definition 152. Let $d \in \mathbb{Z}$ be a non-square. The ring $\mathbb{Z}[\sqrt{d}]$ is the set $\{x + y\sqrt{d} : x, y \in \mathbb{Z}\} \subseteq \mathbb{C}$, with the usual complex $+$ and \times .

Check that these give binary operations on our set:

We have $(x_1 + y_1\sqrt{d}) + (x_2 + y_2\sqrt{d}) = (x_1 + x_2) + (y_1 + y_2)\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$,

and $(x_1 + y_1\sqrt{d})(x_2 + y_2\sqrt{d}) = (x_1x_2 + dy_1y_2) + (x_1y_2 + x_2y_1)\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$.

The ring axioms are left as an exercise.

Example: $\mathbb{Z}[\sqrt{-1}] = \mathbb{Z}[i]$ is the set $\{x + yi : x, y \in \mathbb{Z}\}$, the ring of *Gaussian integers*.

Proposition 153. *Let d be a non-square in \mathbb{Z} . If $x_1 + y_1\sqrt{d} = x_2 + y_2\sqrt{d}$ for $x_1, x_2, y_1, y_2 \in \mathbb{Z}$, then $x_1 = x_2$ and $y_1 = y_2$.*

PROOF. Since d is a non-square in \mathbb{Z} , we know that $\sqrt{d} \notin \mathbb{Q}$. Suppose that $x_1 + y_1\sqrt{d} = x_2 + y_2\sqrt{d}$. Then $x_1 - x_2 = (y_2 - y_1)\sqrt{d}$. If $y_2 \neq y_1$, then $\sqrt{d} = \frac{x_1 - x_2}{y_2 - y_1} \in \mathbb{Q}$, which is a contradiction. So $y_2 = y_1$. Now $x_1 - x_2 = 0$, and so $x_2 = x_1$, as well. ■

Definition 154. Define a function $N : \mathbb{Z}[\sqrt{d}] \rightarrow \mathbb{Z}$, by $N(x + y\sqrt{d}) = x^2 - dy^2$. N is the *norm map* on $\mathbb{Z}[\sqrt{d}]$.

Examples 155.

- (i) In $\mathbb{Z}[i]$, the norm map is $x + iy \mapsto x^2 + y^2 = |x + iy|^2$.
- (ii) More generally, if $d < 0$, then the norm map is $x + y\sqrt{d} \mapsto x^2 - dy^2 = x^2 + |d|y^2 = |x + y\sqrt{d}|^2$ (since $x + y\sqrt{d} = x + y\sqrt{|d|}i$).
- (iii) In $\mathbb{Z}[\sqrt{2}]$, the norm map is $x + y\sqrt{2} \mapsto x^2 - 2y^2$. So for example $N(1 + \sqrt{2}) = -1$. When $d > 0$, the norm map can take negative values.

Proposition 156. (i) If $r \in \mathbb{Z}[\sqrt{d}]$, and if $N(r) = 0$, then $r = 0 = 0 + 0\sqrt{d}$

(ii) If $r, s \in \mathbb{Z}[\sqrt{d}]$, then $N(rs) = N(r)N(s)$.

PROOF (1). Let $r = x + y\sqrt{d}$, and suppose that $N(r) = 0$. Then $x^2 - dy^2 = 0$, and so $x^2 = dy^2$. Now if $y \neq 0$, then $d = \frac{x^2}{y^2} = \left(\frac{x}{y}\right)^2$, and so $\sqrt{d} \in \mathbb{Q}$. But this is impossible. Hence $y = 0$. Now $N(r) = x^2$, so $x^2 = 0$ and so $x = 0$. Proof of (2) in Homework Sheet. ■

Proposition 157. An element r of $\mathbb{Z}[\sqrt{d}]$ is a unit if and only if $N(r) = \pm 1$.

Lecture 28

PROOF. (only if) Suppose that r is a unit. Then r has an inverse r^{-1} . Now $rr^{-1} = 1$, and so $N(r)N(r^{-1}) = N(1)$. But $1 = 1 + 0\sqrt{d}$, so $N(1) = 1$. Hence $N(r)$ has the inverse $N(r^{-1})$ in \mathbb{Z} , so $N(r)$ is a unit in \mathbb{Z} . So $N(r) = \pm 1$.

(if) Let $r = x + y\sqrt{d}$, and suppose that $N(r) = \pm 1$. Let $s = x - y\sqrt{d}$. Now $rs = x^2 - dy^2 = N(r) = \pm 1$. If $N(r) = 1$ then s is an inverse for r . If $N(r) = -1$ then $-s$ is an inverse for r . In either case, r is a unit in $\mathbb{Z}[\sqrt{d}]$. ■

Remark. If $d = -1$, the $\mathbb{Z}[\sqrt{d}] = \mathbb{Z}[i]$ has 4 roots, $\pm 1, \pm i$. If $d < -1$, then $\mathbb{Z}[\sqrt{d}]$ has only two units, ± 1 . If $d > 0$ then $\mathbb{Z}[\sqrt{d}]$ has infinitely many units.

Remark. Elements that are irreducible in \mathbb{Z} [primes] need not be irreducible in $\mathbb{Z}[\sqrt{d}]$.

Examples 158.

- (i) 3 is not irreducible in $\mathbb{Z}[\sqrt{3}]$ since $3 = \sqrt{3} \times \sqrt{3}$, and $\sqrt{3}$ is not a unit.
- (ii) 2 is not irreducible in $\mathbb{Z}[i]$, since $2 = (1 + i)(1 - i)$.
- (iii) *Question:* Is 11 irreducible in $\mathbb{Z}[\sqrt{3}]$?

Answer: We have $N(11) = 11^2 = 121$. Suppose that $11 = ab$ in $\mathbb{Z}[\sqrt{-3}]$, where a and b are non-units. $N(a), N(b) \neq \pm 1$. We have $N(a)N(b) = N(11) = 121$. So we must have $N(a) = N(b) = \pm 11$. Since N takes non-negative values on $\mathbb{Z}[\sqrt{-3}]$, we have $N(a) = 11$. Let $a = x + y\sqrt{d}$. Then

$x^2 + 3y^2 = 11$. But it is easy to check that no such x, y exist in \mathbb{Z} , which is a contradiction. So no such a, b exist. So 11 is irreducible in $\mathbb{Z}[\sqrt{-3}]$.

Highest Common Factor / Greatest Common Divisor

Definition 159. Let R be a ring. Let $a, b, c \in R$. Then c is a *highest common factor* (hcf) or *greatest common divisor* (gcd) for a and b if

- (i) $c \mid a$ and $c \mid b$ (c is a common factor for a and b .)
- (ii) If $d \mid a$ and $d \mid b$ then $d \mid c$ for $d \in R$.

Remarks. We do not claim that a highest common factor necessarily exists for all $a, b \in R$. There exists rings R and elements a, b where this doesn't happen. Where hcf's do exist, they are not usually unique. For instance if $R = \mathbb{Z}$, and $a = 4$, $b = 6$, then the hcf's of a and b are $2, -2$.

Proposition 160. Let R be an integral domain. Let c be an hcf for a and b in R . Then d is a highest common factor for a and b if and only if $d = cu$, for some unit $u \in R$.

PROOF. (only if) Suppose that c and d are both hcf's for a and b . So by the 2nd condition in the definition of hcf, we have $c \mid d$ and $d \mid c$. Now by Proposition 146, we have $d = cu$ for a unit u .

(if) Let $d = cu$, where u is a unit. Since $c \mid a$ and $c \mid b$, we have $cu \mid a$ and $cu \mid b$, by Proposition 141.5. So $d \mid a$ and $d \mid b$, so d is a common factor for a and b . To show the 2nd hcf condition holds, let e be any common factor of a and b . So $e \mid a$ and $e \mid b$. Then $e \mid c$, since c is a hcf. So $e \mid cu$ by Proposition 141.3. Hence $e \mid d$. Hence d is an hcf for a and b . ■

Lecture 29 *Reminder.* Euclid's Lemma: If we have any two $a, b \in \mathbb{Z}$, with $b \neq 0$, then there exists q and $r \in \mathbb{Z}$, with $0 \leq r < |b|$, such that $a = qb + r$.

Lemma 161. Let $f(x)$ and $g(x)$ be polynomials with coefficients from F , (\mathbb{Q}, \mathbb{C}) , with $g(x) \neq 0$. There exist polynomials $q(x)$ and $r(x) \in F[x]$ such that $f(x) = q(x)g(x) + r(x)$, and either $r(x) = 0$ or else $\deg r(x) < \deg g(x)$.

Example. Take $F = \mathbb{Q}$, $f(x) = x^4 + x^3 + 2x + 3$ and $g(x) = x^2 - 1$. Dividing $f(x)$ by $g(x)$, we get $f(x) = g(x)(x^2 + x + 1) + (3x + 4)$. So here $q(x) = x^2 + x + 1$, and $r(x) = 3x + 4$. Here $\deg r(x) = 1 < 2 = \deg g(x)$.

PROOF OF LEMMA 161. Take $\deg f(x) = m$ and $\deg g(x) = n$. If $m < n$, take $q(x) = 0$ and $r(x) = f(x)$, and this satisfies the lemma. So we will assume that $m \geq n$. Now we argue by induction on $m = \deg f$. The base case is $m = 0$ (and so $n = 0$ too). $f(x) = a$ and $g(x) = b$ for $a, b \in F$, with $b \neq 0$. Take $q(x) = \frac{a}{b}$, $r(x) = 0$, and this satisfies the lemma.

Inductive step: Assume the lemma holds for all polynomials $f(x)$ of degree $< m$. Suppose $\deg f(x) = m$. Put

$$f(x) = a_m x^m + \cdots + a_1 x + a_0 \text{ and } g(x) = b_n x^n + \cdots + b_1 x + b_0$$

Define $f'(x) = f(x) - \frac{am}{b_n} x^{m-n} g(x)$. The coefficient of x^m in $f'(x)$ is 0, so $\deg f'(x) < m$. So the inductive assumption applies to $f'(x)$, and so there exists $q'(x)$ and $r(x)$ such that $f'(x) = q'(x)g(x) + r(x)$, and with $r(x) = 0$ or $\deg r(x) < \deg g(x)$.

But now $f(x) = f'(x) + \frac{a_m}{b_n} x^{m-n} g(x) = (q'(x) + \frac{a_m}{b_n} x^{m-n})g(x) + r(x)$. So take $q(x) = q'(x) + \frac{a_m}{b_n} x^{m-n}$, and this satisfies the lemma. \blacksquare

Example: $f(x) = 3x^2 + 5x + 1$, $g(x) = 2x - 1$. Calculate $q(x)$ and $r(x)$ using polynomial long division:

[illegible]

So $q(x) = \frac{3}{2}x + \frac{13}{14}$, $r(x) = \frac{17}{4}$.

Proposition 162. *Let R be a ring. Let $a, b, q, r \in R$ such that $a = bq + r$. Then d is a highest common factor for a and b if and only if d is a highest common factor for b and r .*

PROOF. We actually show that the common factors of a and b are the same as those of b and r . Suppose d divides b and r . Then d divides bq and r . So d divides $bq + r = a$. So d divides a and b .

Conversely, suppose d divides a and b . Note that $r = a - bq$. Now d divides a and bq , so d divides $a - bq = r$. \blacksquare

Propositions 161 and 162 give us a Euclidean Algorithm for polynomials.

Example 163 (Euclidean Algorithm for polynomials.).

$$f(x) = x^3 - 2x^2 - 5x + 6, g(x) = x^2 - 2x - 3.$$

First find $q(x)$ and $r(x)$:

$$\begin{array}{r} x^2 - 2x - 3 \big) \overline{x^3 - 2x^2 - 5x + 6} \\ \underline{-x^3 + 2x^2 + 3x} \\ 0x^2 - 2x + 6 \\ \underline{0x^2 + 0x + 0} \\ -2x + 6 \end{array}$$

We have $q(x) = x$, $r(x) = -2x + 6$. Now we look for a hcf of $g(x)$ and $r(x)$. (A “smaller” problem than the original.) So applying Euclid’s Algorithm again, so now quotient is $-\frac{1}{2}x - \frac{1}{2}$, now remainder is 0. We’ve shown that $r(x) = -2x + 6$ divides $g(x)$. So $r(x)$ is a hcf for $r(x)$ and $g(x)$. So $r(x)$ is also a hcf for $f(x)$ and $g(x)$ by Proposition 162.

N.B. -2 is a unit in $\mathbb{Q}[x]$, with inverse $-\frac{1}{2}$. So $x-3$ is another hcf for $f(x)$ & $g(x)$.

Definition 164. Let R be an integral domain. A *Euclidean function* on R is a function $f : R \setminus \{0\} \rightarrow \mathbb{N} \cup \{0\}$, which satisfies the following two conditions:

- (i) $f(ab) \geq f(a)$ for $a, b \in R \setminus \{0\}$
- (ii) For all $a, b \in R$, $b \neq 0$, there exists $q, r \in R$ such that $a = qb + r$, either $r = 0$, or else $f(r) < f(b)$.

Examples 165.

- (i) If $R = \mathbb{Z}$ then $f(n) = |n|$ is a Euclidean function.
- (ii) If $R = F[x]$ then $\deg f(x)$ is a Euclidean function.

Euclidean functions are what we need for Euclidean algorithms:

Algorithm 166. To find a hcf for a and b in a ring R with a Euclidean function f .

Step 1. Find q and r such that $a = qb + r$ and r is 0 or $f(r) < f(b)$

Step 2. If $r = 0$ then b is a hcf for a and b . Otherwise:

Step 3. Start the algorithm again, replacing a with b and b with r .

Since $f(b) \in \mathbb{N} \cup \{0\}$, this algorithm must eventually terminate.

Definition 167. An integral domain with at least one Euclidean function is called a *Euclidean Domain*.

It can be very difficult to decide whether an Integral Domain is Euclidean.

Can we find a Euclidean function on $\mathbb{Z}[\sqrt{d}]$? Sometimes!

Proposition 168. Let $d \in \{-2, -1, 2, 3\}$. Then the function $f(a) = |N(a)|$ is a Euclidean function on $\mathbb{Z}[\sqrt{d}]$.

PROOF. The first condition from Definition 164 is easy. Since $|N(b)| \geq 1$ for $b \neq 0$, we have $|N(ab)| = |N(a)||N(b)| \geq |N(a)|$.

For the second condition, let $a = x + y\sqrt{d}$ and $b = v + w\sqrt{d}$, with $b \neq 0$. In \mathbb{C} , calculate:

$$\frac{a}{b} = \frac{x + y\sqrt{d}}{v + w\sqrt{d}} = \frac{(x + y\sqrt{d})(v - w\sqrt{d})}{v^2 - dw^2} = \frac{1}{N(b)}((xv - ywd) + (yv - xw)\sqrt{d}).$$

$$\text{Put } \alpha = \frac{xv - ywd}{N(b)}, \beta = \frac{yv - xw}{N(b)}, \in \mathbb{Q}$$

Set m, n to be the integers closest to α, β respectively. So

$$|\alpha - m| \leq \frac{1}{2} \quad \text{and} \quad |\beta - n| \leq \frac{1}{2} \quad (*)$$

Put $q = m + n\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$, and $r = a - bq$

We show that $|N(r)| < |N(b)|$:

Define $c = N(b)(\frac{a}{b} - q) = N(b)\frac{r}{b} \in \mathbb{C}$. So $bc = N(b)r$. We have:

$$\begin{aligned} c &= N(b)(\alpha + \beta\sqrt{d} - m - n\sqrt{d}) \\ &= N(b)(\alpha - m) + N(b)(\beta - n)\sqrt{d} \\ &\in \mathbb{Z}[\sqrt{d}], \text{ since } N(b)\alpha \text{ and } N(b)\beta \in \mathbb{Z} \end{aligned}$$

We have $N(bc) = N(b)N(c) = N(N(b)r) = N(b)^2N(r)$.

So $N(c) = N(b)N(r)$. But $N(c) = N(b)^2(\alpha - m)^2 - N(b)^2(\beta - n)^2d$. $N(r) = N(b)(\alpha - m)^2 - d(\beta - n)^2$.

Now from (*), $|\alpha - m|$ and $|\beta - n| \leq \frac{1}{2}$. So if $-2 \leq d \leq 3$, it is easy to see that $|(\alpha - m)^2 - d(\beta - n)^2| < 1$. So $|N(r)| < |N(b)|$ as required. ■

Example 169. Find a hcf of $a = 4 + \sqrt{2}$ and $b = 2 - 2\sqrt{2}$ in $\mathbb{Z}[\sqrt{2}]$.

Lecture 31

In \mathbb{C} we calculate:

$$\frac{a}{b} = \frac{4 + \sqrt{2}}{2 - 2\sqrt{2}} = \frac{(4 + \sqrt{2})(2 + 2\sqrt{2})}{(2 - 2\sqrt{2})(2 + 2\sqrt{2})} = -3 - \frac{5}{2}\sqrt{2}$$

So set $q = -3 - 2\sqrt{2}$. ($q = -3 - 3\sqrt{2}$ would also work.)

Now $r = a - bq = 2 - \sqrt{2}$. (Notice $|N(r)| = 4 - 2 = 2$, $|N(b)| = |4 - 8| = 4$). Continue, replacing a with b and b with r . In \mathbb{C} :

$$\frac{b}{r} = \frac{2 - 2\sqrt{2}}{2 - \sqrt{2}} = \sqrt{2} \in \mathbb{Z}[\sqrt{2}]$$

. So $q' = -\sqrt{2}$, $r' = 0$. So $r = 2 - \sqrt{2}$ divides $b = 2 - 2\sqrt{2}$, and so r is a hcf for b and r , hence r is a hcf for a and b .

Lemma 170 (Bézout's Lemma). *Let R be a Euclidean domain. Let a and b be elements of R and let d be a hcf for a and b . Then there exists $s, t \in R$ such that $as + bt = d$.*

PROOF. Define $X = X_{a,b} = \{ax + by : x, y \in R\}$. So $X \subseteq R$. The ring R has a Euclidean function f . So if a and b are not both 0, then X has non-zero elements. So there exists some least $n \in \mathbb{N} \cup \{0\}$ such that $f(x) = n$ for some $x \in X \setminus \{0\}$. Let c be some elements of $X \setminus \{0\}$ such that $f(c) = n$. We show that c is a hcf for a and b :

Since $c = ax + by$ for some x, y , it is clear that any common factor of x and y must divide c .

We know that there exists q and $r \in R$ such that $a = qc + r$, and either $r = 0$ or $f(r) < f(c)$. Now $c = ax + by$, so $r = a - qc = a - q(ax + by) = a(1 - qx) - b(qy) \in X$. We can't have $f(r) < f(c)$, since $f(c)$ is the smallest possible f for elements of X . Hence $r = 0$, and so $c \mid a$. A similar argument shows that $c \mid b$, and so c is a common factor - and hence a hcf - for a and b .

Now let d be any hcf of a and b . Then $d = cu$ for some unit $u \in R$. Now $d = (ax + by)u = axu + byu$. So put $r = xu$ and $s = yu$, and we're done. ■

Unique Factorisation

Definition 171. Let R be an integral domain, and let $a, b \in R$. We say a and b are *coprime* if 1 is a hcf for a and b . (Equivalently, any unit is a hcf.)

Proposition 172. Let R be a Euclidean Domain. Suppose a and b are coprime in R , and suppose a divides bc . Then a divides c .

PROOF. Since $a \mid bc$, we can write $ad = bc$, for some $d \in R$. Since R is Euclidean, and a and b are coprime, we use Bézout's Lemma: there exist $s, t \in R$ such that $as + bt = 1$. Now $c = 1c = (as + bt)c = asc + btc = asc + (bc)t = asc + (ad)t = a(sc + dt)$, which is divisible by a . ■

Proposition 172 can fail if R is not Euclidean.

Example: $R = \mathbb{Z}[\sqrt{-3}]$. $a = 2$, $b = 1 + \sqrt{-3}$, $c = 1 - \sqrt{-3}$. Then a and b are coprime, and 2 divides $bc = 4$. But a does not divide c . (This shows $\mathbb{Z}[\sqrt{-3}]$ is not Euclidean.)

Definition 173. A *unique factorisation domain* is an integral domain R with the following properties:

- (i) For every $a \in R$, not zero and not a unit, there exists irreducible elements p_1, \dots, p_s in R such that $a = p_1 \dots p_s$.
- (ii) Let p_1, \dots, p_s and q_1, \dots, q_t be irreducible in R , such that $p_1 \dots p_s = q_1 \dots q_t$. Then $t = s$, and reordering the q_i if necessary, we have $p_i = q_i u_i$ for some unit u_i , for all $i \in \{1, \dots, s\}$. (Factorisation is unique "up to units".)

Example: $R = \mathbb{Z}$. Then every element except 0, 1, -1 can be written as a product of irreducibles (primes), unique up to sign. $30 = 2 \times 3 \times 5 = (-2) \times 3 \times (-5)$.

Lecture 32 Not every integral domain is a UFD. We've seen that $\mathbb{Z}[\sqrt{-z}]$ is not a UFD.

Example 174. In $\mathbb{Z}[\sqrt{-5}]$ we can factorise 6 as 2×3 , and also as $(1 + \sqrt{-5})(1 - \sqrt{-5})$. Are these factors irreducible?

We have $N(2) = 4$, $N(3) = 9$. $N(1 \pm \sqrt{-5}) = 6$. If $ab = 2$ or 3 or $(1 \pm \sqrt{-5})$, and if neither a nor b is a unit, then $N(a)$ must be ± 2 or ± 3 . But it is easy to check that the equations $x^2 + 5y^2 = \pm 2$ or ± 3 has no solutions for $x, y \in \mathbb{Z}$. So all of

$2, 3, 1 \pm \sqrt{-5}$ are irreducible in $\mathbb{Z}[\sqrt{-5}]$.

Are the factorisations of 6 the same “up to units”? No, since $N(2) \neq N(1 \pm \sqrt{-5})$. So 6 is not uniquely factorisable in $\mathbb{Z}[\sqrt{-5}]$

Theorem 175

Every Euclidean Domain is a Unique Factorisation Domain.

Proof. Omitted. (In Algebra II (?)) 

It follows that $\mathbb{Z}[\sqrt{-5}]$ is not a Euclidean Domain, by Example 174.

Proposition 176. Let R be a UFD. Let p be an irreducible element of R , and let $a, b \in R$ be such that $p \mid ab$. Then either $p \mid a$ or $p \mid b$.

PROOF. If a or b is 0 or a unit, then the result is clear. So assume otherwise. So by the first UFD property, there exist irreducible elements $q_1, \dots, q_s, r_1, \dots, r_t$, such that $a = q_1 \dots q_s, b = r_1 \dots r_t$. So $ab = q_1 \dots q_s r_1 \dots r_t$.

Now $p \mid ab$, so $ab = pc$ for some $c \in R$. Suppose that $c = p_1 \dots p_u$ as a product of irreducibles. Then $ab = pp_1 \dots p_u$. Now by the second (uniqueness) property of UFDs, we have $p = q_i w$ or $r_i w$ for some i , and some unit, w . If $p = q_i w$ then $p \mid a$, and if $p = r_i w$ then $p \mid b$, as required. ■

Definition 177. A non-unit element, r , of a ring R is *prime* if it has the property that whenever $r \mid ab$ we have either $r \mid a$ or $r \mid b$.

In any integral domain, any prime element is irreducible. Why?

If r is prime, and $r = ab$, then $r \mid ab$, and then $r \mid a$ or $r \mid b$ by the prime property. Suppose $r \mid a$. Then since $a \mid r$, we have $r = au$ for a unit u . So $ab = au$, and so $a(b - u) = 0$. But R is an integral domain, so $b - u = 0$. Hence $b = u$, a unit.

In general the converse is not true - we’ve seen that 2 is irreducible but not prime in $\mathbb{Z}[\sqrt{-5}]$. (Example 174).

Felina. At the start of the Ring Theory section, we mentioned two equations:

- (i) $x^2 - 2 = -1$ we could have solved (in \mathbb{Z}) then.
- (ii) The other was $y^3 = x^2 + 2$. Can we find all solutions $x, y \in \mathbb{Z}$?

First notice that if $x^2 + 2 = y^3$ then both x and y are odd. (It is clear that if one of them is even, then so is the other. Suppose both are even. Then mod 4, we have $x^2 + 2 \equiv y^3$, and $x^2, y^3 \equiv 0$, so $2 \equiv 0 \pmod{4}$, a contradiction.)

Move into $\mathbb{Z}[\sqrt{-2}]$. We have $x^2 + 2 = (x + \sqrt{-2})(x - \sqrt{-2})$. Put $a = (x + \sqrt{-2})$ and $b = (x - \sqrt{-2})$. So $y^3 = ab$. Hence $N(y^3) = N(a)N(b)$.

Let d the hcf(a, b) in $\mathbb{Z}[\sqrt{-2}]$. Then d divides $a - b = 2\sqrt{-2}$. So $N(d)$ divides $N(2\sqrt{-2})$, so $N(d)$ divides 8. But $N(d)$ divides $N(a)$, so divides $N(y)^3 = y^6$, which is odd. So $N(d) = 1$.

So $\text{hcf}(a, b)$ is a unit, and so a and b are coprime. So we have $y^3 = ab$, where a, b are co-prime. Since $\mathbb{Z}[\sqrt{-2}]$ is a UFD, we can factorise $y^3 = r_1^3 \dots r_t^3$, where r_1, \dots, r_t are irreducible. Now for all i , we must have $r_i^3 \mid a$ or $r_i^3 \mid b$. So it's easy to see that $a = c^3$ and $b = d^3$ for $b, d \in \mathbb{Z}[\sqrt{-2}]$.

Let $c = m + n\sqrt{-2}$. So $(m + n\sqrt{-2})^3 = x + \sqrt{-2}$. So

$$(m + n\sqrt{-2})^3 = (m^3 - 6mn^2) + (3m^2n - 2n^3)\sqrt{-2} = x + \sqrt{-2}$$

So $m^3 - 6mn^2 = x$ and $3m^2n - 2n^3 = 1$. We have $(3m^2 - 2n^3)n = 1$, so $n \mid 1$, hence $n = \pm 1$, and $3m^2 - 2n^2 = n$, so $3m^2 - 2 = \pm 1$. So $n = \pm 1, m = \pm 1$. Now we have $x = \pm 5$. So $x^2 + 2 = 27, y = 3$. So the only solutions are $x = \pm 5, y = 3$.

- End of Algebra I -