Algebra I – Advanced Linear Algebra (MA251) Lecture Notes

Damiano Testa

Autumn 2013

Contents

1 Review of Some Linear Algebra

Students will need to be familiar with the whole of the contents of the First Year Linear Algebra module (MA106). In this section, we shall review the material on matrices of linear maps and change of basis. Other material will be reviewed as it arises.

1.1 The matrix of a linear map with respect to a fixed basis

Let V and W be vector spaces over a field[1] K. Let $T : V \to W$ be a linear map, where $\dim(V) = n$, $\dim(W) = m$. Choose a basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of V and a basis $\mathbf{f}_1, \ldots, \mathbf{f}_m$ of W.

Now, for $1 \leq j \leq n$, $T(\mathbf{e}_j) \in W$, so $T(\mathbf{e}_j)$ can be written uniquely as a linear combination of $\mathbf{f}_1, \ldots, \mathbf{f}_m$. Let

$$\begin{array}{rcl} T(\mathbf{e}_1) & = & \alpha_{11}\mathbf{f}_1 + \alpha_{21}\mathbf{f}_2 + \cdots + \alpha_{m1}\mathbf{f}_m \\ T(\mathbf{e}_2) & = & \alpha_{12}\mathbf{f}_1 + \alpha_{22}\mathbf{f}_2 + \cdots + \alpha_{m2}\mathbf{f}_m \\ \cdots & & \\ T(\mathbf{e}_n) & = & \alpha_{1n}\mathbf{f}_1 + \alpha_{2n}\mathbf{f}_2 + \cdots + \alpha_{mn}\mathbf{f}_m \end{array}$$

where the coefficients $\alpha_{ij} \in K$ (for $1 \leq i \leq m,\ 1 \leq j \leq n$) are uniquely determined.

The coefficients α_{ij} form an $m \times n$ matrix

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ & & \ldots & \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{pmatrix}$$

over K. Then A is called the matrix of the linear map T with respect to the chosen bases of V and W. Note that the columns of A are the images $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)$ of the basis vectors of V represented as column vectors with respect to the basis $\mathbf{f}_1, \ldots, \mathbf{f}_m$ of W.

It was shown in MA106 that T is uniquely determined by A, and so there is a one-one correspondence between linear maps $T : V \to W$ and $m \times n$ matrices over K, which depends on the choice of bases of V and W.

For $\mathbf{v} \in V$, we can write $\mathbf{v}$ uniquely as a linear combination of the basis vectors $\mathbf{e}_i$; that is, $\mathbf{v} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, where the x_i are uniquely determined by $\mathbf{v}$ and the basis $\mathbf{e}_i$. We shall call x_i the *coordinates* of $\mathbf{v}$ with respect to the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$. We associate the column vector

$$\underline{\mathbf{v}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in K^{n,1},$$

to $\mathbf{v}$, where $K^{n,1}$ denotes the space of $n \times 1$-column vectors with entries in K. Notice that $\underline{\mathbf{v}}$ is equal to $(x_1, x_2, \ldots, x_n)^{\mathrm{T}}$, the transpose of the row vector $(x_1, x_2, \ldots, x_n)$. To simplify the typography, we shall often write column vectors in this manner.

It was proved in MA106 that if A is the matrix of the linear map T, then for $\mathbf{v} \in V$, we have $T(\mathbf{v}) = \mathbf{w}$ if and only if $A\underline{\mathbf{v}} = \underline{\mathbf{w}}$, where $\underline{\mathbf{w}} \in K^{m,1}$ is the column vector associated with $\mathbf{w} \in W$.

[1]It is conventional to use either F or K as the letter to denote a field. F stands for a field, while K comes from the German word "körper".

1.2 Change of basis

Let V be a vector space of dimension n over a field K, and let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ be two bases of V. Then there is an invertible $n \times n$ matrix $P = (\sigma_{ij})$ such that

$$\mathbf{e}'_j = \sum_{i=1}^{n} \sigma_{ij} \mathbf{e}_i \quad \text{for } 1 \leq j \leq n. \qquad (*)$$

P is called the *basis change matrix* or *transition matrix* for the original basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and the new basis $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$. Note that the columns of P are the new basis vectors $\mathbf{e}'_i$ written as column vectors in the old basis vectors $\mathbf{e}_i$. (Recall also that P is the matrix of the identity map $V \to V$ using basis $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ in the domain and basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ in the codomain.)

Usually the original basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ will be the standard basis of K^n.

Example. Let $V = \mathbb{R}^3$, $\mathbf{e}_1 = (1\ 0\ 0)$, $\mathbf{e}_2 = (0\ 1\ 0)$, $\mathbf{e}_3 = (0\ 0\ 1)$ (the standard basis) and $\mathbf{e}'_1 = (0\ 1\ 2)$, $\mathbf{e}'_2 = (1\ 2\ 0)$, $\mathbf{e}'_3 = (-1\ 0\ 0)$. Then

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

The following result was proved in MA106.

Proposition 1.1 *With the above notation, let $\mathbf{v} \in V$, and let $\underline{\mathbf{v}}$ and $\underline{\mathbf{v}'}$ denote the column vectors associated with $\mathbf{v}$ when we use the bases $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$, respectively. Then $P\underline{\mathbf{v}'} = \underline{\mathbf{v}}$.*

So, in the example above, if we take $\mathbf{v} = (1\ \ -2\ 4) = \mathbf{e}_1 - 2\mathbf{e}_2 + 4\mathbf{e}_3$ then $\mathbf{v} = 2\mathbf{e}'_1 - 2\mathbf{e}'_2 - 3\mathbf{e}'_3$, and you can check that $P\underline{\mathbf{v}'} = \underline{\mathbf{v}}$.

This equation $P\underline{\mathbf{v}'} = \underline{\mathbf{v}}$ describes the change of coordinates associated with the basis change. In Section 3 below, such basis changes will arise as changes of coordinates, so we will use this relationship quite often.

Now let $T : V \to W$, $\mathbf{e}_i$, $\mathbf{f}_i$ and A be as in Subsection 1.1 above, and choose new bases $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ of V and $\mathbf{f}'_1, \ldots, \mathbf{f}'_m$ of W. Then

$$T(\mathbf{e}'_j) = \sum_{i=1}^{m} \beta_{ij} \mathbf{f}'_i \text{ for } 1 \leq j \leq n,$$

where $B = (\beta_{ij})$ is the $m \times n$ matrix of T with respect to the bases $\{\mathbf{e}'_i\}$ and $\{\mathbf{f}'_i\}$ of V and W. Let the $n \times n$ matrix $P = (\sigma_{ij})$ be the basis change matrix for original basis $\{\mathbf{e}_i\}$ and new basis $\{\mathbf{e}'_i\}$, and let the $m \times m$ matrix $Q = (\tau_{ij})$ be the basis change matrix for original basis $\{\mathbf{f}_i\}$ and new basis $\{\mathbf{f}'_i\}$. The following theorem was proved in MA106:

Theorem 1.2 *With the above notation, we have $AP = QB$, or equivalently $B = Q^{-1}AP$.*

In most of the applications in this course we will have $V = W(= K^n)$, $\{\mathbf{e}_i\} = \{\mathbf{e}'_i\}$, $\{\mathbf{f}_i\} = \{\mathbf{f}'_i\}$ and $P = Q$, and hence $B = P^{-1}AP$.

2 The Jordan Canonical Form

2.1 Introduction

Throughout this section V will be a vector space of dimension n over a field K, $T : V \to V$ will be a linear operator[2], and A will be the matrix of T with respect to a fixed basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$

[2] i.e. a linear map from a space to itself

of V. Our aim is to find a new basis $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ for V, such that the matrix of T with respect to the new basis is as simple as possible. Equivalently (by Theorem 1.2), we want to find an invertible matrix P (the associated basis change matrix) such that $P^{-1}AP$ is a simple as possible.

Our preferred form of matrix is a diagonal matrix, but we saw in MA106 that the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, for example, is not similar to a diagonal matrix. We shall generally assume that $K = \mathbb{C}$. This is to ensure that the characteristic polynomial of A factorises into linear factors. Under this assumption, it can be proved that A is always similar to a matrix $B = (\beta_{ij})$ of a certain type (called the *Jordan canonical form* or sometimes *Jordan normal form* of the matrix), which is not far off being diagonal. In fact β_{ij} is zero except when $j = i$ or $j = i+1$, and $\beta_{i,i+1}$ is either 0 or 1.

We start by summarising some definitions and results from MA106. We shall use $\mathbf{0}$ both for the zero vector in V and the zero $n \times n$ matrix. The zero linear operator $0_V : V \to V$ corresponds to the zero matrix $\mathbf{0}$, and the identity linear operator $I_V : V \to V$ corresponds to the identity $n \times n$ matrix I_n.

Because of the correspondence between linear maps and matrices, which respects addition and multiplication, all statements about A can be rephrased as equivalent statements about T. For example, if $p(x)$ is a polynomial equation in a variable x, then $p(A) = \mathbf{0} \Leftrightarrow p(T) = 0_V$.

If $T\mathbf{v} = \lambda\mathbf{v}$ for $\lambda \in K$ and $\mathbf{0} \neq \mathbf{v} \in V$, or equivalently, if $A\underline{\mathbf{v}} = \lambda\underline{\mathbf{v}}$, then λ is an *eigenvalue*, and $\mathbf{v}$ a corresponding *eigenvector* of T and A. The eigenvalues can be computed as the roots of the *characteristic polynomial* $c_A(x) = \det(A - xI_n)$ of A.

The eigenvectors corresponding to λ are the non-zero elements in the nullspace (= kernel) of the linear operator $T - \lambda I_V$ This nullspace is called the *eigenspace* of T with respect to the eigenvalue λ. In other words, the eigenspace is equal to $\{ \mathbf{v} \in V \mid T(\mathbf{v}) = \lambda\mathbf{v} \}$, which is equal to the set of eigenvectors together with $\mathbf{0}$.

The dimension of the eigenspace, which is called the *nullity* of $T - \lambda I_V$ is therefore equal to the number of linearly independent eigenvectors corresponding to λ. This number plays an important role in the theory of the Jordan canonical form. From the *Dimension Theorem*, proved in MA106, we know that

$$\mathrm{rank}(T - \lambda I_V) + \mathrm{nullity}(T - \lambda I_V) = n,$$

where $\mathrm{rank}(T - \lambda I_V)$ is equal to the dimension of the image of $T - \lambda I_V$.

For the sake of completeness, we shall now repeat the results proved in MA106 about the diagonalisability of matrices. We shall use the theorem that a set of n linearly independent vectors of V form a basis of V without further explicit reference.

Theorem 2.1 *Let $T : V \to V$ be a linear operator. Then the matrix of T is diagonal with respect to some basis of V if and only if V has a basis consisting of eigenvectors of T.*

Proof: Suppose that the matrix $A = (\alpha_{ij})$ of T is diagonal with respect to the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of V. Recall from Subsection 1.1 that the image of the i-th basis vector of V is represented by the i-th column of A. But since A is diagonal, this column has the single non-zero entry α_{ii}. Hence $T(\mathbf{e}_i) = \alpha_{ii}\mathbf{e}_i$, and so each basis vector $\mathbf{e}_i$ is an eigenvector of A.

Conversely, suppose that $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis of V consisting entirely of eigenvectors of T. Then, for each i, we have $T(\mathbf{e}_i) = \lambda_i\mathbf{e}_i$ for some $\lambda_i \in K$. But then the matrix of A with respect to this basis is the diagonal matrix $A = (\alpha_{ij})$ with $\alpha_{ii} = \lambda_i$ for each i. □

Theorem 2.2 *Let $\lambda_1, \ldots, \lambda_r$ be distinct eigenvalues of $T : V \to V$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_r$ be corresponding eigenvectors. (So $T(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ for $1 \leq i \leq r$.) Then $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly independent.*

PROOF: We prove this by induction on r. It is true for $r = 1$, because eigenvectors are non-zero by definition. For $r > 1$, suppose that for some $\alpha_1, \ldots, \alpha_r \in K$ we have

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_r \mathbf{v}_r = \mathbf{0}.$$

Then, applying T to this equation gives

$$\alpha_1 \lambda_1 \mathbf{v}_1 + \alpha_2 \lambda_2 \mathbf{v}_2 + \cdots + \alpha_r \lambda_r \mathbf{v}_r = \mathbf{0}.$$

Now, subtracting λ_1 times the first equation from the second gives

$$\alpha_2(\lambda_2 - \lambda_1)\mathbf{v}_2 + \cdots + \alpha_r(\lambda_r - \lambda_1)\mathbf{v}_r = \mathbf{0}.$$

By inductive hypothesis, $\mathbf{v}_2, \ldots, \mathbf{v}_r$ are linearly independent, so $\alpha_i(\lambda_i - \lambda_1) = 0$ for $2 \leq i \leq r$. But, by assumption, $\lambda_i - \lambda_1 \neq 0$ for $i > 1$, so we must have $\alpha_i = 0$ for $i > 1$. But then $\alpha_1 \mathbf{v}_1 = \mathbf{0}$, so α_1 is also zero. Thus $\alpha_i = 0$ for all i, which proves that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly independent. □

Corollary 2.3 *If the linear operator $T : V \to V$ (or equivalently the $n \times n$ matrix A) has n distinct eigenvalues, where $n = \dim(V)$, then T (or A) is diagonalisable.*

PROOF: Under the hypothesis, there are n linearly independent eigenvectors, which therefore form a basis of V. The result follows from Theorem 2.1. □

2.2 The Cayley-Hamilton theorem

This theorem says that a matrix satisfies its own characteristic equation. It is easy to visualise with the following "non-proof":

$$c_A(A) = \det(A - AI) = \det(0) = 0.$$

This argument is faulty because you cannot really plug the matrix A into $\det(A - xI)$: you must compute this polynomial first.

Theorem 2.4 (Cayley-Hamilton) *Let $c_A(x)$ be the characteristic polynomial of the $n \times n$ matrix A over an arbitrary field K. Then $c_A(A) = \mathbf{0}$.*

PROOF: Recall from MA106 that, for any $n \times n$ matrix B, we have $B\text{adj}(B) = \det(B)I_n$, where $\text{adj}(B)$ is the $n \times n$ matrix whose (j, i)-th entry is the cofactor $c_{ij} = (-1)^{i+j}\det(B_{ij})$, and B_{ij} is the $(n-1) \times (n-1)$ matrix obtained by deleting the i-th row and the j-th column of B.

By definition, $c_A(x) = \det(A - xI_n)$, and $(A - xI_n)\text{adj}(A - xI_n) = \det(A - xI_n)I_n$. Now $\det(A - xI_n)$ is a polynomial of degree n in x; that is $\det(A - xI_n) = a_0x^0 + a_1x^1 + \cdots + a_nx^n$, with $a_i \in K$. Similarly, putting $B = A - xI_n$ in the last paragraph, we see that the (j, i)-th entry $(-1)^{i+j}\det(B_{ij})$ of $\text{adj}(B)$ is a polynomial of degree at most $n - 1$ in x. Hence $\text{adj}(A - xI_n)$ is itself a polynomial of degree at most $n - 1$ in x in which the coefficients are $n \times n$ matrices over K. That is, $\text{adj}(A - xI_n) = B_0x^0 + B_1x + \cdots + B_{n-1}x^{n-1}$, where each B_i is an $n \times n$ matrix over K. So we have

$$(A - xI_n)(B_0x^0 + B_1x + \cdots + B_{n-1}x^{n-1}) = (a_0x^0 + a_1x^1 + \cdots + a_nx^n)I_n.$$

Since this is a polynomial identity, we can equate coefficients of the powers of x on the left and right hand sides. In the list of equations below, the equations on the left are the result of equating coefficients of x^i for $0 \leq i \leq n$, and those on right are obtained by multiplying A^i by the corresponding left hand equation.

$$\begin{array}{rcrcr} AB_0 & & & = & a_0 I_n, \\ AB_1 & - & B_0 & = & a_1 I_n, \\ AB_2 & - & B_1 & = & a_2 I_n, \\ & & & \cdots & \\ AB_{n-1} & - & B_{n-2} & = & a_{n-1} I_n, \\ & - & B_{n-1} & = & a_n I_n, \end{array} \qquad \begin{array}{rcrcr} AB_0 & & & = & a_0 I_n \\ A^2B_1 & - & AB_0 & = & a_1 A \\ A^3B_2 & - & A^2B_1 & = & a_2 A^2 \\ & & & \cdots & \\ A^nB_{n-1} & - & A^{n-1}B_{n-2} & = & a_{n-1}A^{n-1} \\ & & -A^nB_{n-1} & = & a_n A^n \end{array}$$

Now summing all of the equations in the right hand column gives

$$0 = a_0 A^0 + a_1 A + \ldots + a_{n-1} A^{n-1} + a_n A^n$$

(remember $A^0 = I_n$), which says exactly that $c_A(A) = 0$. □

By the correspondence between linear maps and matrices, we also have $c_A(T) = 0$.

2.3 The minimal polynomial

We start this section with a brief general discussion of polynomials in a single variable x with coefficients in a field K, such as $p = p(x) = 2x^2 - 3x + 11$. The set of all such polynomials is denoted by $K[x]$. There are two *binary operations* on this set: addition and multiplication of polynomials. These operations turn $K[x]$ into *a ring*, which will be studied in great detail in *Algebra-II*.

As a ring $K[x]$ has a number of properties in common[3] with the integers $\mathbb{Z}$. The notation $a|b$ mean a divides b. It can be applied to integers (for instance, $3|12$), and also to polynomials (for instance, $(x-3)|(x^2-4x+3)$).

We can divide one polynomial p (with $p \neq 0$) into another polynomial q and get a remainder with degree less than p. For example, if $q = x^5 - 3$, $p = x^2 + x + 1$, then we find $q = sp + r$ with $s = x^3 - x^2 + 1$ and $r = -x - 4$. For both $\mathbb{Z}$ and $K[x]$, this is known as the *Euclidean Algorithm*.

A polynomial r is said to be a *greatest common divisor* of $p, q \in K[x]$ if $r|p$, $r|q$, and, for any polynomial r' with $r'|p$, $r'|q$, we have $r'|r$. Any two polynomials $p, q \in K[x]$ have a greatest common divisor and a least common multiple (which is defined similarly), but these are only determined up to multiplication by a constant. For example, $x - 1$ is a greatest common divisor of $x^2 - 2x + 1$ and $x^2 - 3x + 2$, but so is $1 - x$ and $2x - 2$. To resolve this ambiguity, we make the following definition.

Definition. A polynomial with coefficients in a field K is called *monic* if the coefficient of the highest power of x is 1.

For example, $x^3 - 2x^2 + x + 11$ is monic, but $2x^2 - x - 1$ is not.

Now we can define $\gcd(p, q)$ to be the unique monic greatest common divisor of p and q, and similarly for $\operatorname{lcm}(p, q)$.

As with the integers, we can use the Euclidean Algorithm to compute $\gcd(p, q)$. For example, if $p = x^4 - 3x^3 + 2x^2$, $q = x^3 - 2x^2 - x + 2$, then $p = q(x-1) + r$ with $r = x^2 - 3x + 2$, and $q = r(x+1)$, so $\gcd(p, q) = r$.

Theorem 2.5 *Let A be an $n \times n$ matrix over K representing the linear operator $T : V \to V$. The following statements hold:*

(i) there is a unique monic non-zero polynomial $p(x)$ with minimal degree and coefficients in K such that $p(A) = 0$,

[3]Technically speaking, they are both *Euclidean Domains* that is an important topic in Algebra-II.

(ii) if $q(x)$ is any polynomial with $q(A) = 0$, then $p|q$.

PROOF: (i) If we have any polynomial $p(x)$ with $p(A) = 0$, then we can make p monic by multiplying it by a constant. By Theorem 2.4, there exists such a $p(x)$, namely $c_A(x)$. If we had two distinct monic polynomials $p_1(x)$, $p_2(x)$ of the same minimal degree with $p_1(A) = p_2(A) = 0$, then $p = p_1 - p_2$ would be a non-zero polynomial of smaller degree with $p(A) = 0$, contradicting the minimality of the degree, so p is unique.

(ii) Let $p(x)$ be the minimal monic polynomial in (i) and suppose that $q(A) = 0$. As we saw above, we can write $q = sp + r$ where r has smaller degree than p. If r is non-zero, then $r(A) = q(A) - s(A)p(A) = 0$ contradicting the minimality of p, so $r = 0$ and $p|q$. □

Definition. The unique monic polynomial $\mu_A(x)$ of minimal degree with $\mu_A(A) = 0$ is called the *minimal polynomial* of A or of the corresponding linear operator T. (Note that $p(A) = 0 \iff p(T) = 0$ for $p \in K[x]$.)

By Theorem 2.4 and Theorem 2.5 (ii), we have:

Corollary 2.6 *The minimal polynomial of a square matrix A divides its characteristic polynomial.*

Similar matrices A and B represent the same linear operator T, and so their minimal polynomial is the same as that of T. Hence we have

Proposition 2.7 *Similar matrices have the same minimal polynomial.*

For a vector $\mathbf{v} \in V$, we can also define *a relative minimal polynomial* $\mu_{A,\mathbf{v}}$ as the unique monic polynomial p of minimal degree for which $p(T)(\mathbf{v}) = \mathbf{0}_V$. Since $p(T) = 0$ if and only if $p(T)(\mathbf{v}) = \mathbf{0}_V$ for all $\mathbf{v} \in V$, μ_A is the least common multiple of the polynomials $\mu_{A,\mathbf{v}}$ for all $\mathbf{v} \in V$.

But $p(T)(\mathbf{v}) = \mathbf{0}_V$ for all $\mathbf{v} \in V$ if and only if $p(T)(\mathbf{b}_i) = \mathbf{0}_V$ for all $\mathbf{b}_i$ in a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of V (exercise), so μ_A is the least common multiple of the polynomials $\mu_{A,\mathbf{b}_i}$.

This gives a method of calculating μ_A. For any $\mathbf{v} \in V$, we can compute $\mu_{A,\mathbf{v}}$ by calculating the sequence of vectors $\mathbf{v}$, $T(\mathbf{v})$, $T^2(\mathbf{v})$, $T^3(\mathbf{v})$ and stopping when it becomes linearly dependent. In practice, we compute $T(\mathbf{v})$ etc. as $A\mathbf{v}$ for the corresponding column vector $\mathbf{v} \in K^{n,1}$.

For example, let $K = \mathbb{R}$ and

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

Using the standard basis $\mathbf{b}_1 = (1\ 0\ 0\ 0)^{\mathrm{T}}$, $\mathbf{b}_2 = (0\ 1\ 0\ 0)^{\mathrm{T}}$, $\mathbf{b}_1 = (0\ 0\ 1\ 0)^{\mathrm{T}}$, $\mathbf{b}_4 = (0\ 0\ 0\ 1)^{\mathrm{T}}$ of $\mathbb{R}^{4,1}$, we have:

$A\mathbf{b}_1 = (3\ 1\ 0\ 0)^{\mathrm{T}}$, $A^2\mathbf{b}_1 = A(A\mathbf{b}_1) = (8\ 4\ 0\ 0)^{\mathrm{T}} = 4A\mathbf{b}_1 - 4\mathbf{b}_1$, so $(A^2 - 4A + 4)\mathbf{b}_1 = 0$, and hence $\mu_{A,\mathbf{b}_1} = x^2 - 4x + 4 = (x-2)^2$.

$A\mathbf{b}_2 = (-1\ 1\ 0\ 0)^{\mathrm{T}}$, $A^2\mathbf{b}_2 = (-4\ 0\ 0\ 0)^{\mathrm{T}} = 4A\mathbf{b}_2 - 4\mathbf{b}_2$, so $\mu_{A,\mathbf{b}_2} = x^2 - 4x + 4$.

$A\mathbf{b}_3 = \mathbf{b}_3$, so $\mu_{A,\mathbf{b}_3} = x - 1$.

$A\mathbf{b}_4 = (1\ 1\ 0\ 1)^T$, $A^2\mathbf{b}_4 = (3\ 3\ 0\ 1)^{\mathrm{T}} = 3A\mathbf{b}_4 - 2\mathbf{b}_4$, so $\mu_{A,\mathbf{b}_4} = x^2 - 3x + 2 = (x-2)(x-1)$.

So we have $\mu_A = \mathrm{lcm}(\mu_{A,\mathbf{b}_1}, \mu_{A,\mathbf{b}_2}, \mu_{A,\mathbf{b}_3}, \mu_{A,\mathbf{b}_4}) = (x-2)^2(x-1)$.

2.4 Jordan chains and Jordan blocks

The Cayley-Hamilton theorem and the theory of minimal polynomials are valid for any matrix over an arbitrary field K, but the theory of Jordan forms will require an additional assumption that the characteristic polynomial $c_A(x)$ is *split* in $K[x]$, i.e. it factorises into linear factors. If the field $K = \mathbb{C}$ then all polynomials in $K[x]$ factorise into linear factors by the Fundamental Theorem of Algebra and JCF works for any matrix.

Definition. A Jordan chain of length k is a sequence of non-zero vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in K^{n,1}$ that satisfies

$$A\mathbf{v}_1 = \lambda\mathbf{v}_1, \quad A\mathbf{v}_i = \lambda\mathbf{v}_i + \mathbf{v}_{i-1}, \quad 2 \le i \le k,$$

for some eigenvalue λ of A.

Equivalently, $(A - \lambda I_n)\mathbf{v}_1 = \mathbf{0}$ and $(A - \lambda I_n)\mathbf{v}_i = \mathbf{v}_{i-1}$ for $2 \le i \le k$, so $(A - \lambda I_n)^i\mathbf{v}_i = \mathbf{0}$ for $1 \le i \le k$.

It is instructive to keep in mind the following model of a Jordan chain that works over complex or real field. Let V be the vector space of functions in the form $f(z)e^{\lambda z}$ where $f(z)$ is the polynomial of degree less than k. Consider the derivative, that is, the linear operator $T : V \to V$ given by $T(\phi(z)) = \phi'(z)$. Vectors $\mathbf{v}_i = z^{i-1}e^{\lambda z}/(i-1)!$ form the Jordan chain for T and the basis of V. In particular, the matrix of T in this basis is the Jordan block defined below.

Definition. A non-zero vector $\mathbf{v} \in V$ such that $(A - \lambda I_n)^i\mathbf{v} = \mathbf{0}$ for some $i > 0$ is called a *generalised eigenvector* of A with respect to the eigenvalue λ.

Note that, for fixed $i > 0$, $\{\, \mathbf{v} \in V \mid (A - \lambda I_n)^i\mathbf{v} = \mathbf{0} \,\}$ is the nullspace of $(A - \lambda I_n)^i$, and is called the *generalised eigenspace* of index i of A with respect to λ. When $i = 1$, this is the ordinary eigenspace of A with respect to λ.

Notice that $\mathbf{v} \in V$ is an eigenvector with eigenvalue λ if and only if $\mu_{A,\mathbf{v}} = x - \lambda$. Similarly, generalised eigenvectors are characterised by the property $\mu_{A,\mathbf{v}} = (x - \lambda)^i$.

For example, consider the matrix

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

We see that, for the standard basis of $K^{3,1}$, we have $A\mathbf{b}_1 = 3\mathbf{b}_1$, $A\mathbf{b}_2 = 3\mathbf{b}_2 + \mathbf{b}_1$, $A\mathbf{b}_3 = 3\mathbf{b}_3 + \mathbf{b}_2$, so $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ is a Jordan chain of length 3 for the eigenvalue 3 of A. The generalised eigenspaces of index 1, 2, and 3 are respectively $\langle \mathbf{b}_1 \rangle$, $\langle \mathbf{b}_1, \mathbf{b}_2 \rangle$, and $\langle \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3 \rangle$.

Notice that the dimension of a generalised eigenspace of A is the nullity of $(T - \lambda I_V)^i$, which is a a function of the linear operator T associated with A. Since similar matrices represent the same linear operator, we have

Proposition 2.8 *The dimensions of corresponding generalised eigenspaces of similar matrices are the same.*

Definition. We define a *Jordan block* with eigenvalue λ of degree k to be a $k \times k$ matrix $J_{\lambda,k} = (\gamma_{ij})$, such that $\gamma_{ii} = \lambda$ for $1 \le i \le k$, $\gamma_{i,i+1} = 1$ for $1 \le i < k$, and $\gamma_{ij} = 0$ if j is not equal to i or $i + 1$. So, for example,

$$J_{1,2} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad J_{\lambda,3} = \begin{pmatrix} \frac{3-i}{2} & 1 & 0 \\ 0 & \frac{3-i}{2} & 1 \\ 0 & 0 & \frac{3-i}{2} \end{pmatrix}, \qquad \text{and} \qquad J_{0,4} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

are Jordan blocks, where $\lambda = \frac{3-i}{2}$ in the second example.

It should be clear that the matrix of T with respect to the basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ of $K^{n,1}$ is a Jordan block of degree n if and only if $\mathbf{v}_1, \dots, \mathbf{v}_n$ is a Jordan chain for A.

Note also that for $A = J_{\lambda,k}$, $\mu_{A,\mathbf{v}_i} = (x-\lambda)^i$, so $\mu_A = (x-\lambda)^k$. Since $J_{\lambda,k}$ is an upper triangular matrix with entries λ on the diagonal, we see that the characteristic polynomial c_A of A is also equal to $(\lambda - x)^k$.

Warning: Some authors put the 1's below rather than above the main diagonal in a Jordan block. This corresponds to either writing the Jordan chain in the reversed order or using rows instead of columns for the standard vector space. However, if an author does both (uses rows and reverses the order) then 1's will go back above the diagonal.

2.5 Jordan bases and the Jordan canonical form

Definition. A *Jordan basis* for A is a basis of $K^{n,1}$ which is a disjoint union of Jordan chains.

We denote the $m \times n$ matrix in which all entries are 0 by $\mathbf{0}_{m,n}$. If A is an $m \times m$ matrix and B an $n \times n$ matrix, then we denote the $(m+n) \times (m+n)$ matrix with block form

$$\left(\begin{array}{c|c} A & \mathbf{0}_{m,n} \\ \hline \mathbf{0}_{n,m} & B \end{array}\right),$$

by $A \oplus B$. For example

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

So, if

$$w_{1,1}, \dots, w_{1,k_1}, w_{2,1}, \dots, w_{2,k_2}, \dots, w_{s,1}, \dots, w_{s,k_s}$$

is a Jordan basis for A in which $w_{i,1}, \dots, w_{i,k_i}$ is a Jordan chain for the eigenvalue λ_i for $1 \le i \le s$, then the matrix of T with respect to this basis is the direct sum $J_{\lambda_1,k_1} \oplus J_{\lambda_2,k_2} \oplus \cdots \oplus J_{\lambda_s,k_s}$ of the corresponding Jordan blocks.

We can now state the main theorem of this section, which says that Jordan bases exist.

Theorem 2.9 *Let A be an $n \times n$ matrix over K such that $c_A(x)$ splits into linear factors in $K[x]$. Then there exists a Jordan basis for A, and hence A is similar to a matrix J which is a direct sum of Jordan blocks. The Jordan blocks occurring in J are uniquely determined by A.*

The matrix J in the theorem is said to be the *Jordan canonical form* (JCF) or sometimes Jordan normal form of A. It is uniquely determined by A up to the order of the blocks.

We will prove the theorem later. First we derive some consequences and study methods for calculating the JCF of a matrix. As we have discussed before, polynomials over $\mathbb{C}$ always split. The gives the following corollary.

Corollary 2.10 *Let A be an $n \times n$ matrix over $\mathbb{C}$. Then there exists a Jordan basis for A.*

The proof of the following corollary requires algebraic techniques beyond the scope of this course. You can try to prove yourself after you have done *Algebra-II*[4]. The trick is to find

[4] Or you can take *Galois Theory* next year and this should become obvious.

a field extension $F \geq K$ such that $c_A(x)$ splits in $F[x]$. For example, consider the rotation by 90 degrees matrix $A = \left(\begin{array}{c|c} 0 & 1 \\ -1 & 0 \end{array}\right)$. Since $c_A(x) = x^2 + 1$, its eigenvalues are imaginary numbers i and $-i$. Hence, it admits no JCF over $\mathbb{R}$ but over complex numbers it has JCF $\left(\begin{array}{c|c} i & 0 \\ 0 & -i \end{array}\right)$.

Corollary 2.11 *Let A be an $n \times n$ matrix over K. Then there exists a field extension $F \geq K$ and a Jordan basis for A in $F^{n,1}$.*

The next two corollaries are immediate[5] consequences of Theorem 2.9 but they are worth stating because of their computational significance. The first one needs Theorem 1.2 as well.

Corollary 2.12 *Let A be an $n \times n$ matrix over K that admits a Jordan basis. If P is the matrix having a Jordan basis as columns, then $P^{-1}AP$ is the JCF of A.*

Notice that a Jordan basis is not, in general, unique. Thus, there exists multiple matrices P such that $J = P^{-1}AP$ is the JCF of A. Suppose now that the eigenvalues of A are $\lambda_1, \dots, \lambda_t$, and that the Jordan blocks in J for the eigenvalue λ_i are $J_{\lambda_i,k_{i,1}}, \dots, J_{\lambda_i,k_{i,j_i}}$, where $k_{i,1} \geq k_{i,2} \geq \cdots \geq k_{i,j_i}$. The final corollary follows from an explicit calculation[6] for J because both minimal and characteristic polynomials of J and A are the same.

Corollary 2.13 *The characteristic polynomial $c_A(x) = \prod_{i=1}^{t}(\lambda_i - x)^{k_i}$, where $k_i = k_{i,1} + \cdots + k_{i,j_i}$ for $1 \leq i \leq t$. The minimal polynomial $\mu_A(x) = \prod_{i=1}^{t}(x - \lambda_i)^{k_{i,1}}$.*

2.6 The JCF when $n = 2$ and 3

When $n = 2$ and $n = 3$, the JCF can be deduced just from the minimal and characteristic polynomials. Let us consider these cases.

When $n = 2$, we have either two distinct eigenvalues λ_1, λ_2, or a single repeated eigenvalue λ_1. If the eigenvalues are distinct, then by Corollary 2.3 A is diagonalisable and the JCF is the diagonal matrix $J_{\lambda_1,1} \oplus J_{\lambda_2,1}$.

Example 1. $A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$. We calculate $c_A(x) = x^2 - 2x - 3 = (x-3)(x+1)$, so there are two distinct eigenvalues, 3 and -1. Associated eigenvectors are $(2\ 1)^{\mathrm{T}}$ and $(-2\ 1)^{\mathrm{T}}$, so we put $P = \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix}$ and then $P^{-1}AP = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$.

If the eigenvalues are equal, then there are two possible JCF's, $J_{\lambda_1,1} \oplus J_{\lambda_1,1}$, which is a scalar matrix, and $J_{\lambda_1,2}$. The minimal polynomial is respectively $(x - \lambda_1)$ and $(x - \lambda_1)^2$ in these two cases. In fact, these cases can be distinguished without any calculation whatsoever, because in the first case $A = PJP^{-1} = J$ so A is its own JCF.

In the second case, a Jordan basis consists of a single Jordan chain of length 2. To find such a chain, let $\mathbf{v}_2$ be any vector for which $(A - \lambda_1 I_2)\mathbf{v}_2 \neq \mathbf{0}$ and let $\mathbf{v}_1 = (A - \lambda_1 I_2)\mathbf{v}_2$. (In practice, it is often easier to find the vectors in a Jordan chain in reverse order.)

Example 2. $A = \begin{pmatrix} 1 & 4 \\ -1 & -3 \end{pmatrix}$. We have $c_A(x) = x^2 + 2x + 1 = (x+1)^2$, so there is a single

[5] This means I am not proving them here but I expect you to be able to prove them

[6] The characteristic polynomial of J is the product of the characteristic polynomials of the Jordan blocks and the minimal polynomial of J is the least common multiple of characteristic polynomials of the Jordan blocks

eigenvalue -1 with multiplicity 2. Since the first column of $A+I_2$ is non-zero, we can choose $\mathbf{v}_2 = (1\ 0)^{\mathrm{T}}$ and $\mathbf{v}_1 = (A+I_2)\mathbf{v}_2 = (2\ \ -1)^{\mathrm{T}}$, so $P = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$ and $P^{-1}AP = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$.

Now let $n = 3$. If there are three distinct eigenvalues, then A is diagonalisable.

Suppose that there are two distinct eigenvalues, so one has multiplicity 2, and the other has multiplicity 1. Let the eigenvalues be λ_1, λ_1, λ_2, with $\lambda_1 \neq \lambda_2$. Then there are two possible JCF's for A, $J_{\lambda_1,1} \oplus J_{\lambda_1,1} \oplus J_{\lambda_2,1}$ and $J_{\lambda_1,2} \oplus J_{\lambda_2,1}$, and the minimal polynomial is $(x-\lambda_1)(x-\lambda_2)$ in the first case and $(x-\lambda_1)^2(x-\lambda_2)$ in the second.

In the first case, a Jordan basis is a union of three Jordan chains of length 1, each of which consists of an eigenvector of A.

Example 3. $A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 5 & 2 \\ -2 & -6 & -2 \end{pmatrix}$. Then

$$c_A(x) = (2-x)[(5-x)(-2-x)+12] = (2-x)(x^2-3x+2) = (2-x)^2(1-x).$$

We know from the theory above that the minimal polynomial must be $(x-2)(x-1)$ or $(x-2)^2(x-1)$. We can decide which simply by calculating $(A-2I_3)(A-I_3)$ to test whether or not it is 0. We have

$$A - 2I_3 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 3 & 2 \\ -2 & -6 & -4 \end{pmatrix}, \quad A - I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 4 & 2 \\ -2 & -6 & -3 \end{pmatrix},$$

and the product of these two matrices is 0, so $\mu_A = (x-2)(x-1)$.

The eigenvectors $\mathbf{v}$ for $\lambda_1 = 2$ satisfy $(A-2I_3)\mathbf{v} = \mathbf{0}$, and we must find two linearly independent solutions; for example we can take $\mathbf{v}_1 = (0\ 2\ \ -3)^{\mathrm{T}}$, $\mathbf{v}_2 = (1\ \ -1\ 1)^{\mathrm{T}}$. An eigenvector for the eigenvalue 1 is $\mathbf{v}_3 = (0\ 1\ \ -2)^{\mathrm{T}}$, so we can choose

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 2 & -1 & 1 \\ -3 & 1 & -2 \end{pmatrix}$$

and then $P^{-1}AP$ is diagonal with entries $2, 2, 1$.

In the second case, there are two Jordan chains, one for λ_1 of length 2, and one for λ_2 of length 1. For the first chain, we need to find a vector $\mathbf{v}_2$ with $(A-\lambda_1 I_3)^2\mathbf{v}_2 = \mathbf{0}$ but $(A-\lambda_1 I_3)\mathbf{v}_2 \neq \mathbf{0}$, and then the chain is $\mathbf{v}_1 = (A-\lambda_1 I_3)\mathbf{v}_2, \mathbf{v}_2$. For the second chain, we simply need an eigenvector for λ_2.

Example 4. $A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 3 & 1 \\ -1 & -4 & -1 \end{pmatrix}$. Then

$$c_A(x) = (3-x)[(3-x)(-1-x)+4] - 2 + (3-x) = -x^3 + 5x^2 - 8x + 4 = (2-x)^2(1-x),$$

as in Example 3. We have

$$A-2I_3 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & -4 & -3 \end{pmatrix}, \quad (A-2I_3)^2 = \begin{pmatrix} 0 & 0 & 0 \\ -1 & -3 & -2 \\ 2 & 6 & 4 \end{pmatrix}, \quad (A-I_3) = \begin{pmatrix} 2 & 2 & 1 \\ 0 & 2 & 1 \\ -1 & -4 & -2 \end{pmatrix}.$$

and we can check that $(A-2I_3)(A-I_3)$ is non-zero, so we must have $\mu_A = (x-2)^2(x-1)$.

For the Jordan chain of length 2, we need a vector with $(A-2I_3)^2\mathbf{v}_2 = \mathbf{0}$ but $(A-2I_3)\mathbf{v}_2 \neq \mathbf{0}$, and we can choose $\mathbf{v}_2 = (2\ 0\ \ {-1})^{\mathrm{T}}$. Then $\mathbf{v}_1 = (A-2I_3)\mathbf{v}_2 = (1\ \ {-1}\ 1)^{\mathrm{T}}$. An eigenvector for the eigenvalue 1 is $\mathbf{v}_3 = (0\ 1\ \ {-2})^{\mathrm{T}}$, so we can choose

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

and then

$$P^{-1}AP = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Finally, suppose that there is a single eigenvalue, λ_1, so $c_A = (\lambda_1 - x)^3$. There are three possible JCF's for A, $J_{\lambda_1,1} \oplus J_{\lambda_1,1} \oplus J_{\lambda_1,1}$, $J_{\lambda_1,2} \oplus J_{\lambda_1,1}$, and $J_{\lambda_1,3}$, and the minimal polynomials in the three cases are $(x-\lambda_1)$, $(x-\lambda_1)^2$, and $(x-\lambda_1)^3$, respectively.

In the first case, J is a scalar matrix, and $A = PJP^{-1} = J$, so this is recognisable immediately.

In the second case, there are two Jordan chains, one of length 2 and one of length 1. For the first, we choose $\mathbf{v}_2$ with $(A-\lambda_1 I_3)\mathbf{v}_2 \neq 0$, and let $\mathbf{v}_1 = (A-\lambda_1 I_3)\mathbf{v}_2$. (This case is easier than the case illustrated in Example 4, because we have $(A-\lambda_1 I_3)^2\mathbf{v} = \mathbf{0}$ for all $\mathbf{v} \in \mathbb{C}^{3,1}$.) For the second Jordan chain, we choose $\mathbf{v}_3$ to be an eigenvector for λ_1 such that $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly independent.

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

$$c_A(x) = -x[(3+x)x+2] - 2(x+1) - 2 + (3+x) = -x^3 - 3x^2 - 3x - 1 = -(1+x)^3.$$

We have

$$A + I_3 = \begin{pmatrix} 1 & 2 & 1 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix},$$

and we can check that $(A+I_3)^2 = 0$. The first column of $A+I_3$ is non-zero, so $(A+I_3)(1\ 0\ 0)^{\mathrm{T}} \neq \mathbf{0}$, and we can choose $\mathbf{v}_2 = (1\ 0\ 0)^{\mathrm{T}}$ and $\mathbf{v}_1 = (A+I_3)\mathbf{v}_2 = (1\ \ {-1}\ 1)^{\mathrm{T}}$. For $\mathbf{v}_3$ we need to choose a vector which is not a multiple of $\mathbf{v}_1$ such that $(A+I_3)\mathbf{v}_3 = \mathbf{0}$, and we can choose $\mathbf{v}_3 = (0\ 1\ \ {-2})^{\mathrm{T}}$. So we have

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

and then

$$P^{-1}AP = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

In the third case, there is a single Jordan chain, and we choose $\mathbf{v}_3$ such that $(A-\lambda_1 I_3)^2\mathbf{v}_3 \neq 0$, $\mathbf{v}_2 = (A-\lambda_1 I_3)\mathbf{v}_3$, $\mathbf{v}_1 = (A-\lambda_1 I_3)^2\mathbf{v}_3$.

Example 6. $A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix}$. Then

$$c_A(x) = -x[(2+x)(1+x)] - (2+x) + 1 = -(1+x)^3.$$

We have

$$A + I_3 = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix}, \quad (A + I_3)^2 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 1 & 1 \end{pmatrix},$$

so $(A + I_3)^2 \neq 0$ and $\mu_A = (x+1)^3$. For $\mathbf{v}_3$, we need a vector that is not in the nullspace of $(A+I_3)^2$. Since the second column, which is the image of $(0\ 1\ 0)^{\mathrm{T}}$ is non-zero, we can choose $\mathbf{v}_3 = (0\ 1\ 0)^{\mathrm{T}}$, and then $\mathbf{v}_2 = (A + I_3)\mathbf{v}_3 = (1\ 0\ 0)^{\mathrm{T}}$ and $\mathbf{v}_1 = (A + I_3)\mathbf{v}_2 = (1\ \ {-1}\ 1)^{\mathrm{T}}$. So we have

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

and then

$$P^{-1}AP = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

2.7 The general case

For dimensions higher than 3, we cannot always determine the JCF just from the characteristic and minimal polynomials. For example, when $n = 4$, the JCF's $J_{\lambda,2} \oplus J_{\lambda,2}$ and $J_{\lambda,2} \oplus J_{\lambda,1} \oplus J_{\lambda,1}$ both have $c_A = (\lambda - x)^4$ and $\mu_A = (x - \lambda)^2$.

In general, we can compute the JCF from the dimensions of the generalised eigenspaces.

Let $J_{\lambda,k}$ be a Jordan block and let $A = J_{\lambda,k} - \lambda I_k$. Then we calculate that, for $1 \leq i < k$, A^i has $(k - i)$ 1's on the i-th diagonal upwards from the main diagonal, and $A^k = 0$. For example, when $k = 4$,

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^4 = 0.$$

(A matrix A for which $A^k = 0$ for some $k > 0$ is called *nilpotent*.)

It should be clear from this that, for $1 \leq i \leq k$, $\mathrm{rank}(A^i) = k - i$, so $\mathrm{nullity}(A^i) = i$, and for $i \geq k$, $\mathrm{rank}(A^i) = 0$, $\mathrm{nullity}(A^i) = k$.

On the other hand, if $\mu \neq \lambda$ and $A = J_{\lambda,k} - \mu I_k$, then, for any integer i, A^i is an upper triangular matrix with non-zero entries $(\lambda - \mu)^i$ on the diagonal, and so $\mathrm{rank}(A^i) = k$, $\mathrm{nullity}(A^i) = 0$.

It is easy to see that, for square matrices A and B, $\mathrm{rank}(A \oplus B) = \mathrm{rank}(A) + \mathrm{rank}(B)$ and $\mathrm{nullity}(A \oplus B) = \mathrm{nullity}(A) + \mathrm{nullity}(B)$. So, for a matrix J in JCF, we can determine the sizes of the Jordan blocks for an eigenvalue λ of J from a knowledge of the nullities of the matrices $(J - \lambda)^i$ for $i > 0$.

For example, suppose that $J = J_{-2,3} \oplus J_{-2,3} \oplus J_{-2,1} \oplus J_{1,2}$. Then $\mathrm{nullity}(J + 2I_9) = 3$, $\mathrm{nullity}(J + 2I_9)^2 = 5$, $\mathrm{nullity}(J + 2I_9)^i = 7$ for $i \geq 3$, $\mathrm{nullity}(J - I_9) = 1$, $\mathrm{nullity}(J - I_9)^i = 2$ for $i \geq 2$.

First observe that the total number of Jordan blocks with eigenvalue λ is equal to $\mathrm{nullity}(J - \lambda I_n)$.

More generally, the number of Jordan blocks $J_{\lambda,j}$ for λ with $j \geq i$ is equal to $\mathrm{nullity}((J - \lambda I_n)^i) - \mathrm{nullity}((J - \lambda I_n)^{i-1})$.

The nullspace of $(J - \lambda I_n)^i$ was defined earlier to be the generalised eigenspace of index i of J with respect to the eigenvalue λ. If J is the JCF of a matrix A, then A and J are similar matrices, so it follows from Proposition 2.8 that $\mathrm{nullity}((J - \lambda I_n)^i) = \mathrm{nullity}((A - \lambda I_n)^i)$.

So, summing up, we have:

Theorem 2.14 *Let λ be an eigenvalue of a matrix A and let J be the JCF of A. The following hold.*

(i) The number of Jordan blocks of J with eigenvalue λ is equal to $\mathrm{nullity}(A - \lambda I_n)$.

(ii) More generally, for $i > 0$, the number of Jordan blocks of J with eigenvalue λ and degree at least i is equal to $\mathrm{nullity}((A - \lambda I_n)^i) - \mathrm{nullity}((A - \lambda I_n)^{i-1})$.

Note that this proves the uniqueness part of Theorem 2.9.

2.8 Examples

Example 7. $A = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 1 & 0 & -2 & -2 \end{pmatrix}$. Then $c_A(x) = (-2 - x)^4$, so there is a single eigenvalue -2 with multiplicity 4. We find $(A+2I_4) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & -2 & 0 \end{pmatrix}$, and $(A+2I_4)^2 = 0$, so $\mu_A = (x+2)^2$, and the JCF of A could be $J_{-2,2} \oplus J_{-2,2}$ or $J_{-2,2} \oplus J_{-2,1} \oplus J_{-2,1}$.

To decide which case holds, we calculate the nullity of $A + 2I_4$ which, by Theorem 2.14, is equal to the number of Jordan blocks with eigenvalue -2. Since $A+2I_4$ has just two non-zero rows, which are distinct, its rank is clearly 2, so its nullity is $4 - 2 = 2$, and hence the JCF of A is $J_{-2,2} \oplus J_{-2,2}$.

A Jordan basis consists of a union of two Jordan chains, which we will call $\mathbf{v}_1, \mathbf{v}_2$, and $\mathbf{v}_3, \mathbf{v}_4$, where $\mathbf{v}_1$ and $\mathbf{v}_3$ are eigenvectors and $\mathbf{v}_2$ and $\mathbf{v}_4$ are generalised eigenvectors of index 2. To find such chains, it is probably easiest to find $\mathbf{v}_2$ and $\mathbf{v}_4$ first and then to calculate $\mathbf{v}_1 = (A + 2I_4)\mathbf{v}_2$ and $\mathbf{v}_3 = (A + 2I_4)\mathbf{v}_4$.

Although it is not hard to find $\mathbf{v}_2$ and $\mathbf{v}_4$ in practice, we have to be careful, because they need to be chosen so that no linear combination of them lies in the nullspace of $(A + 2I_4)$. In fact, since this nullspace is spanned by the second and fourth standard basis vectors, the obvious choice is $\mathbf{v}_2 = (1\ 0\ 0\ 0)^{\mathrm{T}}$, $\mathbf{v}_4 = (0\ 0\ 1\ 0)^{\mathrm{T}}$, and then $\mathbf{v}_1 = (A + 2I_4)\mathbf{v}_2 = (0\ 0\ 0\ 1)^{\mathrm{T}}$, $\mathbf{v}_3 = (A + 2I_4)\mathbf{v}_4 = (0\ 1\ 0\ \ -2)^{\mathrm{T}}$, so to transform A to JCF, we put

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & -2 & 0 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 0 & 2 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad P^{-1}AP = \begin{pmatrix} -2 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

Example 8. $A = \begin{pmatrix} -1 & -3 & -1 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 3 & 1 & -1 \end{pmatrix}$. Then $c_A(x) = (-1 - x)^2(2 - x)^2$, so there are two eigenvalue -1, 2, both with multiplicity 2. There are four possibilities for the JCF (one or two blocks for each of the two eigenvalues). We could determine the JCF by computing the minimal polynomial μ_A but it is probably easier to compute the nullities of the eigenspaces and use Theorem 2.14. We have

$$A+I_4 = \begin{pmatrix} 0 & -3 & -1 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 3 & 1 & 0 \end{pmatrix}, \quad (A-2I_4) = \begin{pmatrix} -3 & -3 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & -3 \end{pmatrix}, \quad (A-2I_4)^2 = \begin{pmatrix} 9 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -9 & 0 & 9 \end{pmatrix}.$$

The rank of $A + I_4$ is clearly 2, so its nullity is also 2, and hence there are two Jordan blocks with eigenvalue -1. The three non-zero rows of $(A - 2I_4)$ are linearly independent, so its rank is 3, hence its nullity 1, so there is just one Jordan block with eigenvalue 2, and the JCF of A is $J_{-1,1} \oplus J_{-1,1} \oplus J_{2,2}$.

For the two Jordan chains of length 1 for eigenvalue -1, we just need two linearly independent eigenvectors, and the obvious choice is $\mathbf{v}_1 = (1\ 0\ 0\ 0)^{\mathrm{T}}$, $\mathbf{v}_2 = (0\ 0\ 0\ 1)^{\mathrm{T}}$. For the Jordan chain $\mathbf{v}_3, \mathbf{v}_4$ for eigenvalue 2, we need to choose $\mathbf{v}_4$ in the nullspace of $(A - 2I_4)^2$ but not in the nullspace of $A - 2I_4$. (This is why we calculated $(A - 2I_4)^2$.) An obvious choice here is $\mathbf{v}_4 = (0\ 0\ 1\ 0)^{\mathrm{T}}$, and then $\mathbf{v}_3 = (-1\ 1\ 0\ 1)^{\mathrm{T}}$, and to transform A to JCF, we put

$$P = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

2.9 Proof of Theorem 2.9 (non-examinable)

We proceed by induction on $n = \dim(V)$. The case $n = 1$ is clear.

Let λ be an eigenvalue of T and let $U = \mathrm{im}(T - \lambda I_V)$ and $m = \dim(U)$. Then $m = \mathrm{rank}(T - \lambda I_V) = n - \mathrm{nullity}(T - \lambda I_V) < n$, because the eigenvectors for λ lie in the nullspace of $T - \lambda I_V$. For $\mathbf{u} \in U$, we have $\mathbf{u} = (T - \lambda I_V)(\mathbf{v})$ for some $\mathbf{v} \in V$, and hence $T(\mathbf{u}) = T(T - \lambda I_V)(\mathbf{v}) = (T - \lambda I_V)T(\mathbf{v}) \in U$. So T restricts to $T_U : U \to U$, and we can apply our inductive hypothesis to T_U to deduce that U has a basis $\mathbf{e}_1, \ldots, \mathbf{e}_m$, which is a disjoint union of Jordan chains for T_U.

We now show how to extend the Jordan basis of U to one of V. We do this in two stages. For the first stage, suppose that l of the Jordan chains of T_U are for the eigenvalue λ (possibly $l = 0$). For each such chain $\mathbf{v}_1, \ldots, \mathbf{v}_k$ with $T(\mathbf{v}_1) = \lambda\mathbf{v}_1,\ \ T(\mathbf{v}_i) = \lambda\mathbf{v}_i + \mathbf{v}_{i-1},\ \ 2 \le i \le k$, since $\mathbf{v}_k \in U = \mathrm{im}(T - \lambda I_V)$, we can find $\mathbf{v}_{k+1} \in V$ with $T(\mathbf{v}_{k+1}) = \lambda\mathbf{v}_{k+1} + \mathbf{v}_k$, thereby extending the chain by an extra vector. So far we have adjoined l new vectors to the basis, by extending the length l Jordan chains by 1. Let us call these new vectors $\mathbf{w}_1, \ldots, \mathbf{w}_l$.

For the second stage, observe that the first vector in each of the l chains lies in the eigenspace of T_U for λ. We know that the dimension of the eigenspace of T for λ is the nullspace of $(T - \lambda I_V)$, which has dimension $n - m$. So we can adjoin $(n - m) - l$ further eigenvectors of T to the l that we have already to complete a basis of the nullspace of $(T - \lambda I_V)$. Let us call these $(n - m) - l$ new vectors $\mathbf{w}_{l+1}, \ldots, \mathbf{w}_{n-m}$. They are adjoined to our basis of V in the second stage. They each form a Jordan chain of length 1, so we now have a collection of n vectors which form a disjoint union of Jordan chains.

To complete the proof, we need to show that these n vectors form a basis of V, for which is it is enough to show that they are linearly independent.

Partly because of notational difficulties, we provide only a sketch proof of this, and leave the details to the student. Suppose that $\alpha_1\mathbf{w}_1 + \cdots + \alpha_{n-m}\mathbf{w}_{n-m} + \mathbf{x} = \mathbf{0}$, where $\mathbf{x}$ is a linear combination of the basis vectors of U. Applying $T - \lambda I_n$ gives

$$\alpha_1(T - \lambda I_n)(\mathbf{w}_1) + \cdots + \alpha_l(T - \lambda I_n)(\mathbf{w}_l) + (T - \lambda I_n)(\mathbf{x}).$$

Each of $\alpha_i(T-\lambda I_n)(\mathbf{w}_i)$ for $1 \le i \le l$ is the last member of one of the l Jordan chains for T_U. When we apply $(T-\lambda I_n)$ to one of the basis vectors of U, we get a linear combination of the basis vectors of U other than $\alpha_i(T-\lambda I_n)(\mathbf{w}_i)$ for $1 \le i \le l$. Hence, by the linear independence of the basis of U, we deduce that $\alpha_i = 0$ for $1 \le i \le l$ This implies that $(T-\lambda I_n)(\mathbf{x}) = \mathbf{0}$, so $\mathbf{x}$ is in the eigenspace of T_U for the eigenvalue λ. But, by construction, $\mathbf{w}_{l+1}, \ldots, \mathbf{w}_{n-m}$ extend a basis of this eigenspace of T_U to that the eigenspace of V, so we also get $\alpha_i = 0$ for $l+1 \le i \le n-m$, which completes the proof.

2.10 Powers of matrices

The theory we developed can be used to compute powers of matrices efficiently. Suppose we need to compute A^{2012} where $A = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 1 & 0 & -2 & -2 \end{pmatrix}$ from section 2.8.

There are two practical ways of computing A^n for a general matrix. The first one involves Jordan forms. If $J = P^{-1}AP$ is the JCF of A then it is sufficient to compute J^n because of the telescoping product:

$$A^n = (PJP^{-1})^n = PJP^{-1}PJP^{-1}P \ldots JP^{-1} = PJ^nP^{-1}.$$

$$\text{If } J = \begin{pmatrix} J_{k_1,\lambda_1} & 0 & \cdots & 0 \\ 0 & J_{k_2,\lambda_2} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & J_{k_t,\lambda_t} \end{pmatrix} \text{ then } J^n = \begin{pmatrix} J^n_{k_1,\lambda_1} & 0 & \cdots & 0 \\ 0 & J^n_{k_2,\lambda_2} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & J^n_{k_t,\lambda_t} \end{pmatrix}.$$

Finally, the power of an individual Jordan block can be computed as

$$J^n_{k,\lambda} = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \cdots & C^n_{k-2}\lambda^{n-k+2} & C^n_{k-1}\lambda^{n-k+1} \\ 0 & \lambda^n & \cdots & C^n_{k-3}\lambda^{n-k+3} & C^n_{k-2}\lambda^{n-k+2} \\ \ddots & \ddots & \ddots & \ddots & \\ 0 & 0 & \cdots & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & \cdots & 0 & \lambda^n \end{pmatrix}$$

where $C^n_t = n!/(n-t)!t!$ is the Choose-function, interpreted as $C^n_t = 0$ whenever $t > n$.

Let us apply it to the matrix from example 7 in 2.8:

$$A^n = PJ^nP^{-1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & -2 & 0 \end{pmatrix} \begin{pmatrix} -2 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix}^n \begin{pmatrix} 0 & 2 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} =$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & -2 & 0 \end{pmatrix} \begin{pmatrix} (-2)^n & n(-2)^{n-1} & 0 & 0 \\ 0 & (-2)^n & 0 & 0 \\ 0 & 0 & (-2)^n & n(-2)^{n-1} \\ 0 & 0 & 0 & (-2)^n \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} =$$

$$\begin{pmatrix} (-2)^n & 0 & 0 & 0 \\ 0 & (-2)^n & n(-2)^{n-1} & 0 \\ 0 & 0 & (-2)^n & 0 \\ n(-2)^{n-1} & 0 & n(-2)^n & (-2)^n \end{pmatrix}.$$

The second method of computing A^n uses Lagrange's interpolation polynomial. It is less labour intensive and more suitable for pen-and-paper calculations. Suppose $\psi(A) = 0$ for a polynomial $\psi(z)$, in practice, $\psi(z)$ is either minimal or characteristic polynomial. Dividing with a remainder $z^n = q(z)\psi(z) + h(z)$, we conclude that

$$A^n = q(A)\psi(A) + h(A) = h(A).$$

Division with a remainder may appear problematic[7] for large n but there is a shortcut. If we know the roots of $\psi(z)$, say $\alpha_1, \ldots, \alpha_k$ with their multiplicities $m_1, \ldots, m_k$, then $h(z)$ can be found by solving the system of simultaneous equations in coefficients of $h(z)$:

$$f^{(t)}(\alpha_j) = h^{(t)}(\alpha_j),\ 1 \leq j \leq k,\ 0 \leq t < m_j$$

where $f(z) = z^n$ and $f^{(t)} = f^{(t-1)\prime}$ is the t-th derivative. In other words, $h(z)$ is Lagrange's interpolation polynomial for the function z^n at the roots of $\psi(z)$.

We know that $\mu_A(z) = (z+2)^2$ for the matrix A above. Suppose the Lagrange interpolation of z^n at the roots of $(z+2)^2$ is $h(z) = \alpha z + \beta$. The condition on the coefficients is given by

$$\begin{cases} (-2)^n & = & h(-2) & = & -2\alpha + \beta \\ n(-2)^{n-1} & = & h'(-2) & = & \alpha \end{cases}$$

Solving them gives $\alpha = n(-2)^{n-1}$ and $\beta = (1-n)(-2)^n$. It follows that

$$A^n = n(-2)^{n-1}A + (1-n)(-2)^n I = \begin{pmatrix} (-2)^n & 0 & 0 & 0 \\ 0 & (-2)^n & n(-2)^{n-1} & 0 \\ 0 & 0 & (-2)^n & 0 \\ n(-2)^{n-1} & 0 & n(-2)^n & (-2)^n \end{pmatrix}.$$

2.11 Applications to difference equations

Let us consider *an initial value problem* for an *autonomous* system with discrete time:

$$x(n+1) = Ax(n),\ n \in \mathbb{N},\ x(0) = w.$$

Here $x(n) \in K^m$ is a sequence of vectors in a vector space over a field K. One thinks of $x(n)$ as a state of the system at time n. The initial state is $x(0) = w$. The $n \times n$-matrix A with coefficients in K describes the evolution of the system. The adjective *autonomous* means that the evolution equation does not change with the time[8].

It takes longer to formulate this problem then to solve it. The solution is a no-brainer:

$$x(n) = Ax(n-1) = A^2 x(n-2) = \cdots = A^n x(0) = A^n w.$$

As a working example, let us consider a 2-step linearly recursive sequence. It is determined by a quadruple $(a, b, c, d) \in K^4$ and the rules

$$s_0 = a,\ s_1 = b,\ s_n = cs_{n-1} + ds_{n-2} \text{ for } n \geq 2.$$

Such sequences are ubiquitous. Arithmetic sequences form a subclass with $c = 2$, $d = -1$. In general, $(a, b, 2, -1)$ determines the arithmetic sequence starting at a with the difference $b - a$. For instance, $(0, 1, 2, -1)$ determines the sequence of natural numbers $s_n = n$.

A geometric sequence starting at a with ratio q admits a non-unique description. One obvious quadruples giving it is $(a, aq, q, 0)$. However, it is conceptually better to use quadruple $(a, aq, 2q, -q^2)$ because the sequences coming from $(a, b, 2q, -q^2)$ include both arithmetic and geometric sequences and can be called *arithmo-geometric sequences*.

If $c = d = 1$ then this is a Fibonacci type sequence. For instance, $(0, 1, 1, 1)$ determines Fibonacci numbers F_n while $(2, 1, 1, 1)$ determines Lucas numbers L_n.

[7] Try to divide z^{2012} by $z^2 + z + 1$ without reading any further.

[8] A nonautonomous system would be described by $x(n+1) = A(n)x(n)$ here.

All of these examples admit *closed*[9] formulae for a generic term s_n. Can we find a closed formula for s_n, in general? Yes, we can because this problem is reduced to an initial value problem with discrete time if we set

$$x(n) = \begin{pmatrix} s_n \\ s_{n+1} \end{pmatrix}, \ w = \begin{pmatrix} a \\ b \end{pmatrix}, \ A = \begin{pmatrix} 0 & 1 \\ d & c \end{pmatrix} .$$

Computing the characteristic polynomial, $c_A(z) = z^2 - cz - d$. If $c^2 + 4d = 0$, the JCF of A is $J = \begin{pmatrix} c/2 & 1 \\ 0 & c/2 \end{pmatrix}$. Let $q = c/2$. Then $d = -q^2$ and we are dealing with the arithmo-geometric sequence $(a, b, 2q, -q^2)$. Let us find the closed formula for s_n in this case using Jordan forms. As $A = \begin{pmatrix} 0 & 1 \\ -q^2 & 2q \end{pmatrix}$ one can choose the Jordan basis $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $e_1 = \begin{pmatrix} 1 \\ q \end{pmatrix}$. If $P = \begin{pmatrix} 1 & 0 \\ q & 1 \end{pmatrix}$ then $P^{-1} = \begin{pmatrix} 1 & 0 \\ -q & 1 \end{pmatrix}$ and

$$A^n = (PJP^{-1})^n = PJ^nP^{-1} = P \begin{pmatrix} q^n & nq^{n-1} \\ 0 & q^n \end{pmatrix} P^{-1} = \begin{pmatrix} (1-n)q^n & nq^{n-1} \\ -nq^{n+1} & (1+n)q^n \end{pmatrix}.$$

This gives the closed formula for arithmo-geometric sequence we were seeking:

$$s_n = (1-n)q^n a + nq^{n-1} b.$$

If $c^2+4d \neq 0$, the JCF of A is $\begin{pmatrix} (c+\sqrt{c^2+4d})/2 & 0 \\ 0 & (c-\sqrt{c^2+4d})/2 \end{pmatrix}$ and the closed formula for s_n will involve the sum of two geometric sequences. Let us see it through for Fibonacci and Lucas numbers using Lagrange's polynomial. Since $c = d = 1$, $c^2 + 4d = 5$ and the roots of $c_A(z)$ are the golden ratio $\lambda = (1+\sqrt{5})/2$ and $1-\lambda = (1-\sqrt{5})/2$. It is useful to observe that $2\lambda - 1 = \sqrt{5}$ and $\lambda(1-\lambda) = -1$. Let us introduce the number $\mu_n = \lambda^n - (1-\lambda)^n$. Suppose the Lagrange interpolation of z^n at the roots of $z^2 - z - 1$ is $h(z) = \alpha z + \beta$. The condition on the coefficients is given by

$$\begin{cases} \lambda^n & = & h(\lambda) & = & \alpha\lambda + \beta \\ (1-\lambda)^n & = & h(1-\lambda) & = & \alpha(1-\lambda) + \beta \end{cases}$$

Solving them gives $\alpha = \mu_n/\sqrt{5}$ and $\beta = \mu_{n-1}/\sqrt{5}$. It follows that

$$A^n = \alpha A + \beta = \mu_n/\sqrt{5}A + \mu_{n-1}/\sqrt{5}I_2 = \begin{pmatrix} \mu_{n-1}/\sqrt{5} & \mu_n/\sqrt{5} \\ \mu_n/\sqrt{5} & (\mu_n + \mu_{n-1})/\sqrt{5} \end{pmatrix} .$$

Since $\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = A^n \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, it immediately implies that

$$A^n = \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix} \text{ and } F_n = \mu_n/\sqrt{5} .$$

Similarly for the Lucas numbers, we get $\begin{pmatrix} L_n \\ L_{n+1} \end{pmatrix} = A^n \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and

$$L_n = 2F_{n-1} + F_n = F_{n-1} + F_{n+1} = (\mu_{n-1} + \mu_{n+1})/\sqrt{5}.$$

[9] Closed means non-recursive, for instance, $s_n = a + n(b-a)$ for the arithmetic sequence

2.12 Functions of matrices

We restrict to $K = \mathbb{R}$ in this section. Let us consider a power series $\sum_n a_n z^n$, $a_n \in \mathbb{R}$ with a positive radius of convergence ε. It defines a function $f : (-\varepsilon, \varepsilon) \to \mathbb{R}$ by

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

and the power series is Taylor's series of $f(x)$ at zero. In particular,

$$a_n = f^{[n]}(0) = f^{(n)}(0)/n!,$$

where $f^{[n]}(z) = f^{(n)}(z)/n!$ is a divided derivative. We extend the function $f(z)$ to matrices by the formula

$$f(A) = \sum_{n=0}^{\infty} f^{[n]}(0) A^n \ .$$

The right hand side of this formula is a matrix whose entries are series. All these series need to converge for $f(A)$ to be well defined. If the norm[10] of A is less than ε then $f(A)$ is well defined. Alternatively, if all eigenvalues of A belong to $(-\varepsilon, \varepsilon)$ then $f(A)$ is well defined as can be seen from the JCF method of computing $f(A)$. If

$$J = \begin{pmatrix} J_{k_1,\lambda_1} & 0 & \cdots & 0 \\ 0 & J_{k_2,\lambda_2} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & J_{k_t,\lambda_t} \end{pmatrix} = P^{-1}AP$$

is the JCF of A then

$$f(A) = Pf(J)P^{-1} = P \begin{pmatrix} f(J_{k_1,\lambda_1}) & 0 & \cdots & 0 \\ 0 & f(J_{k_2,\lambda_2}) & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & f(J_{k_t,\lambda_t}) \end{pmatrix}$$

while

$$f(J_{k,\lambda}) = \begin{pmatrix} f(\lambda) & f^{[1]}(\lambda) & \cdots & f^{[k-1]}(\lambda) \\ 0 & f(\lambda) & \cdots & f^{[k-2]}(\lambda) \\ & & \ddots & \\ 0 & 0 & \cdots & f(\lambda) \end{pmatrix} .$$

Lagrange's method of computing $f(A)$ works as well despite the fact that there is no sensible way to divide with a remainder in analytic functions. For instance,

$$e^z = \frac{e^z}{z^2+1}\psi(z) + 0 = \frac{e^z - 1}{z^2+1}\psi(z) + 1 = \frac{e^z - z}{z^2+1}\psi(z) + z$$

for $\psi(z) = z^2 + 1$. Thus, there are infinitely many ways to divide with a remainder as $f(z) = q(z)\psi(z) + h(z)$. The point is that $f(A) = h(A)$ only if $q(A)$ is well defined. Notice that the naive expression $q(A) = (f(A) - h(A))\psi(A)^{-1}$ involves division by zero. However, if $h(z)$ is the interpolation polynomial then $q(A)$ is well defined and the calculation $f(A) = q(A)\psi(A) + h(A) = h(A)$ carries through.

Let us compute e^A for the matrix A from example 7, section 2.8. Recall that Taylor's series for exponent $e^x = \sum_{n=0}^{\infty} x^n/n!$ converges for all x. Consequently the matrix exponent $e^A = \sum_{n=0}^{\infty} A^n/n!$ is defined for all real m-by-m matrices.

[10]this notion is beyond the scope of this module and will be discussed in *Differentiation*

Suppose the Lagrange interpolation of e^z at the roots of $\mu_A(z) = (z+2)^2$ is $h(z) = \alpha z + \beta$. The condition on the coefficients is given by

$$\begin{cases} e^{-2} &= h(-2) &= -2\alpha + \beta \\ e^{-2} &= h'(-2) &= \alpha \end{cases}$$

Solving them gives $\alpha = e^{-2}$ and $\beta = 3e^{-2}$. It follows that

$$e^A = e^{-2}A + 3e^{-2}I = \begin{pmatrix} e^{-2} & 0 & 0 & 0 \\ 0 & e^{-2} & e^{-2} & 0 \\ 0 & 0 & e^{-2} & 0 \\ e^{-2} & 0 & -2e^{-2} & e^{-2} \end{pmatrix}.$$

2.13 Applications to differential equations

Let us now consider an initial value problem for an autonomous system with continuous time:

$$\frac{dx(t)}{dt} = Ax(t),\ t \in [0,\infty),\ x(0) = w.$$

Here $A \in \mathbb{R}^{n\times n}$, $w \in \mathbb{R}^n$ are given, $x : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ is a smooth function to be found. One thinks of $x(t)$ as a state of the system at time t. The solution to this problem is

$$x(t) = e^{tA}w$$

because, as one can easily check,

$$\frac{d}{dt}(x(t)) = \sum_n \frac{d}{dt}\left(\frac{t^n}{n!}A^n w\right) = \sum_n \frac{t^{n-1}}{(n-1)!}A^n w = A\sum_k \frac{t^k}{k!}A^k w = Ax(t).$$

Let us consider a harmonic oscillator described by equation $y''(t) + y(t) = 0$. The general solution $y(t) = \alpha\sin(t) + \beta\cos(t)$ is well known. Let us obtain it using matrix exponents. Setting

$$x(t) = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix},\ A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

the harmonic oscillator becomes the initial value problem with a solution $x(t) = e^{tA}x(0)$. The eigenvalues of A are i and $-i$. Interpolating e^{zt} at these values of z gives the following condition on $h(z) = \alpha z + \beta$

$$\begin{cases} e^{it} &= h(i) &= \alpha i + \beta \\ e^{-it} &= h(-i) &= -\alpha i + \beta \end{cases}$$

Solving them gives $\alpha = (e^{it} - e^{-it})/2i = \sin(t)$ and $\beta = (e^{it} + e^{-it})/2 = \cos(t)$. It follows that

$$e^{tA} = \sin(t)A + \cos(t)I_2 = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$

and $y(t) = \cos(t)y(0) + \sin(t)y'(0)$.

As another example, let us consider a system of differential equations

$$\begin{cases} y_1' &= y_1 - 3y_3 \\ y_2' &= y_1 - y_2 - 6y_3 \\ y_3' &= -y_1 + 2y_2 + 5y_3 \end{cases} \quad \text{with initial condition} \quad \begin{cases} y_1(0) &= 1 \\ y_2(0) &= 1 \\ y_3(0) &= 0 \end{cases}$$

Using matrices

$$x(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix},\ w = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},\ A = \begin{pmatrix} 1 & 0 & -3 \\ 1 & -1 & -6 \\ -1 & 2 & 5 \end{pmatrix},$$

it becomes an initial value problem. The characteristic polynomial is $c_A(z) = -z^3 + 5z^2 - 8z + 4 = (1-z)(2-z)^2$. We need to interpolate e^{tz} at 1 and 2 by $h(z) = \alpha z^2 + \beta z + \gamma$. At the multiple root 2 we need to interpolate up to order 2 that involves tracking the derivative $(e^{tz})' = te^{tz}$:

$$\begin{cases} e^t & = & h(1) & = & \alpha + \beta + \gamma \\ e^{2t} & = & h(2) & = & 4\alpha + 2\beta + \gamma \\ te^{2t} & = & h'(2) & = & 4\alpha + \beta \end{cases}$$

Solving, $\alpha = (t-1)e^{2t} + e^t$, $\beta = (4-3t)e^{2t} - 4e^t$, $\gamma = (2t-3)e^{2t} + 4e^t$. It follows that

$$e^{tA} = e^{2t} \begin{pmatrix} 3t-3 & -6t+6 & -9t+6 \\ 3t-2 & -6t+4 & -9t+3 \\ -t & 2t & 3t+1 \end{pmatrix} + e^t \begin{pmatrix} 4 & -6 & -6 \\ 2 & -3 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$x(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = e^{tA} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} (3-3t)e^{2t} - 2e^t \\ (2-3t)e^{2t} - e^t \\ te^{2t} \end{pmatrix}.$$

3 Bilinear Maps and Quadratic Forms

3.1 Bilinear maps: definitions

Let V and W be vector spaces over a field K.

Definition. A *bilinear map* on W and V is a map $\tau : W \times V \to K$ such that

(i) $\tau(\alpha_1 \mathbf{w}_1 + \alpha_2 \mathbf{w}_2, \mathbf{v}) = \alpha_1 \tau(\mathbf{w}_1, \mathbf{v}) + \alpha_2 \tau(\mathbf{w}_2, \mathbf{v})$ and

(ii) $\tau(\mathbf{w}, \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2) = \alpha_1 \tau(\mathbf{w}, \mathbf{v}_1) + \alpha_2 \tau(\mathbf{w}, \mathbf{v}_2)$

for all $\mathbf{w}, \mathbf{w}_1, \mathbf{w}_2 \in W$, $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in V$, and $\alpha_1, \alpha_2 \in K$. Notice the difference between linear and bilinear maps. For instance, let $V = W = K$. Addition is a linear map but not bilinear. On the other hand, multiplication is bilinear but not linear.

Let us choose a basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of V and a basis $\mathbf{f}_1, \ldots, \mathbf{f}_m$ of W.

Let $\tau : W \times V \to K$ be a bilinear map, and let $\alpha_{ij} = \tau(\mathbf{f}_i, \mathbf{e}_j)$, for $1 \le i \le m$, $1 \le j \le n$. Then the $m \times n$ matrix $A = (\alpha_{ij})$ is defined to be the matrix of τ with respect to the bases $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and $\mathbf{f}_1, \ldots, \mathbf{f}_m$ of V and W.

For $\mathbf{v} \in V$, $\mathbf{w} \in W$, let $\mathbf{v} = x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n$ and $\mathbf{w} = y_1 \mathbf{f}_1 + \cdots + y_m \mathbf{f}_m$, and hence

$$\underline{\mathbf{v}} = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix} \in K^{n,1}, \quad \text{and} \quad \underline{\mathbf{w}} = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_m \end{pmatrix} \in K^{m,1}.$$

Then, by using the equations (i) and (ii) above, we get

$$\tau(\mathbf{w}, \mathbf{v}) = \sum_{i=1}^{m} \sum_{j=1}^{n} y_i \, \tau(\mathbf{f}_i, \mathbf{e}_j) \, x_j = \sum_{i=1}^{m} \sum_{j=1}^{n} y_i \, \alpha_{ij} \, x_j = \underline{\mathbf{w}}^{\mathrm{T}} A \underline{\mathbf{v}} \tag{2.1}$$

For example, let $V = W = \mathbb{R}^2$ and use the natural basis of V. Suppose that $A = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$. Then

$$\tau((y_1, y_2), (x_1, x_2)) = (y_1, y_2) \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = y_1 x_1 - y_1 x_2 + 2 y_2 x_1.$$

3.2 Bilinear maps: change of basis

We retain the notation of the previous subsection. As in Subsection 1.2 above, suppose that we choose new bases $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ of V and $\mathbf{f}'_1, \ldots, \mathbf{f}'_m$ of W, and let $P = (\sigma_{ij})$ and $Q = (\tau_{ij})$ be the associated basis change matrices. Then, by Proposition 1.1, if $\underline{\mathbf{v}'}$ and $\underline{\mathbf{w}'}$ are the column vectors representing the vectors $\mathbf{v}$ and $\mathbf{w}$ with respect to the bases $\{\mathbf{e}'_i\}$ and $\{\mathbf{f}'_i\}$, we have $P\underline{\mathbf{v}'} = \underline{\mathbf{v}}$ and $Q\underline{\mathbf{w}'} = \underline{\mathbf{w}}$, and so

$$\underline{\mathbf{w}}^{\mathrm{T}} A \underline{\mathbf{v}} = \underline{\mathbf{w}'}^{\mathrm{T}} Q^{\mathrm{T}} A P \underline{\mathbf{v}'},$$

and hence, by Equation (2.1):

Theorem 3.1 *Let A be the matrix of the bilinear map $\tau : W \times V \to K$ with respect to the bases $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and $\mathbf{f}_1, \ldots, \mathbf{f}_m$ of V and W, and let B be its matrix with respect to the bases $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ and $\mathbf{f}'_1, \ldots, \mathbf{f}'_m$ of V and W. Let P and Q be the basis change matrices, as defined above. Then $B = Q^{\mathrm{T}} A P$.*

Compare this result with Theorem 1.2.

We shall be concerned from now on only with the case where $V = W$. A bilinear map $\tau : V \times V \to K$ is called a *bilinear form* on V. Theorem 3.1 then becomes:

Theorem 3.2 *Let A be the matrix of the bilinear form τ on V with respect to the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of V, and let B be its matrix with respect to the basis $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ of V. Let P the basis change matrix with original basis $\{\mathbf{e}_i\}$ and new basis $\{\mathbf{e}'_i\}$. Then $B = P^{\mathrm{T}} A P$.*

So, in the example at the end of Subsection 3.1, if we choose the new basis $\mathbf{e}'_1 = (1\ \ {-1})$, $\mathbf{e}'_2 = (1\ 0)$ then $P = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$, $P^{\mathrm{T}} A P = \begin{pmatrix} 0 & -1 \\ 2 & 1 \end{pmatrix}$, and

$$\tau\Big((y'_1\mathbf{e}'_1 + y'_2\mathbf{e}'_2\,,\ x'_1\mathbf{e}'_1 + x'_2\mathbf{e}'_2)\Big) = -y'_1x'_2 + 2y'_2x'_1 + y'_2x'_2.$$

Definition. Matrices A and B are called *congruent* if there exists an invertible matrix P with $B = P^{\mathrm{T}} A P$.

Definition. A bilinear form τ on V is called *symmetric* if $\tau(\mathbf{w}, \mathbf{v}) = \tau(\mathbf{v}, \mathbf{w})$ for all $\mathbf{v}, \mathbf{w} \in V$. An $n \times n$ matrix A is called symmetric if $A^{\mathrm{T}} = A$.

We then clearly have:

Proposition 3.3 *The bilinear form τ is symmetric if and only if its matrix (with respect to any basis) is symmetric.*

The best known example is when $V = \mathbb{R}^n$, and τ is defined by

$$\tau((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

This form has matrix equal to the identity matrix I_n with respect to the standard basis of $\mathbb{R}^n$. Geometrically, it is equal to the normal scalar product $\tau(\mathbf{v}, \mathbf{w}) = |\mathbf{v}||\mathbf{w}|\cos\theta$, where θ is the angle between the vectors $\mathbf{v}$ and $\mathbf{w}$.

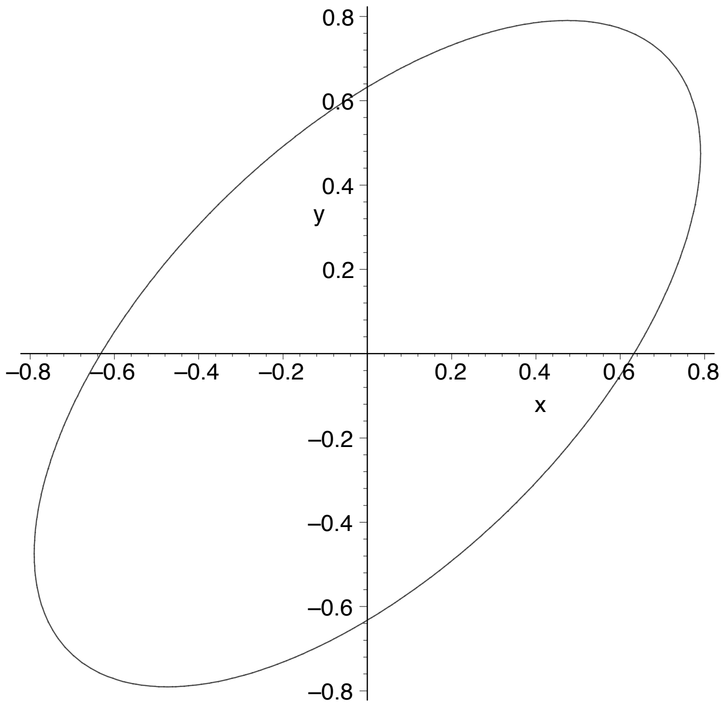


Figure 1: $5x^2 + 5y^2 - 6xy = 2$

3.3 Quadratic forms: introduction

A quadratic form on the standard vector space K^n is a polynomial function of several variables $x_1, \ldots, x_n$ in which each term has total degree two, such as $3x^2 + 2xz + z^2 - 4yz + xy$. One motivation comes to study them comes from the geometry of curves or surfaces defined by quadratic equations. Consider, for example, the equation $5x^2 + 5y^2 - 6xy = 2$ (see Fig. 1).

This represents an ellipse, in which the two principal axes are at an angle of $\pi/4$ with the x- and y-axes. To study such curves in general, it is desirable to change variables (which will turn out to be equivalent to a change of basis) so as to make the principal axes of the ellipse coincide with the x- and y-axes. This is equivalent to eliminating the xy-term in the equation. We can do this easily by completing the square.

In the example

$$5x^2 + 5y^2 - 6xy = 2 \Rightarrow 5(x - 3y/5)^2 - 9y^2/5 + 5y^2 = 2 \Rightarrow 5(x - 3y/5)^2 + 16y^2/5 = 2$$

so if we change variables, and put $x' = x - 3y/5$ and $y' = y$, then the equation becomes $5x'^2 + 16y'^2/5 = 2$ (see Fig. 2).

Here we have allowed an arbitrary basis change. We shall be studying this situation in Subsection 3.5.

One disadvantage of doing this is that the shape of the curve has become distorted. If we wish to preserve the shape, then we should restrict our basis changes to those that preserve distance and angle. These are called *orthogonal* basis changes, and we shall be studying that situation in Subsection 3.6. In the example, we can use the change of variables $x' = (x + y)/\sqrt{2}$, $y' = (x - y)/\sqrt{2}$ (which represents a non-distorting rotation through an angle of $\pi/4$), and the equation becomes $x'^2 + 4y'^2 = 1$. See Fig. 3.

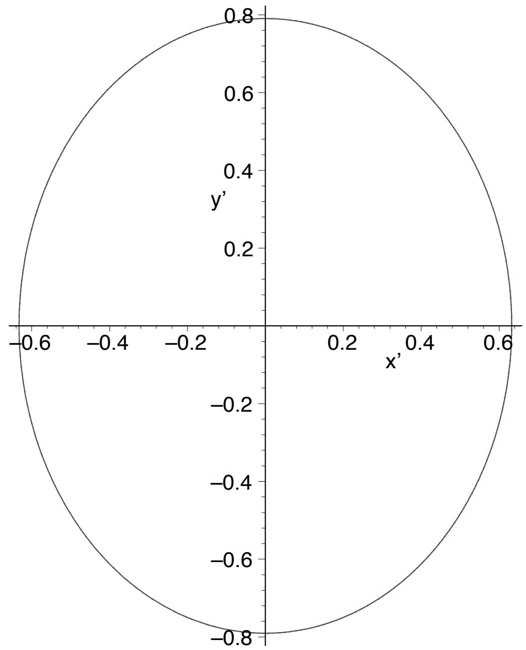


Figure 2: $5x'^2 + 16y'^2/5 = 2$

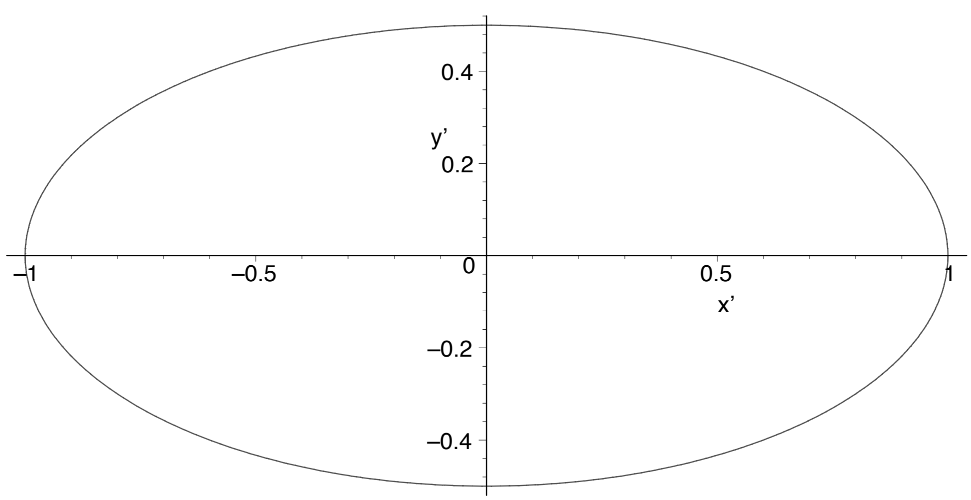


Figure 3: $x'^2 + 4y'^2 = 1$

3.4 Quadratic forms: definitions

Definition. Let V be a vector space over the field K. *A quadratic form* on V is a function $q : V \to K$ that is defined by $q(\mathbf{v}) = \tau(\mathbf{v}, \mathbf{v})$, where $\tau : V \times V \to K$ is a bilinear form.

As this is the official definition of a quadratic form we will use, we do not really need to observe that it yields the same notion for the standard vector space K^n as the definition in the previous section. However, it is a good exercise that an inquisitive reader should definitely do. The key is to observe that the function $x_i x_j$ comes from the bilinear form $\tau_{i,j}$ such that $\tau_{i,j}(e_i, e_j) = 1$ and zero elsewhere.

In Proposition 3.4 we need to be able to divide by 2 in the field K. This means that we must assume[11] that $1 + 1 \neq 0$ in K. For example, we would like to avoid the field of two elements. If you prefer to avoid worrying about such technicalities, then you can safely assume that K is either $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$.

Let us consider the following three sets. The first set $\mathrm{Q}(V, K)$ consists of all quadratic forms on V. It is a subset of the set of all functions from V to K. The second set $\mathrm{Bil}(V \times V, K)$ consists of all bilinear forms on V. It is a subset of the set of all functions from $V \times V$ to K. Finally, we need $\mathrm{Sym}(V \times V, K)$, the subset of $\mathrm{Bil}(V \times V, K)$ consisting of symmetric bilinear forms.

There are two interesting functions connecting these sets. We have already defined *a square function* $\Phi : \mathrm{Bil}(V \times V, K) \to \mathrm{Q}(V, K)$ by $\Phi(\tau)(\mathbf{v}) = \tau(\mathbf{v}, \mathbf{v})$. The second function $\Psi : \mathrm{Q}(V, K) \to \mathrm{Bil}(V \times V, K)$ is *a polarisation*[12] defined by $\Psi(q)(\mathbf{u}, \mathbf{v}) = q(\mathbf{u} + \mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v})$.

Proposition 3.4 *The following statements hold for all $q \in Q(V, K)$ and $\tau \in Sym(V \times V, K)$:*

(i) $\Psi(q) \in Sym(V \times V, K)$,
(ii) $\Phi(\Psi(q)) = 2q$,
(iii) $\Psi(\Phi(\tau)) = 2\tau$,
(iv) *if* $1 + 1 \neq 0 \in K$ *then there are natural*[13] *bijections between* $Q(V, K)$ *and* $Sym(V \times V, K)$.

PROOF: Observe that $q = \Phi(\tau)$ for some bilinear form τ. For $\mathbf{u}, \mathbf{v} \in V$, $q(\mathbf{u} + \mathbf{v}) = \tau(\mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}) = \tau(\mathbf{u}, \mathbf{u}) + \tau(\mathbf{v}, \mathbf{v}) + \tau(\mathbf{u}, \mathbf{v}) + \tau(\mathbf{v}, \mathbf{u}) = q(\mathbf{u}) + q(\mathbf{v}) + \tau(\mathbf{u}, \mathbf{v}) + \tau(\mathbf{v}, \mathbf{u})$. It follows that $\Psi(q)(\mathbf{u}, \mathbf{v}) = \tau(\mathbf{u}, \mathbf{v}) + \tau(\mathbf{v}, \mathbf{u})$ and that $\Psi(q)$ is a symmetric bilinear form. Besides it follows that $\Psi(\Phi(\tau)) = 2\tau$ if τ is symmetric.

Since $q(\alpha\mathbf{v}) = \alpha^2 q(\mathbf{v})$ for all $\alpha \in K$, $\mathbf{v} \in V$, $\Phi(\Psi(q))(\mathbf{v}) = \Psi(q)(\mathbf{v}, \mathbf{v}) = q(2\mathbf{v}) - q(\mathbf{v}) - q(\mathbf{v}) = 2q(\mathbf{v})$. Finally, if we can divide by 2 then Φ and $\Psi/2$ defined by $\Psi/2(q)(\mathbf{u}, \mathbf{v}) = (q(\mathbf{u} + \mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v}))/2$ provide inverse bijection between symmetric bilinear forms on V and quadratic forms on V. □

Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be a basis of V. Recall that the coordinates of $\mathbf{v}$ with respect to this basis are defined to be the field elements x_i such that $\mathbf{v} = \sum_{i=1}^{n} x_i \mathbf{e}_i$.

Let $A = (\alpha_{ij})$ be the matrix of τ with respect to this basis. We will also call A the matrix of q with respect to this basis. Then A is symmetric because τ is, and by Equation (2.1) of Subsection 3.1, we have

[11] Fields with $1 + 1 = 0$ are fields of characteristic 2. One can actually do quadratic and bilinear forms over them but the theory is quite specific. It could be a good topic for a second year essay.

[12] Some authors call it linearisation.

[13] There is a precise mathematical way of defining *natural* using *Category Theory* but it is far beyond the scope of this course. The only meaning we can endow this word with is that we do not make any choices for this bijection.

$$q(\mathbf{v}) = \underline{\mathbf{v}}^{\mathrm{T}} A \underline{\mathbf{v}} = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \alpha_{ij} x_j = \sum_{i=1}^{n} \alpha_{ii} x_i^2 + 2 \sum_{i=1}^{n} \sum_{j=1}^{i-1} \alpha_{ij} x_i x_j. \tag{3.1}$$

When $n \le 3$, we shall usually write x, y, z instead of x_1, x_2, x_3. For example, if $n = 2$ and $A = \begin{pmatrix} 1 & 3 \\ 3 & -2 \end{pmatrix}$, then $q(\mathbf{v}) = x^2 - 2y^2 + 6xy$.

Conversely, if we are given a quadratic form as in the right hand side of Equation (3.1), then it is easy to write down its matrix A. For example, if $n = 3$ and $q(\mathbf{v}) = 3x^2 + y^2 - 2z^2 + 4xy - xz$, then $A = \begin{pmatrix} 3 & 2 & -1/2 \\ 2 & 1 & 0 \\ -1/2 & 0 & -2 \end{pmatrix}$.

3.5 Change of variable under the general linear group

Our general aim is to make a change of basis so as to eliminate the terms in $q(\mathbf{v})$ that involve $x_i x_j$ for $i \ne j$, leaving only terms of the form $\alpha_{ii} x_i^2$. In this section, we will allow arbitrary basis changes; in other words, we allow basis change matrices P from the general linear group $\mathrm{GL}(n, K)$. It follows from Theorem 3.2 that when we make such a change, the matrix A of q is replaced by $P^{\mathrm{T}} A P$.

As with other results in linear algebra, we can formulate theorems either in terms of abstract concepts like quadratic forms, or simply as statements about matrices.

Theorem 3.5 *Assume that* $1 + 1 \ne 0 \in K$.

Let q be a quadratic form on V. Then there is a basis $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ of V such that $q(\mathbf{v}) = \sum_{i=1}^{n} \alpha_i (x'_i)^2$, where the x'_i are the coordinates of $\mathbf{v}$ with respect to $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$.

Equivalently, given any symmetric matrix A, there is an invertible matrix P such that $P^{\mathrm{T}} A P$ is a diagonal matrix; that is, A is congruent to a diagonal matrix.

PROOF: This is by induction on n. There is nothing to prove when $n = 1$. As usual, let $A = (\alpha_{ij})$ be the matrix of q with respect to the initial basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Case 1. First suppose that $\alpha_{11} \ne 0$. As in the example in Subsection 3.3, we can complete the square. We have

$$q(\mathbf{v}) = \alpha_{11} x_1^2 + 2\alpha_{12} x_1 x_2 + \cdots + 2\alpha_{1n} x_1 x_n + q_0(\mathbf{v}),$$

where q_0 is a quadratic form involving only the coordinates $x_2, \ldots, x_n$. So

$$q(\mathbf{v}) = \alpha_{11}(x_1 + \frac{\alpha_{12}}{\alpha_{11}} x_2 + \cdots + \frac{\alpha_{1n}}{\alpha_{11}} x_n)^2 + q_1(\mathbf{v}),$$

where $q_1(\mathbf{v})$ is another quadratic form involving only $x_2, \ldots, x_n$.

We now make the change of coordinates $x'_1 = x_1 + \frac{\alpha_{12}}{\alpha_{11}} x_2 + \cdots + \frac{\alpha_{1n}}{\alpha_{11}} x_n$, $x'_i = x_i$ for $2 \le i \le n$. Then we have $q(\mathbf{v}) = \alpha_1 (x'_1)^2 + q_1(\mathbf{v})$, where $\alpha_1 = \alpha_{11}$ and $q_1(\mathbf{v})$ involves only $x'_2, \ldots, x'_n$. By inductive hypothesis (applied to the subspace of V spanned by $\mathbf{e}_2, \ldots, \mathbf{e}_n$), we can change the coordinates of q_1 from $x'_2, \ldots, x'_n$ to $x''_2, \ldots, x''_n$, say, to bring it to the required form, and then we get $q(\mathbf{v}) = \sum_{i=1}^{n} \alpha_i (x''_i)^2$ (where $x''_1 = x'_1$) as required.

Case 2. $\alpha_{11} = 0$ but $\alpha_{ii} \ne 0$ for some $i > 1$. In this case, we start by interchanging $\mathbf{e}_1$ with $\mathbf{e}_i$ (or equivalently x_1 with x_i), which takes us back to Case 1.

Case 3. $\alpha_{ii} = 0$ for $1 \le i \le n$. If $\alpha_{ij} = 0$ for all i and j then there is nothing to prove, so assume that $\alpha_{ij} \ne 0$ for some i, j. Then we start by making a coordinate change $x_i = x'_i + x'_j$,

$x_j = x'_i - x'_j$, $x_k = x'_k$ for $k \neq i, j$. This introduces terms $2\alpha_{ij}((x'_i)^2 - (x'_j)^2)$ into q, taking us back to Case 2. □

Notice that, in the first change of coordinates in Case 1 of the proof, we have

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} 1 & \frac{\alpha_{12}}{\alpha_{11}} & \frac{\alpha_{13}}{\alpha_{11}} & \dots & \frac{\alpha_{1n}}{\alpha_{11}} \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & & \dots & & \\ & & \dots & & \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{or equivalently}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & -\frac{\alpha_{12}}{\alpha_{11}} & -\frac{\alpha_{13}}{\alpha_{11}} & \dots & -\frac{\alpha_{1n}}{\alpha_{11}} \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & & \dots & & \\ & & \dots & & \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}.$$

In other words, $\underline{\mathbf{v}} = P\underline{\mathbf{v}'}$, where

$$P = \begin{pmatrix} 1 & -\frac{\alpha_{12}}{\alpha_{11}} & -\frac{\alpha_{13}}{\alpha_{11}} & \dots & -\frac{\alpha_{1n}}{\alpha_{11}} \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix},$$

so by Proposition 1.1, P is the basis change matrix with original basis $\{\mathbf{e}_i\}$ and new basis $\{\mathbf{e}'_i\}$.

Example. Let $n = 3$ and $q(\mathbf{v}) = xy + 3yz - 5xz$, so $A = \begin{pmatrix} 0 & 1/2 & -5/2 \\ 1/2 & 0 & 3/2 \\ -5/2 & 3/2 & 0 \end{pmatrix}$.

Since we are using x, y, z for our variables, we can use x_1, y_1, z_1 (rather than x', y', z') for the variables with respect to a new basis, which will make things typographically simpler!

We are in Case 3 of the proof above, and so we start with a coordinate change $x = x_1 + y_1$, $y = x_1 - y_1$, $z = z_1$, which corresponds to the basis change matrix $P_1 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then we get $q(\mathbf{v}) = x_1^2 - y_1^2 - 2x_1z_1 - 8y_1z_1$.

We are now in Case 1 of the proof above, and the next basis change, from completing the square, is $x_2 = x_1 - z_1$, $y_2 = y_1$, $z_2 = z_1$, or equivalently, $x_1 = x_2 + z_2$, $y_1 = y_2$, $z_1 = z_2$, and then the associated basis change matrix is $P_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, and $q(\mathbf{v}) = x_2^2 - y_2^2 - 8y_2z_2 - z_2^2$.

We now proceed by induction on the 2-coordinate form in y_2, z_2, and completing the square again leads to the basis change $x_3 = x_2$, $y_3 = y_2 + 4z_2$, $z_3 = z_2$, which corresponds to the basis change matrix $P_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix}$, and $q(\mathbf{v}) = x_3^2 - y_3^2 + 15z_3^2$.

The total basis change in moving from the original basis with coordinates x, y, z to the final basis with coordinates x_3, y_3, z_3 is

$$P = P_1P_2P_3 = \begin{pmatrix} 1 & 1 & -3 \\ 1 & -1 & 5 \\ 0 & 0 & 1 \end{pmatrix},$$

and you can check that $P^{\mathrm{T}}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 15 \end{pmatrix}$, as expected.

Since P is an invertible matrix, P^{T} is also invertible (its inverse is $(P^{-1})^{\mathrm{T}}$), and so the matrices $P^{\mathrm{T}}AP$ and A are equivalent, and hence have the same rank. (This was proved in MA106.) The rank of the quadratic form q is defined to be the rank of its matrix A. So we have just shown that the rank of q is independent of the choice of basis used for the matrix A. If $P^{\mathrm{T}}AP$ is diagonal, then its rank is equal to the number of non-zero terms on the diagonal.

Notice that both statements of Theorem 3.5 fail in characteristic 2. Let K be the field of two elements. The quadratic form $q(x, y) = xy$ cannot be diagonalised. Similarly, the symmetric matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is not congruent to a diagonal matrix. Do it as an exercise: there are 6 possible change of variable (you need to choose two among three possible variables x, y and $x + y$) and you can observe directly what happens with each change of variables.

In the case $K = \mathbb{C}$, after reducing q to the form $q(\mathbf{v}) = \sum_{i=1}^{n} \alpha_{ii}x_i^2$, we can permute the coordinates if necessary to get $\alpha_{ii} \neq 0$ for $1 \leq i \leq r$ and $\alpha_{ii} = 0$ for $r + 1 \leq i \leq n$, where $r = \mathrm{rank}(q)$. We can then make a further coordinates change $x_i' = \sqrt{\alpha_{ii}}x_i$ $(1 \leq i \leq r)$, giving $q(\mathbf{v}) = \sum_{i=1}^{r}(x_i')^2$. Hence we have proved:

Proposition 3.6 *A quadratic form q over $\mathbb{C}$ has the form $q(\mathbf{v}) = \sum_{i=1}^{r} x_i^2$ with respect to a suitable basis, where $r = \mathrm{rank}(q)$.*

Equivalently, given a symmetric matrix $A \in \mathbb{C}^{n,n}$, there is an invertible matrix $P \in \mathbb{C}^{n,n}$ such that $P^{\mathrm{T}}AP = B$, where $B = (\beta_{ij})$ is a diagonal matrix with $\beta_{ii} = 1$ for $1 \leq i \leq r$, $\beta_{ii} = 0$ for $r + 1 \leq i \leq n$, and $r = \mathrm{rank}(A)$.

When $K = \mathbb{R}$, we cannot take square roots of negative numbers, but we can replace each positive α_i by 1 and each negative α_i by -1 to get:

Proposition 3.7 (Sylvester's Theorem) *A quadratic form q over $\mathbb{R}$ has the form $q(\mathbf{v}) = \sum_{i=1}^{t} x_i^2 - \sum_{i=1}^{u} x_{t+i}^2$ with respect to a suitable basis, where $t + u = \mathrm{rank}(q)$.*

Equivalently, given a symmetric matrix $A \in \mathbb{R}^{n,n}$, there is an invertible matrix $P \in \mathbb{R}^{n,n}$ such that $P^{\mathrm{T}}AP = B$, where $B = (\beta_{ij})$ is a diagonal matrix with $\beta_{ii} = 1$ for $1 \leq i \leq t$, $\beta_{ii} = -1$ for $t + 1 \leq i \leq t + u$, and $\beta_{ii} = 0$ for $t + u + 1 \leq i \leq n$, and $t + u = \mathrm{rank}(A)$.

We shall now prove that the numbers t and u of positive and negative terms are invariants of q. The difference $t - u$ between the numbers of positive and negative terms is called the *signature* of q.

Theorem 3.8 (Sylvester's Law of Intertia) *Suppose that q is a quadratic form over the vector space V over $\mathbb{R}$, and that $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and $\mathbf{e}_1', \ldots, \mathbf{e}_n'$ are two bases of V with associated coordinates x_i and x_i', such that*

$$q(\mathbf{v}) = \sum_{i=1}^{t} x_i^2 - \sum_{i=1}^{u} x_{t+i}^2 = \sum_{i=1}^{t'} (x_i')^2 - \sum_{i=1}^{u'} (x_{t'+i})^2.$$

Then $t = t'$ and $u = u'$.

Proof: We know that $t + u = t' + u' = \text{rank}(q)$, so it is enough to prove that $t = t'$. Suppose not, and suppose that $t > t'$. Let $V_1 = \{\mathbf{v} \in V \mid x_{t+1} = x_{t+2} = \ldots = x_n = 0\}$, and let $V_2 = \{\mathbf{v} \in V \mid x'_1 = x'_2 = \ldots = x'_{t'} = 0\}$. Then V_1 and V_2 are subspaces of V with $\dim(V_1) = t$ and $\dim(V_2) = n - t'$. It was proved in MA106 that

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2).$$

However, $\dim(V_1 + V_2) \le \dim(V) = n$, and so $t > t'$ implies that $\dim(V_1) + \dim(V_2) > n$. Hence $\dim(V_1 \cap V_2) > 0$, and there is a non-zero vector $\mathbf{v} \in V_1 \cap V_2$. But it is easily seen from the expressions for $q(\mathbf{v})$ in the statement of the theorem that $0 \ne \mathbf{v} \in V_1 \Rightarrow q(\mathbf{v}) > 0$, whereas $\mathbf{v} \in V_2 \Rightarrow q(\mathbf{v}) \le 0$, which is a contradiction, and completes the proof. □

3.6 Change of variable under the orthogonal group

In this subsection, we assume throughout that $K = \mathbb{R}$.

Definition. A quadratic form q on V is said to be *positive definite* if $q(\mathbf{v}) > 0$ for all $0 \ne \mathbf{v} \in V$.

It is clear that this is the case if and only if $t = n$ and $u = 0$ in Proposition 3.7; that is, if q has rank and signature n. In this case, Proposition 3.7 says that there is a basis $\{\mathbf{e}_i\}$ of V with respect to which $q(\mathbf{v}) = \sum_{i=1}^n x_i^2$ or, equivalently, such that the matrix A of q is the identity matrix I_n.

The associated symmetric bilinear form τ is also called positive definite when q is. If we use a basis such that $A = I_n$, then τ is just the standard scalar (or inner) product on V.

Definition. A vector space V over $\mathbb{R}$ together with a positive definite symmetric bilinear form τ is called a *Euclidean space*.

We shall assume from now on that V is a Euclidean space, and that the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ has been chosen so that the matrix of τ is I_n. Since τ is the standard scalar product, we shall write $\mathbf{v} \cdot \mathbf{w}$ instead of $\tau(\mathbf{v}, \mathbf{w})$.

Note that $\mathbf{v} \cdot \mathbf{w} = \underline{\mathbf{v}}^{\mathrm{T}} \underline{\mathbf{w}}$ where, as usual, $\underline{\mathbf{v}}$ and $\underline{\mathbf{w}}$ are the column vectors associated with $\mathbf{v}$ and $\mathbf{w}$.

For $\mathbf{v} \in V$, define $|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$. Then $|\mathbf{v}|$ is the length of $\mathbf{v}$. Hence the length, and also the cosine $\mathbf{v} \cdot \mathbf{w}/(|\mathbf{v}||\mathbf{w}|)$ of the angle between two vectors can be defined in terms of the scalar product.

Definition. A linear operator $T : V \to V$ is said to be *orthogonal* if it preserves the scalar product on V. That is, if $T(\mathbf{v}) \cdot T(\mathbf{w}) = \mathbf{v} \cdot \mathbf{w}$ for all $\mathbf{v}, \mathbf{w} \in V$.

Since length and angle can be defined in terms of the scalar product, an orthogonal linear operator preserves distance and angle, so geometrically it is a rigid map. In $\mathbb{R}^2$, for example, an orthogonal operator is a rotation about the origin or a reflection about a line through the origin.

If A is the matrix of T, then $T(\mathbf{v}) = A\underline{\mathbf{v}}$, so $T(\mathbf{v}) \cdot T(\mathbf{w}) = \underline{\mathbf{v}}^{\mathrm{T}} A^{\mathrm{T}} A \underline{\mathbf{w}}$, and hence T is orthogonal if and only if $A^{\mathrm{T}} A = I_n$, or equivalently if $A^{\mathrm{T}} = A^{-1}$.

Definition. An $n \times n$ matrix is called *orthogonal* if $A^{\mathrm{T}} A = I_n$.

So we have proved:

Proposition 3.9 *A linear operator $T : V \to V$ is orthogonal if and only if its matrix A (with respect to a basis such that the matrix of the bilinear form τ is I_n) is orthogonal.*

Incidentally, the fact that $A^{\mathrm{T}}A = I_n$ tells us that A and hence T is invertible, and so we have also proved:

Proposition 3.10 *An orthogonal linear operator is invertible.*

Let $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ be the columns of the matrix A. As we observed in Subsection 1.1, $\mathbf{c}_i$ is equal to the column vector representing $T(\mathbf{e}_i)$. In other words, if $T(\mathbf{e}_i) = \mathbf{f}_i$ then $\underline{\mathbf{f}_i} = \mathbf{c}_i$. Since the (i,j)-th entry of $A^{\mathrm{T}}A$ is $\mathbf{c}_i^{\mathrm{T}}\mathbf{c}_j = \mathbf{f}_i \cdot \mathbf{f}_j$, we see that T and A are orthogonal if and only if[14]

$$\mathbf{f}_i \cdot \mathbf{f}_i = \delta_{i,j}, \quad 1 \le i,j \le n. \qquad (*)$$

Definition. A basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ of V that satisfies $(*)$ is called *orthonormal*.

By Proposition 3.10, an orthogonal linear operator is invertible, so $T(\mathbf{e}_i)$ $(1 \le i \le n)$ form a basis of V, and we have:

Proposition 3.11 *A linear operator T is orthogonal if and only if $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)$ is an orthonormal basis of V.*

Example For any $\theta \in \mathbb{R}$, let $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. (This represents a counter-clockwise rotation through an angle θ.) Then it is easily checked that $A^{\mathrm{T}}A = AA^{\mathrm{T}} = I_2$. Notice that the columns of A are mutually orthogonal vectors of length 1, and the same applies to the rows of A.

The following theorem tells us that we can always complete an orthonormal set of vectors to an orthonormal basis.

Theorem 3.12 (Gram-Schmidt) *Let V be a Euclidean space of dimension n, and suppose that, for some r with $0 \le r \le n$, $\mathbf{f}_1, \ldots, \mathbf{f}_r$ are vectors in V that satisfy the equations $(*)$ for $1 \le i,j \le r$. Then $\mathbf{f}_1, \ldots, \mathbf{f}_r$ can be extended to an orthonormal basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ of V.*

PROOF: We prove first that $\mathbf{f}_1, \ldots, \mathbf{f}_r$ are linearly independent. Suppose that $\sum_{i=1}^{r} x_i\mathbf{f}_i = \mathbf{0}$ for some $x_1, \ldots, x_r \in \mathbb{R}$. Then, for each j with $1 \le j \le r$, the scalar product of the left hand side of this equation with $\mathbf{f}_j$ is $\sum_{i=1}^{r} x_i\mathbf{f}_j \cdot \mathbf{f}_i = x_j$, by $(*)$. Since $\mathbf{f}_j \cdot \mathbf{0} = 0$, this implies that $x_j = 0$ for all j, so the $\mathbf{f}_i$ are linearly independent.

The proof of the theorem will be by induction on $n - r$. We can start the induction with the case $n - r = 0$, when $r = n$, and there is nothing to prove. So assume that $n - r > 0$; i.e. that $r < n$. By a result from MA106, we can extend any linearly independent set of vectors to a basis of V, so there is a basis $\mathbf{f}_1, \ldots, \mathbf{f}_r, \mathbf{g}_{r+1}, \ldots, \mathbf{g}_n$ of V containing the $\mathbf{f}_i$. The trick is to define

$$\mathbf{f}'_{r+1} = \mathbf{g}_{r+1} - \sum_{i=1}^{r} (\mathbf{f}_i \cdot \mathbf{g}_{r+1})\mathbf{f}_i.$$

If we take the scalar product of this equation by $\mathbf{f}_j$ for some $0 \le j \le r$, then we get

$$\mathbf{f}_j \cdot \mathbf{f}'_{r+1} = \mathbf{f}_j \cdot \mathbf{g}_{r+1} - \sum_{i=1}^{r} (\mathbf{f}_i \cdot \mathbf{g}_{r+1})(\mathbf{f}_j \cdot \mathbf{f}_i)$$

and then, by $(*)$, $\mathbf{f}_j \cdot \mathbf{f}_i$ is non-zero only when $j = i$, so the sum on the right hand side simplifies to $\mathbf{f}_j \cdot \mathbf{g}_{r+1}$, and the whole equation simplifies to $\mathbf{f}_j \cdot \mathbf{f}'_{r+1} = \mathbf{f}_j \cdot \mathbf{g}_{r+1} - \mathbf{f}_j \cdot \mathbf{g}_{r+1} = 0$.

[14] We are using Kronecker's delta symbol in the next formula. It is just the identity matrix $I_m = (\delta_{i,j})$ of sufficiently large size. In layman's terms, $\delta_{i,i} = 1$ and $\delta_{i,j} = 0$ if $i \ne j$.

The vector $\mathbf{f}'_{r+1}$ is non-zero by linear independence of the basis, and if we define $\mathbf{f}_{r+1} = \mathbf{f}'_{r+1}/|\mathbf{f}'_{r+1}|$, then we still have $\mathbf{f}_j \cdot \mathbf{f}_{r+1} = 0$ for $1 \leq j \leq r$, and we also have $\mathbf{f}_{r+1}.\mathbf{f}_{r+1} = 1$. Hence $\mathbf{f}_1, \ldots, \mathbf{f}_{r+1}$ satisfy the equations $(*)$, and the result follows by inductive hypothesis. □

Recall from MA106 that if T is a linear operator with matrix A, and $\mathbf{v}$ is a non-zero vector such that $T(\mathbf{v}) = \lambda\mathbf{v}$ (or equivalently $A\underline{\mathbf{v}} = \lambda\underline{\mathbf{v}}$), then λ is called an *eigenvalue* and $\mathbf{v}$ an associated *eigenvector* of T and A. It was proved in MA106 that the eigenvalues are the roots of the characteristic equation $\det(A - xI_n) = 0$ of A.

Proposition 3.13 *Let A be a real symmetric matrix. Then A has an eigenvalue in $\mathbb{R}$, and all complex eigenvalues of A lie in $\mathbb{R}$.*

PROOF: (To simplify the notation, we will write just $\mathbf{v}$ for a column vector $\underline{\mathbf{v}}$ in this proof.)

The characteristic equation $\det(A - xI_n) = 0$ is a polynomial equation of degree n in x, and since $\mathbb{C}$ is an algebraically closed field, it certainly has a root $\lambda \in \mathbb{C}$, which is an eigenvalue for A if we regard A as a matrix over $\mathbb{C}$. We shall prove that any such λ lies in $\mathbb{R}$, which will prove the proposition.

For a column vector $\mathbf{v}$ or matrix B over $\mathbb{C}$, we denote by $\overline{\mathbf{v}}$ or $\overline{B}$ the result of replacing all entries of $\mathbf{v}$ or B by their complex conjugates. Since the entries of A lie in $\mathbb{R}$, we have $\overline{A} = A$.

Let $\mathbf{v}$ be a complex eigenvector associated with λ. Then

$$A\mathbf{v} = \lambda\mathbf{v} \tag{1}$$

so, taking complex conjugates and using $\overline{A} = A$, we get

$$A\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}. \tag{2}$$

Transposing (1) and using $A^T = A$ gives

$$\mathbf{v}^{\mathrm{T}}A = \lambda\mathbf{v}^{\mathrm{T}}, \tag{3}$$

so by (2) and (3) we have

$$\lambda\mathbf{v}^{\mathrm{T}}\overline{\mathbf{v}} = \mathbf{v}^{\mathrm{T}}A\overline{\mathbf{v}} = \overline{\lambda}\mathbf{v}^{\mathrm{T}}\overline{\mathbf{v}}.$$

But if $\mathbf{v} = (\alpha_1, \alpha_2, \ldots, \alpha_n)^{\mathrm{T}}$, then $\mathbf{v}^{\mathrm{T}}\overline{\mathbf{v}} = \alpha_1\overline{\alpha_1} + \cdots + \alpha_n\overline{\alpha_n}$, which is a non-zero real number (eigenvectors are non-zero by definition). Thus $\lambda = \overline{\lambda}$, so $\lambda \in \mathbb{R}$. □

Before coming to the main theorem of this section, we recall the notation $A \oplus B$ for matrices, which we introduced in Subsection 2.5. It is straightforward to check that $(A_1 \oplus B_1)\ (A_2 \oplus B_2)$ $=(A_1A_2 \oplus B_1B_2)$, provided that A_1 and A_2 are matrices with the same dimensions.

Theorem 3.14 *Let q be a quadratic form defined on a Euclidean space V. Then there is an orthonormal basis $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ of V such that $q(\mathbf{v}) = \sum_{i=1}^n \alpha_i (x'_i)^2$, where x'_i are the coordinates of $\mathbf{v}$ with respect to $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$. Furthermore, the numbers α_i are uniquely determined by q.*

Equivalently, given any symmetric matrix A, there is an orthogonal matrix P such that $P^{\mathrm{T}}AP$ is a diagonal matrix. Since $P^T = P^{-1}$, this is saying that A is simultaneously similar and congruent to a diagonal matrix.

PROOF: We start with a general remark about orthogonal basis changes. The matrix q represents a quadratic form on V with respect to the initial orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of V, but it also represents a linear operator $T : V \to V$ with respect to the same basis. When we make an orthogonal basis change with original basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and a new orthonormal basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ with the basis change matrix P, then P is orthogonal, so $P^{\mathrm{T}} = P^{-1}$ and hence $P^{\mathrm{T}}AP = P^{-1}AP$. Hence, by Theorems 1.2 and 3.2, the matrix $P^{\mathrm{T}}AP$ simultaneously represents both the linear operator T and the quadratic form q with respect to the new basis.

Recall from MA106 that two $n \times n$ matrices are called *similar* if there exists an invertible $n \times n$ matrix P with $B = P^{-1}AP$. In particular, if P is orthogonal, then A and $P^{\mathrm{T}}AP$ are similar. It was proved in MA106 that similar matrices have the same eigenvalues. But the α_i are precisely the eigenvalues of the diagonalised matrix $P^{\mathrm{T}}AP$, and so the α_i are the eigenvalues of the original matrix A, and hence are uniquely determined by A and q. This proves the uniqueness part of the theorem.

The equivalence of the two statements in the theorem follows from Proposition 3.11 and Theorem 3.2. Their proof will by induction on $n = \dim(V)$. There is nothing to prove when $n = 1$. By Proposition 3.13, A and its associated linear operator T have a real eigenvalue α_1. Let $\mathbf{v}$ be a corresponding eigenvector (of T). Then $\mathbf{f}_1 = \mathbf{v}/|\mathbf{v}|$ is also an eigenvector (i.e. $T\mathbf{f}_1 = \alpha_1\mathbf{f}_1$), and $\mathbf{f}_1 \cdot \mathbf{f}_1 = 1$. By Theorem 3.12, there is an orthonormal basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ of V containing $\mathbf{f}_1$. Let B be the matrix of q with respect to this basis, so $B = P^{\mathrm{T}}AP$ with P orthogonal. By the remark above, B is also the matrix of T with respect to $\mathbf{f}_1, \ldots, \mathbf{f}_n$, and because $T\mathbf{f}_1 = \alpha_1\mathbf{f}_1$, the first column of B is $(\alpha_1, 0, \ldots, 0)^{\mathrm{T}}$. But B is the matrix of the quadratic form q, so it is symmetric, and hence the first row of B is $(\alpha_1, 0, \ldots, 0)$, and therefore $B = P^{\mathrm{T}}AP = P^{-1}AP = (\alpha_1) \oplus A_1$, where A_1 is an $(n-1) \times (n-1)$ matrix and (α_1) is a 1×1 matrix.

Furthermore, B symmetric implies A_1 symmetric, and by inductive assumption there is an $(n-1) \times (n-1)$ orthogonal matrix Q_1 with $Q_1^{\mathrm{T}}A_1Q_1$ diagonal. Let $Q = (1) \oplus Q_1$. Then Q is also orthogonal (check!) and we have $(PQ)^{\mathrm{T}}A(PQ) = Q^{\mathrm{T}}(P^{\mathrm{T}}AP)Q = (\alpha_1) \oplus Q_1^{\mathrm{T}}A_1Q_1$ is diagonal. But PQ is the product of two orthogonal matrices and so is itself orthogonal. This completes the proof. □

Although it is not used in the proof of the theorem above, the following proposition is useful when calculating examples. It helps us to write down more vectors in the final orthonormal basis immediately, without having to use Theorem 3.12 repeatedly.

Proposition 3.15 *Let A be a real symmetric matrix, and let λ_1, λ_2 be two distinct eigenvalues of A, with corresponding eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$. Then $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.*

Proof: (As in Proposition 3.13, we will write $\mathbf{v}$ rather than $\underline{\mathbf{v}}$ for a column vector in this proof. So $\mathbf{v}_1 \cdot \mathbf{v}_2$ is the same as $\mathbf{v}_1^{\mathrm{T}}\mathbf{v}_2$.) We have

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1 \quad (1) \qquad \text{and} \qquad A\mathbf{v}_2 = \lambda_2\mathbf{v}_2 \quad (2).$$

Transposing (1) and using $A = A^{\mathrm{T}}$ gives $\mathbf{v}_1^{\mathrm{T}}A = \lambda_1\mathbf{v}_1^{\mathrm{T}}$, and so

$$\mathbf{v}_1^{\mathrm{T}}A\mathbf{v}_2 = \lambda_1\mathbf{v}_1^{T}\mathbf{v}_2 \quad (3) \qquad \text{and similarly} \qquad \mathbf{v}_2^{\mathrm{T}}A\mathbf{v}_1 = \lambda_2\mathbf{v}_2^{T}\mathbf{v}_1 \quad (4).$$

Transposing (4) gives $\mathbf{v}_1^{\mathrm{T}}A\mathbf{v}_2 = \lambda_2\mathbf{v}_1^{T}\mathbf{v}_2$ and subtracting (3) from this gives $(\lambda_2 - \lambda_1)\mathbf{v}_1^{T}\mathbf{v}_2 = 0$. Since $\lambda_2 - \lambda_1 \neq 0$ by assumption, we have $\mathbf{v}_1^{T}\mathbf{v}_2 = 0$. □

Example 1. Let $n = 2$ and $q(\mathbf{v}) = x^2 + y^2 + 6xy$, so $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$. Then

$$\det(A - xI_2) = (1-x)^2 - 9 = x^2 - 2x - 8 = (x-4)(x+2),$$

so the eigenvalues of A are 4 and -2. Solving $A\mathbf{v} = \lambda\mathbf{v}$ for $\lambda = 4$ and -2, we find corresponding eigenvectors $(1\ 1)^{\mathrm{T}}$ and $(1\ -1)^{\mathrm{T}}$. Proposition 3.15 tells us that these vectors are orthogonal to each other (which we can of course check directly!), so if we divide them by their lengths to give vectors of length 1, giving $(\frac{1}{\sqrt{2}}\ \frac{1}{\sqrt{2}})^{\mathrm{T}}$ and $(\frac{1}{\sqrt{2}}\ \frac{-1}{\sqrt{2}})^{\mathrm{T}}$ then we get an orthonormal basis consisting of eigenvectors of A, which is what we want. The corresponding basis change matrix P has these vectors as columns, so $P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$, and we can check that $P^{\mathrm{T}}P = I_2$ (hence P is orthogonal) and that $P^{\mathrm{T}}AP = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$.

Example 2. Let $n = 3$ and

$$q(\mathbf{v}) = 3x^2 + 6y^2 + 3z^2 - 4xy - 4yz + 2xz,$$

$$\text{so } A = \begin{pmatrix} 3 & -2 & 1 \\ -2 & 6 & -2 \\ 1 & -2 & 3 \end{pmatrix}.$$

Then, expanding by the first row,

$$\begin{aligned} \det(A - xI_3) &= (3-x)(6-x)(3-x) - 4(3-x) - 4(3-x) + 4 + 4 - (6-x) \\ &= -x^3 + 12x^2 - 36x + 32 = (2-x)(x-8)(x-2), \end{aligned}$$

so the eigenvalues are 2 (repeated) and 8. For the eigenvalue 8, if we solve $A\mathbf{v} = 8\mathbf{v}$ then we find a solution $\mathbf{v} = (1\ \ -2\ 1)^{\mathrm{T}}$. Since 2 is a repeated eigenvalue, we need two corresponding eigenvectors, which must be orthogonal to each other. The equations $A\mathbf{v} = 2\mathbf{v}$ all reduce to $x - 2y + z = 0$, and so any vector $(x, y, z)^{\mathrm{T}}$ satisfying this equation is an eigenvector for $\lambda = 2$. By Proposition 3.15 these eigenvectors will all be orthogonal to the eigenvector for $\lambda = 8$, but we will have to choose them orthogonal to each other. We can choose the first one arbitrarily, so let's choose $(1\ 0\ \ -1)^{\mathrm{T}}$. We now need another solution that is orthogonal to this. In other words, we want x, y and z not all zero satisfying $x - 2y + z = 0$ and $x - z = 0$, and $x = y = z = 1$ is a solution. So we now have a basis $(1\ \ -2\ 1)^{\mathrm{T}}$, $(1\ 0\ \ -1)^{\mathrm{T}}$, $(1\ 1\ 1)^{\mathrm{T}}$ of three mutually orthogonal eigenvectors. To get an orthonormal basis, we just need to divide by their lengths, which are, respectively, $\sqrt{6}$, $\sqrt{2}$, and $\sqrt{3}$, and then the basis change matrix P has these vectors as columns, so

$$P = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

It can then be checked that $P^{\mathrm{T}}P = I_3$ and that $P^{\mathrm{T}}AP$ is the diagonal matrix with entries 8, 2, 2.

3.7 Applications of quadratic forms to geometry

3.7.1 Reduction of the general second degree equation

The general equation of the second degree in n variables $x_1, \ldots, x_n$ is

$$\sum_{i=1}^{n} \alpha_i x_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{i-1} \alpha_{ij} x_i x_j + \sum_{i=1}^{n} \beta_i x_i + \gamma = 0.$$

This defines a *quadric hypersurface*[15] in n-dimensional Euclidean space. To study the possible shapes of the curves and surfaces defined, we first simplify this equation by applying coordinate changes resulting from isometries of $\mathbb{R}^n$.

[15]also called quadric surface if $n = 3$ or quadric curve if $n = 3$.

By Theorem 3.14, we can apply an orthogonal basis change (that is, an isometry of $\mathbb{R}^n$ that fixes the origin) which has the effect of eliminating the terms $\alpha_{ij}x_ix_j$ in the above sum.

Now, whenever $\alpha_i \neq 0$, we can replace x_i by $x_i - \beta_i/(2\alpha_i)$, and thereby eliminate the term $\beta_i x_i$ from the equation. This transformation is just a translation, which is also an isometry.

If $\alpha_i = 0$, then we cannot eliminate the term $\beta_i x_i$. Let us permute the coordinates such that $\alpha_i \neq 0$ for $1 \le i \le r$, and $\beta_i \neq 0$ for $r+1 \le i \le r+s$. Then if $s > 1$, by using Theorem 3.12, we can find an orthogonal transformation that leaves x_i unchanged for $1 \le i \le r$ and replaces $\sum_{i=1}^{s} \beta_{r+j}x_{r+j}$ by βx_{r+1} (where β is the length of $\sum_{i=1}^{s} \beta_{r+j}x_{r+j}$), and then we have only a single non-zero β_i; namely $\beta_{r+1} = \beta$.

Finally, if there is a non-zero $\beta_{r+1} = \beta$, then we can perform the translation that replaces x_{r+1} by $x_{r+1} - \gamma/\beta$, and thereby eliminate γ.

We have now reduced to one of two possible types of equation:

$$\sum_{i=1}^{r} \alpha_i x_i^2 + \gamma = 0 \qquad \text{and} \qquad \sum_{i=1}^{r} \alpha_i x_i^2 + \beta x_{r+1} = 0.$$

In fact, by dividing through by γ or β, we can assume that $\gamma = 0$ or 1 in the first equation, and that $\beta = 1$ in the second. In both cases, we shall assume that $r \neq 0$, because otherwise we have a linear equation. The curve defined by the first equation is called a *central quadric* because it has central symmetry; i.e. if a vector $\mathbf{v}$ satisfies the equation, then so does $-\mathbf{v}$.

We shall now consider the types of curves and surfaces that can arise in the familiar cases $n = 2$ and $n = 3$. These different types correspond to whether the α_i are positive, negative or zero, and whether $\gamma = 0$ or 1.

We shall use x, y, z instead of x_1, x_2, x_3, and α, β, γ instead of α_1, α_2, α_3. We shall assume also that α, β, γ are all strictly positive, and write $-\alpha$, etc., for the negative case.

3.7.2 The case $n = 2$

When $n = 2$ we have the following possibilities.

(i) $\alpha x^2 = 0$. This just defines the line $x = 0$ (the y-axis).

(ii) $\alpha x^2 = 1$. This defines the two parallel lines $x = \pm\frac{1}{\sqrt{\alpha}}$.

(iii) $-\alpha x^2 = 1$. This is the empty curve!

(iv) $\alpha x^2 + \beta y^2 = 0$. The single point $(0, 0)$.

(v) $\alpha x^2 - \beta y^2 = 0$. This defines two straight lines $y = \pm\sqrt{\frac{\alpha}{\beta}}\, x$, which intersect at $(0, 0)$.

(vi) $\alpha x^2 + \beta y^2 = 1$. An ellipse.

(vii) $\alpha x^2 - \beta y^2 = 1$. A hyperbola.

(viii) $-\alpha x^2 - \beta y^2 = 1$. The empty curve again.

(ix) $\alpha x^2 - y = 0$. A parabola.

3.7.3 The case $n = 3$

When $n = 3$, we still get the nine possibilities (i) – (ix) that we had in the case $n = 2$, but now they must be regarded as equations in the three variables x, y, z that happen not to involve z.

So, in Case (i), we now get the plane $x = 0$, in case (ii) we get two parallel planes $x = \pm 1/\sqrt{\alpha}$, in Case (iv) we get the line $x = y = 0$ (the z-axis), in case (v) two intersecting planes $y = \pm\sqrt{\alpha/\beta}x$, and in Cases (vi), (vii) and (ix), we get, respectively, elliptical, hyperbolic and parabolic cylinders.

The remaining cases involve all of x, y and z. We omit $-\alpha x^2 - \beta y^2 - \gamma z^2 = 1$, which is empty.

(x) $\alpha x^2 + \beta y^2 + \gamma z^2 = 0$. The single point $(0, 0, 0)$.

(xi) $\alpha x^2 + \beta y^2 - \gamma z^2 = 0$. See Fig. 4.

This is an elliptical cone. The cross sections parallel to the xy-plane are ellipses of the form $\alpha x^2 + \beta y^2 = c$, whereas the cross sections parallel to the other coordinate planes are generally hyperbolas. Notice also that if a particular point (a, b, c) is on the surface, then so is $t(a, b, c)$ for any $t \in \mathbb{R}$. In other words, the surface contains the straight line through the origin and any of its points. Such lines are called *generators*. When each point of a 3-dimensional surface lies on one or more generators, it is possible to make a model of the surface with straight lengths of wire or string.

(xii) $\alpha x^2 + \beta y^2 + \gamma z^2 = 1$. An ellipsoid. See Fig. 5.

(xiii) $\alpha x^2 + \beta y^2 - \gamma z^2 = 1$. A hyperboloid. See Fig. 6.

There are two types of 3-dimensional hyperboloids. This one is connected, and is known as a *hyperboloid of one sheet*. Although it is not immediately obvious, each point of this surface lies on exactly two generators; that is, lines that lie entirely on the surface. For each $\lambda \in \mathbb{R}$, the line defined by the pair of equations

$$\sqrt{\alpha}\,x - \sqrt{\gamma}\,z = \lambda(1 - \sqrt{\beta}\,y); \qquad \lambda(\sqrt{\alpha}\,x + \sqrt{\gamma}\,z) = 1 + \sqrt{\beta}\,y.$$

lies entirely on the surface; to see this, just multiply the two equations together. The same applies to the lines defined by the pairs of equations

$$\sqrt{\beta}\,y - \sqrt{\gamma}\,z = \mu(1 - \sqrt{\alpha}\,x); \qquad \mu(\sqrt{\beta}\,y + \sqrt{\gamma}\,z) = 1 + \sqrt{\alpha}\,x.$$

It can be shown that each point on the surface lies on exactly one of the lines in each if these two families.

(xiv) $\alpha x^2 - \beta y^2 - \gamma z^2 = 1$. A hyperboloid. See Fig. 7.

This one has two connected components and is called a *hyperboloid of two sheets*. It does not have generators. Besides it is easy to observe that it is disconnected. Substitute $x = 0$ into its equation. The resulting equation $-\beta y^2 - \gamma z^2 = 1$ has no solutions. This means that the hyperboloid does not intersect the plane $x = 0$. A closer inspection confirms that the two parts of the hyperboloid lie one both sides of the plane: intersect the hyperboloid with the line $y = z = 0$ to see two points on both sides.

(xv) $\alpha x^2 + \beta y^2 - z = 0$. An elliptical paraboloid. See Fig. 8.

(xvi) $\alpha x^2 - \beta y^2 - z = 0$. A hyperbolic paraboloid. See Fig. 9.

As in the case of the hyperboloid of one sheet, there are two generators passing through each point of this surface, one from each of the following two families of lines:

$$\lambda(\sqrt{\alpha}\,x - \sqrt{\beta})\,y = z; \qquad \sqrt{\alpha}\,x + \sqrt{\beta}\,y = \lambda.$$
$$\mu(\sqrt{\alpha}\,x + \sqrt{\beta})\,y = z; \qquad \sqrt{\alpha}\,x - \sqrt{\beta}\,y = \mu.$$

3.8 Unitary, hermitian and normal matrices

The results in Subsections 3.1-3.6 apply vector spaces over the real numbers $\mathbb{R}$. A naive reformulation of some of the results to complex numbers fails. For instance, the vector $v = (i, 1)^{\mathrm{T}} \in \mathbb{C}^2$ is *isotropic*, i.e. it has $v^{\mathrm{T}} v = i^2 + 1^2 = 0$, which creates various difficulties.

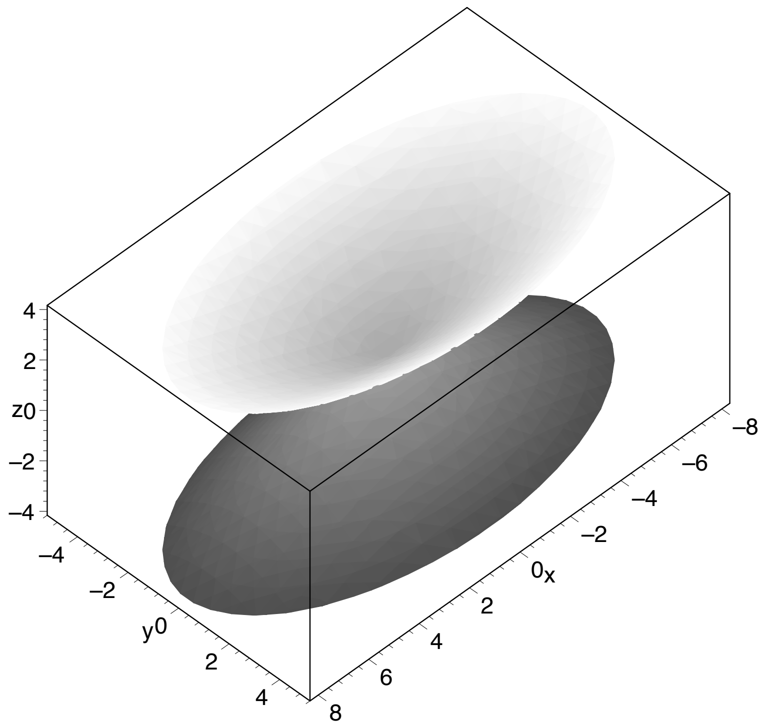


Figure 4: $x^2/4 + y^2 - z^2 = 0$

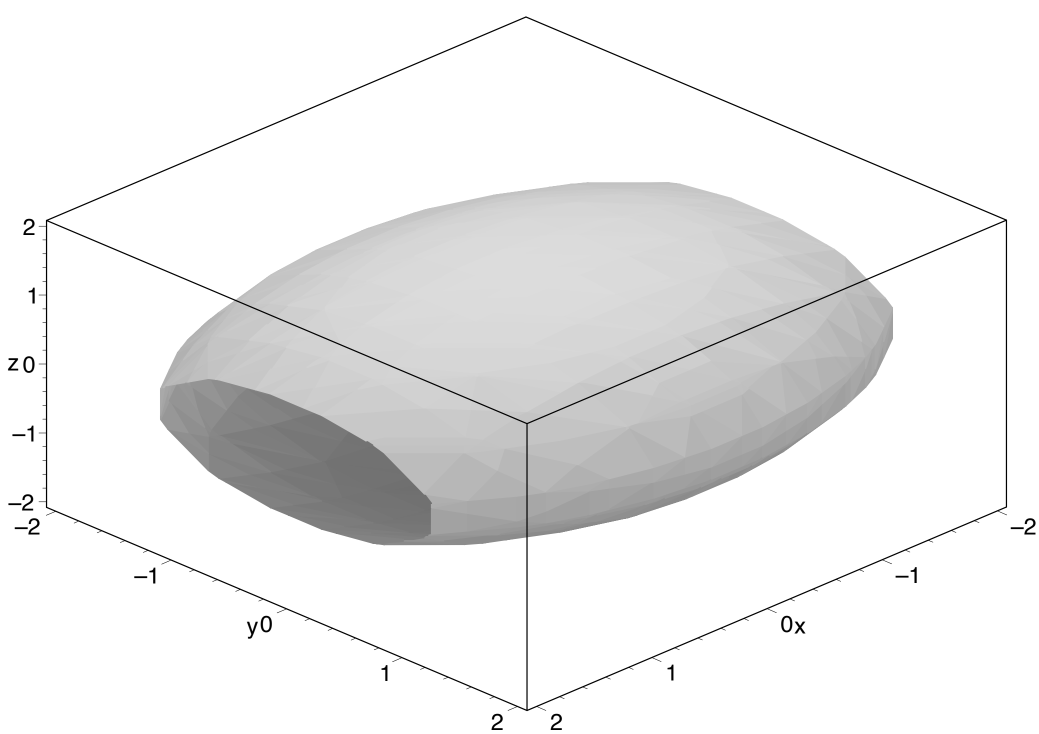


Figure 5: $x^2 + 2y^2 + 4z^2 = 7$

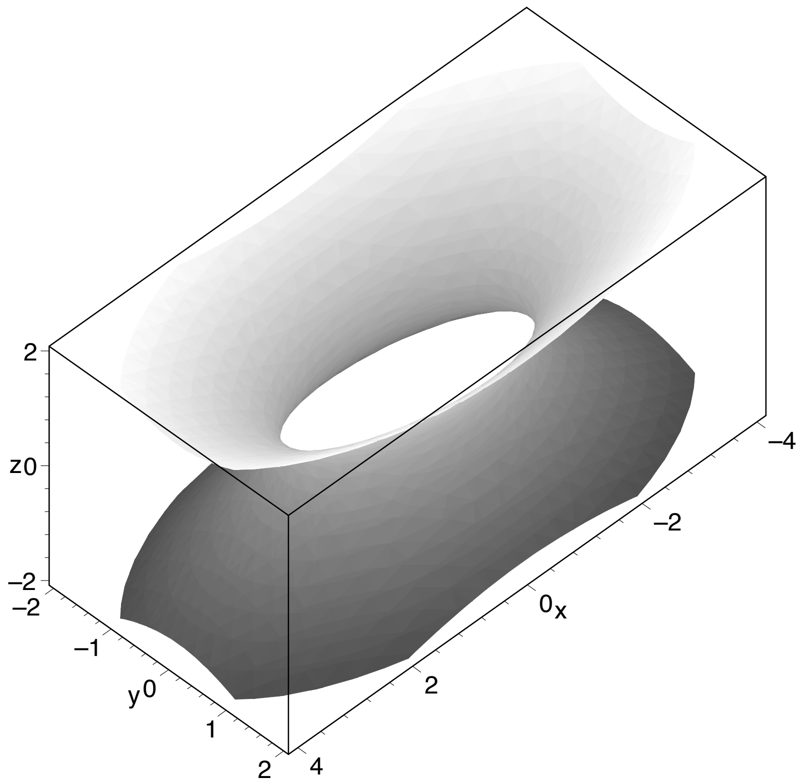


Figure 6: $x^2/4 + y^2 - z^2 = 1$

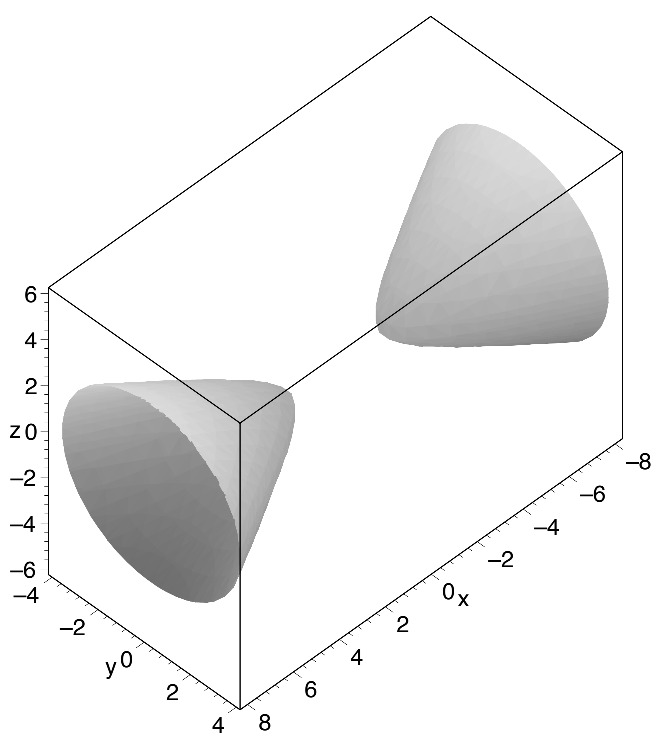


Figure 7: $x^2/4 - y^2 - z^2 = 1$

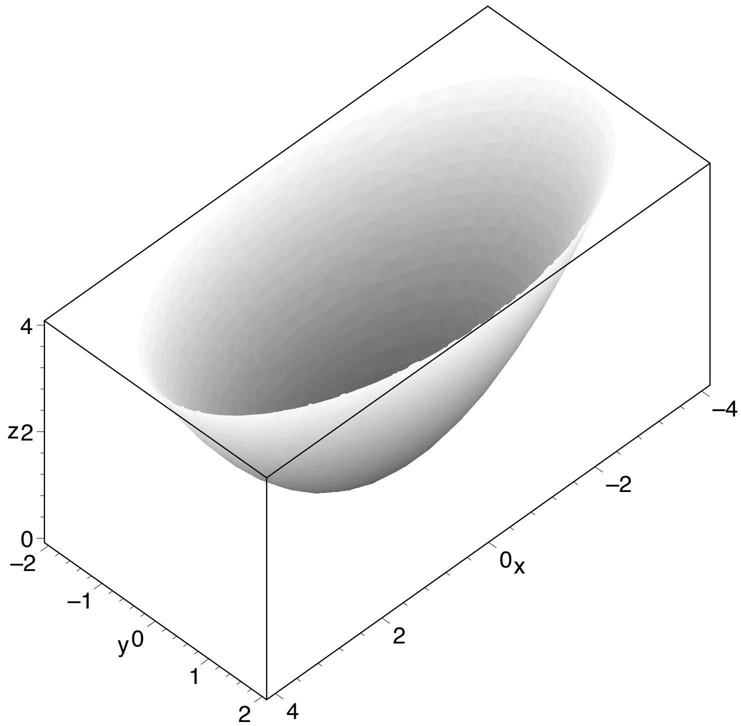


Figure 8: $z = x^2/2 + y^2$

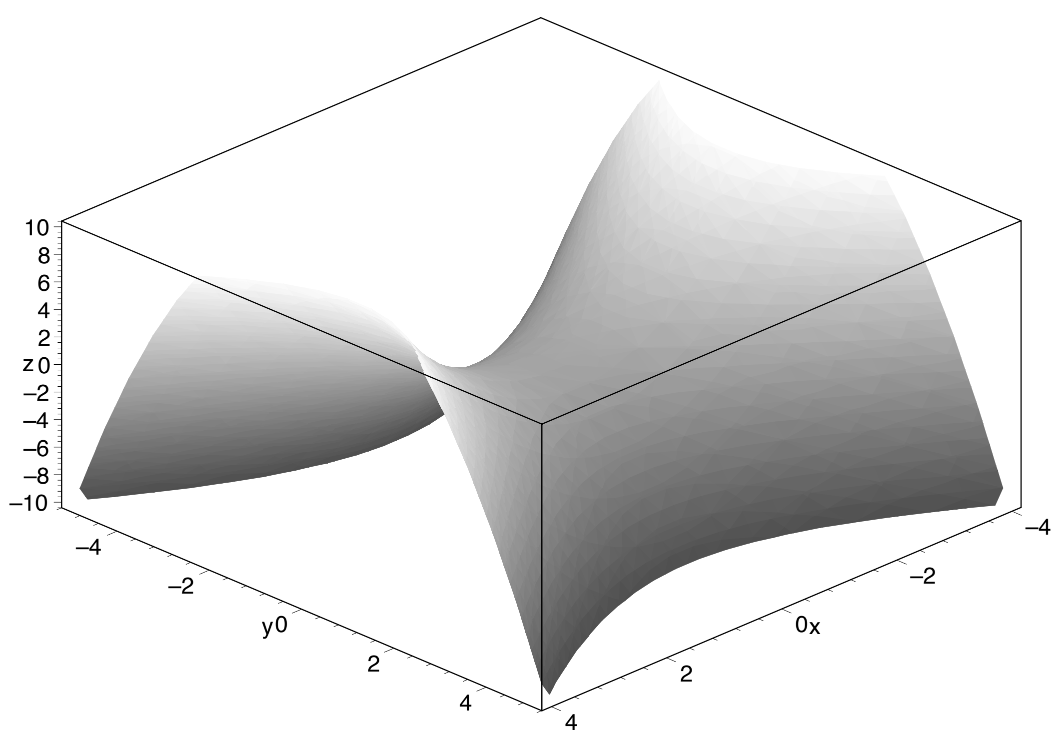


Figure 9: $z = x^2 - y^2$

To build the theory for vector spaces over $\mathbb{C}$ we have to replace bilinear forms with *sesquilinear*[16] form, that is, maps $\tau : W \times V \to \mathbb{C}$ such that

(i) $\tau(\alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2, \mathbf{v}) = \bar{\alpha}_1\tau(\mathbf{w}_1, \mathbf{v}) + \bar{\alpha}_2\tau(\mathbf{w}_2, \mathbf{v})$ and

(ii) $\tau(\mathbf{w}, \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1\tau(\mathbf{w}, \mathbf{v}_1) + \alpha_2\tau(\mathbf{w}, \mathbf{v}_2)$

for all $\mathbf{w}, \mathbf{w}_1, \mathbf{w}_2 \in W$, $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in V$, and $\alpha_1, \alpha_2 \in \mathbb{C}$. As usual, $\bar{z}$ denotes the complex conjugate of z.

Sesquilinear forms can be represented by matrices $A \in \mathbb{C}^{m,n}$ as in Subsection 3.1. Let $A^* = \bar{A}^{\mathrm{T}}$ be the conjugate matrix of A, that is, $(\alpha_{ij})^* = \overline{\alpha_{ji}}$. The representation comes by choosing a basis and writing $\alpha_{ij} = \tau(\mathbf{f}_i, \mathbf{e}_j)$. Similarly to equation (2.1), we get

$$\tau(\mathbf{w}, \mathbf{v}) = \sum_{i=1}^{m}\sum_{j=1}^{n} \bar{y}_i\, \tau(\mathbf{f}_i, \mathbf{e}_j)\, x_j = \sum_{i=1}^{m}\sum_{j=1}^{n} \bar{y}_i\, \alpha_{ij}\, x_j = \underline{\mathbf{w}}^* A \underline{\mathbf{v}}$$

For instance, the standard inner product on $\mathbb{C}^n$ becomes $\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^*\mathbf{w}$ rather than $\mathbf{v}^{\mathrm{T}}\mathbf{w}$. Note that, for $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v}^* = \mathbf{v}^{\mathrm{T}}$, so this definition is compatible with the one for real vectors. The length $|\mathbf{v}|$ of a vector is given by $|\mathbf{v}|^2 = \mathbf{v} \cdot \mathbf{v} = \mathbf{v}^*\mathbf{v}$, which is always a non-negative real number.

We are going to formulate several propositions that generalise results of the previous sections to hermitian matrices. The proofs are very similar and left for you to fill them up as an exercise. The first two propositions are generalisation of Theorems 3.1 and 3.2.

Proposition 3.16 *Let A be the matrix of the sesquilinear map $\tau : W \times V \to \mathbb{C}$ with respect to the bases $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and $\mathbf{f}_1, \ldots, \mathbf{f}_m$ of V and W, and let B be its matrix with respect to the bases $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ and $\mathbf{f}'_1, \ldots, \mathbf{f}'_m$ of V and W. If P and Q are the basis change matrices then $B = Q^*AP$.*

Proposition 3.17 *Let A be the matrix of the sesquilinear form τ on V with respect to the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of V, and let B be its matrix with respect to the basis $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ of V. If P is the basis change matrix then $B = P^*AP$.*

Definition. A matrix $A \in \mathbb{C}^{n,n}$ is called *hermitian* if $A = A^*$. A sesquilinear form τ on V is called *hermitian* if $\tau(\mathbf{w}, \mathbf{v}) = \overline{\tau(\mathbf{v}, \mathbf{w})}$ for all $\mathbf{v}, \mathbf{w} \in V$.

These are the complex analogues of symmetric matrices and symmetric bilinear forms. The following proposition is an analogue of Proposition 3.3.

Proposition 3.18 *A sesquilinear form τ is hermitian if and only if its matrix is hermitian.*

Hermitian matrices A and B are *congruent* if there exists an invertible matrix P with $B = P^*AP$. *A Hermitian quadratic form* is a function $q : V \to \mathbb{C}$ given by $q(\mathbf{v}) = \tau(\mathbf{v}, \mathbf{v})$ for some sesquilinear form τ. The following is a hermitian version of Sylvester's Theorem (Proposition 3.7) and Inertia Law (Theorem 3.8) together.

Proposition 3.19 *A hermitian quadratic form q has the form $q(\mathbf{v}) = \sum_{i=1}^{t} |x_i|^2 - \sum_{i=1}^{u} |x_{t+i}|^2$ with respect to a suitable basis, where $t + u = \mathrm{rank}(q)$.*

Equivalently, given a hermitian matrix $A \in \mathbb{C}^{n,n}$, there is an invertible matrix $P \in \mathbb{C}^{n,n}$ such that $P^*AP = B$, where $B = (\beta_{ij})$ is a diagonal matrix with $\beta_{ii} = 1$ for $1 \le i \le t$, $\beta_{ii} = -1$ for $t + 1 \le i \le t + u$, and $\beta_{ii} = 0$ for $t + u + 1 \le i \le n$, and $t + u = \mathrm{rank}(A)$.

The numbers t and u are uniquely determined by q (or A).

[16]from Latin one and a half

Similarly to the real case, the difference $t-u$ is called the signature of q (or A). We say that a hermitian matrix (or a hermitian form) is positive definite if its signature is equal to the dimension of the space. By Proposition 3.19 a positive definite hermitian form looks like the standard inner product on $\mathbb{C}^n$ in some choice of a basis. *A hermitian vector space* is a vector space over $\mathbb{C}$ equipped with a hermitian positive definite form.

Definition. A linear operator $T : V \to V$ on a hermitian vector space (V, τ) is said to be *unitary* if it preserves the form τ. That is, if $\tau(T(\mathbf{v}), T(\mathbf{w})) = \tau(\mathbf{v}, \mathbf{w})$ for all $\mathbf{v}, \mathbf{w} \in V$.

Definition. A matrix $A \in \mathbb{C}^{n,n}$ is called *unitary* if $A^*A = I_n$.

A basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of a hermitian space (V, τ) is orthonormal if $\tau(\mathbf{e}_i, \mathbf{e}_i) = 1$ and $\tau(\mathbf{e}_i, \mathbf{e}_j) = 0$ for all $i \neq j$. The following is an analogue of Proposition 3.9.

Proposition 3.20 *A linear operator $T : V \to V$ is unitary if and only if its matrix A with respect to an orthonormal basis is unitary.*

The Gram-Schmidt process works perfectly well in hermitian setting, so Theorem 3.12 turns into the following statement.

Proposition 3.21 *Let (V, τ) be a hermitian space of dimension n, and suppose that, for some r with $0 \leq r \leq n$, $\mathbf{f}_1, \ldots, \mathbf{f}_r$ are vectors in V that satisfy $\tau(\mathbf{e}_i, \mathbf{e}_i) = \delta_{i,j}$ for $1 \leq i, j \leq r$. Then $\mathbf{f}_1, \ldots, \mathbf{f}_r$ can be extended to an orthonormal basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$ of V.*

Proposition 3.21 ensures existence of orthonormal bases in hermitian spaces. Proposition 3.13 and Theorem 3.14 have analogues as well.

Proposition 3.22 *Let A be a complex hermitian matrix. Then A has an eigenvalue in $\mathbb{R}$, and all complex eigenvalues of A lie in $\mathbb{R}$.*

Proposition 3.23 *Let q be a hermitian quadratic form defined on a hermitian space V. Then there is an orthonormal basis $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ of V such that $q(\mathbf{v}) = \sum_{i=1}^{n} \alpha_i |x'_i|^2$, where x'_i are the coordinates of $\mathbf{v}$ with respect to $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$. Furthermore, the numbers α_i are real and uniquely determined by q.*

Equivalently, given any hermitian matrix A, there is a unitary matrix P such that P^*AP is a real diagonal matrix.

Notice the crucial difference between Theorem 3.14 and Proposition 3.23. In the former we start with a real matrix to end up with a real diagonal matrix. In the latter we start with a complex matrix but still we end up with a real diagonal matrix. The point is that Theorem 3.14 admits a useful generalisation to a wider class of matrices.

Definition. A matrix $A \in \mathbb{C}^{n,n}$ is called *normal* if $AA^* = A^*A$.

In particular, all Hermitian and all unitary matrices are normal. Consequently, all real symmetric and real orthogonal matrices are normal.

Lemma 3.24 *If $A \in \mathbb{C}^{n,n}$ is normal and $P \in \mathbb{C}^{n,n}$ is unitary, then P^*AP is normal.*

PROOF: If $B = P^*AP = B$ then using $(BC)^* = C^*B^*$ we compute that

$$BB^* = (P^*AP)(P^*AP)^* = P^*APP^*A^*P = P^*AA^*P = P^*A^*AP = (P^*A^*P)(P^*AP) = B^*B$$

as required. □

The following theorem is extremely useful as it is a general criterion for diagonalisability of matrices.

Theorem 3.25 *A matrix $A \in \mathbb{C}^{n,n}$ is normal if and only if there exists a unitary matrix $P \in \mathbb{C}^{n,n}$ such that P^*AP is diagonal*[17].

Proof: The "if part" follows from Lemma 3.24 as diagonal matrices are normal.

For the "only if part" we proceed by induction on n. If $n = 1$, there is nothing to prove. Let us assume we have proved the statement for all dimensions less than n. The matrix A admits an eigenvector $v \in \mathbb{C}^n$ with eigenvalue λ. Let W be the vector subspace of all vectors x satisfying $Ax = \lambda x$. If $W = \mathbb{C}^n$ then A is a scalar matrix and we are done. Otherwise, we have a nontrivial[18] decomposition $\mathbb{C}^n = W \oplus W^\perp$ where $W^\perp = \{\mathbf{v} \in \mathbb{C}^n \mid \forall \mathbf{w} \in W\ \mathbf{v}^* \cdot \mathbf{w} = 0\}$.

Let us notice that $A^*W \subseteq W$ because $AA^*x = A^*Ax = A^*\lambda x = \lambda(A^*x)$ for any $x \in W$. It follows that $AW^\perp \subseteq W^\perp$ since $(Ay)^*x = y^*(A^*x) \in y^*W = 0$ so $(Ay)^*x = 0$ for all $x \in W$, $y \in W^\perp$. Similarly, $A^*W^\perp \subseteq W^\perp$.

Now choose orthonormal bases of W and $W^\perp$. Together they form a new orthonormal basis of $\mathbb{C}^n$. The change of basis matrix P is unitary, hence by Lemma 3.24 the matrix $P^*AP = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}$ is normal. It follows that the matrices B and C are normal of smaller size and we can use the inductive hypothesis to complete the proof. □

Theorem 3.25 is an extremely useful criterion for diagonalisability of matrices. To find P in practice, we use similar methods to those used in the real case.

Example. Let A be the matrix $A = \begin{pmatrix} 6 & 2+2i \\ 2-2i & 4 \end{pmatrix}$. Then

$$c_A(x) = (6-x)(4-x) - (2+2i)(2-2i) = x^2 - 10x + 16 = (x-2)(x-8),$$

so the eigenvalues of A are 2 and 8. (We saw in Proposition 3.22 that the eigenvalues of any Hermitian matrix are real.) The corresponding eigenvectors are $\mathbf{v}_1 = (1+i, -2)^{\mathrm{T}}$ and $\mathbf{v}_2 = (1+i, 1)^T$. We find that $|\mathbf{v}_1|^2 = \mathbf{v}_1^*\mathbf{v}_1 = 6$ and $|\mathbf{v}_2|^2 = 3$, so we divide by their lengths to get an orthonormal basis $\mathbf{v}_1/|\mathbf{v}_1|$, $\mathbf{v}_2/|\mathbf{v}_2|$ of $\mathbb{C}^2$. Then the matrix

$$P = \begin{pmatrix} \frac{1+i}{\sqrt{6}} & \frac{1+i}{\sqrt{3}} \\ \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

having this basis as columns is unitary and satisfies $P^*AP = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$.

3.9 Applications to quantum mechanics (non-examinable)

With all the linear algebra we know it is a little step aside to understand basics of quantum mechanics. We discuss Schrödinger's picture[19] of quantum mechanics and derive (mathematically) Heisenberg's uncertainty principle.

The main ingredient of quantum mechanics is a hermitian vector space $(V, <,>)$. There are physical arguments showing that real Euclidean vector spaces are no good and that V must be infinite-dimensional. Here we just take their conclusions at face value. The states of the system are lines in V. We denote by $[\mathbf{v}]$ the line $\mathbb{C}\mathbf{v}$ spanned by $\mathbf{v} \in V$. We use *normalised* vectors, i.e., $\mathbf{v}$ such that $< \mathbf{v}, \mathbf{v} >= 1$ to present states as this makes formulae slightly easier.

[17] with complex entries

[18] i.e., neither W nor $W^\perp$ is zero.

[19] The alternative is Heisenberg's picture but we have no time to discuss it here.

It is impossible to observe the state of the quantum system but we can try to observe some physical quantities such as momentum, energy, spin, etc. Such physical quantities become *observables*, i.e., hermitian linear operators $\Phi : V \to V$. Hermitian in this context means that $< x, \Phi y >=< \Phi x, y >$ for all $x, y \in V$. Sweeping a subtle mathematical point under the carpet[20], we assume that Φ is diagonalisable with eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots$ and eigenvalues $\phi_1, \phi_2, \ldots$. Proof of Proposition 3.22 goes through in the infinite dimensional case, so we conclude that all ϕ_i belong to $\mathbb{R}$. Back to Physics, if we measure Φ on a state $[\mathbf{v}]$ with normalised $\mathbf{v} = \sum_n \alpha_n \mathbf{e}_n$ then the measurement will return ϕ_n as a result with probability $|\alpha_n|^2$

One observable is energy $H : V \to V$, often called hamiltonian. It is central to the theory because it determines the time evolution $[\mathbf{v}(t)]$ of the system by Schrödinger's equation:

$$\frac{d\mathbf{v}(t)}{dt} = \frac{1}{i\hbar} H\mathbf{v}(t)$$

where $\hbar \approx 10^{-34}$ Joule per second[21] is the reduced Planck constant. We know how to solve this equation: $\mathbf{v}(t) = e^{tH/i\hbar}\mathbf{v}(0)$.

As a concrete example, let us look at the quantum oscillator. The full energy of the classical harmonic oscillator mass m and frequency ω is

$$h = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2$$

where x is the position and $p = mx'$ is the momentum. To quantise it, we have to play with this expression. The vector subspace of the space of all smooth functions $C^\infty(\mathbb{R}, \mathbb{C})$ admits a convenient subspace $V = \{f(x)e^{-x^2/2} \mid f(x) \in \mathbb{C}[x]\}$, which we make hermitian by $< \phi(x), \psi(x) >= \int_{-\infty}^{\infty} \bar{\phi}(x)\psi(x)dx$. Quantum momentum and quantum position are linear operators (observables) on this space:

$$P(f(x)) = -i\hbar f'(x), \quad X(f(x)) = f(x) \cdot x.$$

The quantum Hamiltonian is a second order differential operator operator given by the same equation

$$H = \frac{P^2}{2m} + \frac{1}{2} m\omega^2 X^2 = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2} m\omega^2 x^2.$$

As mathematicians, we can assume that $m = 1$ and $\omega = 1$, so that $H(f) = (fx^2 - f'')/2$. The eigenvectors of H are Hermite functions

$$\Psi_n(x) = (-1)^n e^{x^2/2} (e^{-x^2})^{(n)}\ ,\ n = 0, 1, 2 \ldots$$

with eigenvalues $n + 1/2$ which are discrete *energy levels* of the quantum oscillator. Notice that $< \Psi_k, \Psi_n >= \delta_{k,n} 2^n n! \sqrt{\pi}$, so they are orthogonal but not orthonormal. The states $[\Psi_n]$ are *pure* states: they do not change with time and always give $n + 1/2$ as energy. If we take a system in a state $[\mathbf{v}]$ where

$$\mathbf{v} = \sum_n \alpha_n \frac{1}{\pi^4 2^{n/2} n!} \Psi_n$$

[20]If V were finite-dimensional we could have used Proposition 3.23. But V is infinite dimensional! To ensure diagonalisability V must be complete with respect to the hermitian norm. Such spaces are called Hilbert spaces. Diagonalisability is still subtle as eigenvectors do not span the whole V but only a dense subspace. Furthermore, if V admits no dense countably dimensional subspace, further difficulties arise. . . Pandora box of functional analysis is wide open, so let us try to keep it shut.

[21]Notice the physical dimensions: H is energy, t is time, i dimensionless, $\hbar$ equalises the dimensions in the both sides irrespectively of what $\mathbf{v}$ is.

is normalised then the measurement of energy will return $n + 1/2$ with probability $|\alpha_n|^2$ Notice that the measurement breaks the system!! It changes it to the state $[\Psi_n]$ and all future measurements will return the same energy!

Alternatively, it is possible to model the quantum oscillator on the vector space $W = \mathbb{C}[x]$ of polynomials. One has to use the natural linear bijection

$$\alpha : W \to V,\ \alpha(f(x)) = f(x)e^{-x^2/2}$$

and transfer all the formulae to W. The metric becomes $< f, g >=< \alpha(f), \alpha(g) >= \int_{-\infty}^{\infty} \bar{f}(x)g(x)e^{-x^2}dx$, the formulae for P and X changes accordingly, and at the end one arrives at Hermite polynomials $\alpha^{-1}(\Psi_n(x)) = (-1)^n e^{x^2}(e^{-x^2})^{(n)}$ instead of Hermite functions.

Let us go back to an abstract system with two observables P and Q. It is pointless to measure Q after measuring P as the system is broken. But can we measure them simultaneously? The answer is given by Heisenberg's uncertainty principle. Mathematically, it is a corollary of Schwarz's inequality:

$$||\mathbf{v}||^2 \cdot ||\mathbf{w}||^2 \ =< \mathbf{v}, \mathbf{v} >< \mathbf{w}, \mathbf{w} > \geq \ | < \mathbf{v}, \mathbf{w} > |^2.$$

Let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots$ be eigenvectors for P an let $p_1, p_2, \ldots$ be the corresponding eigenvalues. The probability that p_j is returned after measuring on $[\mathbf{v}]$ with $\mathbf{v} = \sum_n \alpha_n \mathbf{e}_n$ depends on the multiplicity of the eigenvalue:

$$\text{Prob}(p_j \text{ is returned}) = \sum_{p_k = p_j} |\alpha_k|^2.$$

Hence, we should have the expected value

$$\mathcal{E}(P, \mathbf{v}) = \sum_k p_k|\alpha_k|^2 = \sum_k < \alpha_k \mathbf{e}_k, p_k \alpha_k \mathbf{e}_k >=< \mathbf{v}, P(\mathbf{v}) > .$$

To compute the expected quadratic error we use the shifted observable $P_\mathbf{v} = P - \mathcal{E}(P, \mathbf{v})I$:

$$\mathcal{D}(P, \mathbf{v}) = \sqrt{\mathcal{E}({P_\mathbf{v}}^2, \mathbf{v})} = \sqrt{< \mathbf{v}, P_\mathbf{v}(P_\mathbf{v}(\mathbf{v})) >} = \sqrt{< P_\mathbf{v}(\mathbf{v}), P_\mathbf{v}(\mathbf{v}) >} = ||P_\mathbf{v}(\mathbf{v})||$$

where we use the fact that P and $P_\mathbf{v}$ are hermitian. Notice that $\mathcal{D}(P, \mathbf{v})$ has a physical meaning of uncertainty of measurement of P. Notice also that the operator $PQ - QP$ is no longer hermitian in general but we can still talk about its expected value. Here goes Heisenberg's principle.

Theorem 3.26

$$\mathcal{D}(P, \mathbf{v}) \cdot \mathcal{D}(Q, \mathbf{v}) \ \geq \ \frac{1}{2}|\mathcal{E}(PQ - QP, \mathbf{v})|$$

PROOF: In the right hand side, $\mathcal{E}(PQ - QP, \mathbf{v}) = \mathcal{E}(P_\mathbf{v}Q_\mathbf{v} - Q_\mathbf{v}P_\mathbf{v}, \mathbf{v}) =< \mathbf{v}, P_\mathbf{v}Q_\mathbf{v}(\mathbf{v}) > - < \mathbf{v}, Q_\mathbf{v}P_\mathbf{v}(\mathbf{v}) >=< P_\mathbf{v}(\mathbf{v}), Q_\mathbf{v}(\mathbf{v}) > - < Q_\mathbf{v}(\mathbf{v}), P_\mathbf{v}(\mathbf{v}) >$. Remembering that the form is hermitian,

$$\mathcal{E}(PQ - QP, \mathbf{v}) =< P_\mathbf{v}(\mathbf{v}), Q_\mathbf{v}(\mathbf{v}) > - \overline{< P_\mathbf{v}(\mathbf{v}), Q_\mathbf{v}(\mathbf{v}) >} = 2 \cdot \text{Im}(< P_\mathbf{v}(\mathbf{v}), Q_\mathbf{v}(\mathbf{v}) >) \ ,$$

twice the imaginary part. So the right hand side is estimated by Schwarz's inequality:

$$\text{Im}(< P_\mathbf{v}(\mathbf{v}), Q_\mathbf{v}(\mathbf{v}) >) \ \leq \ | < P_\mathbf{v}(\mathbf{v}), Q_\mathbf{v}(\mathbf{v}) > | \ \leq \ ||P_\mathbf{v}(\mathbf{v})|| \cdot ||Q_\mathbf{v}(\mathbf{v})|| \ .$$

□

Two cases of particular physical interest are *commuting observables*, i.e. $PQ = QP$ and *conjugate observables*, i.e. $PQ - QP = i\hbar I$. Commuting observable can be measured simultaneously with any degree of certainty. Conjugate observables obey Heisenberg's uncertainty:

$$\mathcal{D}(P, \mathbf{v}) \cdot \mathcal{D}(Q, \mathbf{v}) \ \geq \ \frac{\hbar}{2}.$$

4 Finitely Generated Abelian Groups

4.1 Definitions

Groups were introduced in the first year in *Foundations*, and will be studied in detail next term in *Algebra II: Groups and Rings*. In this course, we are only interested in abelian (= commutative) groups, which are defined as follows.

Definition. An abelian group is a set G together with a binary operation, which we write as addition, and which satisfies the following properties:

(i) (*Closure*) for all $g, h \in G$, $g + h \in G$;
(ii) (*Associativity*) for all $g, h, k \in G$, $(g + h) + k = g + (h + k)$;
(iii) there exists an element $0_G \in G$ such that:
  (a) (*Identity*) for all $g \in G$, $g + 0_G = g$; and
  (b) (*Inverse*) for all $g \in G$ there exists $-g \in G$ such that $g + (-g) = 0_G$;
(iv) (*Commutativity*) for all $g, h \in G$, $g + h = h + g$.

Usually we just write 0 rather than 0_G. We only write 0_G if we need to distinguish between the zero elements of different groups.

The commutativity axiom (iv) is not part of the definition of a general group, and for general (non-abelian) groups, it is more usual to use multiplicative rather than additive notation. All groups in this course should be assumed to be abelian, although many of the definitions in this section apply equally well to general groups.

Examples. 1. The integers $\mathbb{Z}$.

2. Fix a positive integer $n > 0$ and let

$$\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\} = \{ x \in \mathbb{Z} \mid 0 \le x < n \}.$$

where addition is computed modulo n. So, for example, when $n = 9$, we have $2 + 5 = 7$, $3 + 8 = 2$, $6 + 7 = 4$, etc. Note that the inverse $-x$ of $x \in \mathbb{Z}_n$ is equal to $n - x$ (if $x \neq 0$) in this example.

3. Examples from linear algebra. Let K be a field.
(i) The elements of K form an abelian group under addition.
(ii) The non-zero elements of K form an abelian group $K^\times$ under multiplication.
(iii) The vectors in any vector space form an abelian group under addition.

Proposition 4.1 *(The cancellation law) Let G be any group, and let $g, h, k \in G$. Then $g + h = g + k \Rightarrow h = k$.*

Proof: Add $-g$ to both sides of the equation and use the Associativity and Identity axioms. □

For any group G, $g \in G$, and integer $n > 0$, we define ng to be $g + g + \cdots + g$, with n occurrences of g in the sum. So, for example, $1g = g$, $2g = g + g$, $3g = g + g + g$, etc. We extend this notation to all $n \in \mathbb{Z}$ by defining $0g = 0$ and $(-n)g = -(ng)$ for $-n < 0$. Overall, this defines a scalar action $\mathbb{Z} \times G \to G$ which allows as to think of abelian groups as "vector spaces over $\mathbb{Z}$" (or using precise terminology $\mathbb{Z}$-modules - algebraic modules will play a significant role in *Rings and Modules* in year 3).

Definition. A group G is called *cyclic* if there exists an element $x \in G$ such that every element of G is of the form mx for some $m \in \mathbb{Z}$.

The element x in the definition is called a *generator* of G. Note that $\mathbb{Z}$ and $\mathbb{Z}_n$ are cyclic with generator $x = 1$.

Definition. A bijection $\phi : G \to H$ between two (abelian) groups is called an *isomorphism* if $\phi(g + h) = \phi(g) + \phi(h)$ for all $g, h \in G$, and the groups G and H are called *isomorphic* is there is an isomorphism between them.

The notation $G \cong H$ means that G is isomorphic to H; isomorphic groups are often thought of as being essentially the same group, but with elements having different names.

Note (exercise) that any isomorphism must satisfy $\phi(0_G) = 0_H$ and $\phi(-g) = -\phi(g)$ for all $g \in G$.

Proposition 4.2 *Any cyclic group G is isomorphic either to $\mathbb{Z}$ or to $\mathbb{Z}_n$ for some $n > 0$.*

PROOF: Let G be cyclic with generator x. So $G = \{ mx \mid m \in \mathbb{Z} \}$. Suppose first that the elements mx for $m \in \mathbb{Z}$ are all distinct. Then the map $\phi : \mathbb{Z} \to G$ defined by $\phi(m) = mx$ is a bijection, and it is clearly an isomorphism.

Otherwise, we have $lx = mx$ for some $l < m$, and so $(m-l)x = 0$ with $m - l > 0$. Let n be the least integer with $n > 0$ and $nx = 0$. Then the elements $0x = 0, 1x, 2x, \ldots, (n-1)x$ of G are all distinct, because otherwise we could find a smaller n. Furthermore, for any $mx \in G$, we can write $m = rn + s$ for some $r, s \in \mathbb{Z}$ with $0 \leq s < n$. Then $mx = (rn + s)x = sx$, so $G = \{ 0, 1x, 2x, \ldots, (n-1)x \}$, and the map $\phi : \mathbb{Z}_n \to G$ defined by $\phi(m) = mx$ for $0 \leq m < n$ is a bijection, which is easily seen to be an isomorphism. □

Definition. For an element $g \in G$, the least integer $n > 0$ with $ng = 0$, if it exists, is called the *order* $|g|$ of g. If there is no such n, then g has infinite order and we write $|g| = \infty$.

Exercise. If $\phi : G \to H$ is an isomorphism, then $|g| = |\phi(g)|$ for all $g \in G$.

Definition. A group G is *generated* or *spanned* by a subset X of G if every $g \in G$ can be written as a finite sum $\sum_{i=1}^{k} m_i x_i$, with $m_i \in \mathbb{Z}$ and $x_i \in X$. It is finitely generated if it has a finite generating set $X = \{x_1, \ldots, x_n\}$.

So a group is cyclic if and only if it has a generating set X with $|X| = 1$.

In general, if G is generated by X, then we write $G = \langle X \rangle$ or $G = \langle x_1, \ldots, x_n \rangle$ when $X = \{x_1, \ldots, x_n\}$ is finite.

Definition. The direct sum of groups $G_1, \ldots, G_n$ is defined to be the set $\{ (g_1, g_2, \ldots, g_n) \mid g_i \in G_i \}$ with component-wise addition

$$(g_1, g_2, \ldots, g_n) + (h_1, h_2, \ldots, h_n) = (g_1 + h_1, g_2 + h_2, \ldots, g_n + h_n).$$

This is a group with identity element $(0, 0, \ldots, 0)$ and $-(g_1, g_2, \ldots, g_n) = (-g_1, -g_2, \ldots, -g_n)$.

In general (non-abelian) group theory this is more often known as the direct product of groups.

The main result of this section, known as the *fundamental theorem of finitely generated abelian groups*, is that every finitely generated abelian group is isomorphic to a direct sum of cyclic groups. (This is not true in general for abelian groups, such as the additive group $\mathbb{Q}$ of rational numbers, which are not finitely generated.)

4.2 Subgroups, cosets and quotient groups

Definition. A subset H of a group G is called a *subgroup* of G if it forms a group under the same operation as that of G.

Lemma 4.3 *If H is a subgroup of G, then the identity element 0_H of H is equal to the identity element 0_G of G.*

Proof: Using the identity axioms for H and G, $0_H + 0_H = 0_H = 0_H + 0_G$. Now by the cancellation law, $0_H = 0_G$. □

The definition of a subgroup is *semantic* in its nature. While it precisely pinpoints what a subgroup is, it is quite cumbersome to use. The following proposition gives a usable criterion.

Proposition 4.4 *Let H be a subset of a group G. The following statements are equivalent.*

(i) *H is a subgroup of G.*
(ii) *(a)* *H is nonempty; and*
(b) *$h_1, h_2 \in H \Rightarrow h_1 + h_2 \in H$; and*
(c) *$h \in H \Rightarrow -h \in H$.*
(iii) *(a)* *H is nonempty; and*
(b) *$h_1, h_2 \in H \Rightarrow h_1 - h_2 \in H$.*

Proof: If H is a subgroup of G then it is nonempty as it contains 0_H. Moreover, $h_1 - h_2 = h_1 + (-h_2) \in H$ if so are h_1 and h_2. Thus, (i) implies (iii).

To show that (iii) implies (ii) we pick $x \in H$. Then $0 = x - x \in H$. Now $-h = 0 - h \in H$ for any $h \in H$. Finally, $h_1 + h_2 = h_1 - (-h_2) \in H$ for all $h_1, h_2 \in H$.

To show that (ii) implies (i) we need to verify the four group axioms in H. Two of these, 'Closure', and 'Inverse', are the conditions (b) and (c). The other two axioms are 'Associativity' and 'Identity'. Associativity holds because it holds in G, and H is a subset of G. Since we are assuming that H is nonempty, there exists $h \in H$, and then $-h \in H$ by (c), and $h + (-h) = 0 \in H$ by (b), and so 'Identity' holds, and H is a subgroup. □

Examples. 1. There are two standard subgroups of any group G: the whole group G itself, and the *trivial* subgroup $\{0\}$ consisting of the identity alone. Subgroups other than G are called *proper* subgroups, and subgroups other than $\{0\}$ are called *non-trivial* subgroups.

2. If g is any element of any group G, then the set of all integer multiples $\{ mg \mid m \in \mathbb{Z} \}$ forms a subgroup of G called the cyclic subgroup generated by g.

Let us look at a few specific examples. If $G = \mathbb{Z}$, then $5\mathbb{Z}$, which consists of all multiples of 5, is the cyclic subgroup generated by 5. Of course, we can replace 5 by any integer here, but note that the cyclic groups generated by 5 and -5 are the same.

If $G = \langle g \rangle$ is a finite cyclic group of order n and m is a positive integer dividing n, then the cyclic subgroup generated by mg has order n/m and consists of the elements kmg for $0 \leq k < n/m$.

Exercise. What is the order or the cyclic subgroup generated by mg for general m (where we drop the assumption that $m|n$)?

Exercise. Show that the groups of non-zero complex numbers $\mathbb{C}^\times$ under the operation of multiplication has finite cyclic subgroups of all possible orders.

Definition. Let $g \in G$. Then the *coset* $H + g$ is the subset $\{ h + g \mid h \in H \}$ of G.

(*Note*: Since our groups are abelian, we have $H + g = g + H$, but in general group theory the right and left cosets Hg and gH can be different.)

Examples. 3. $G = \mathbb{Z}$, $H = 5\mathbb{Z}$. There are just 5 distinct cosets $H = H + 0 = \{ 5n \mid n \in \mathbb{Z} \}$, $H + 1 = \{ 5n + 1 \mid n \in \mathbb{Z} \}$, $H + 2$, $H + 3$, $H + 4$. Note that $H + i = H + j$ whenever $i \equiv j$

(mod 5).

4. $G = \mathbb{Z}_6$, $H = \{0, 3\}$. There are 3 distinct cosets, $H = H + 3 = \{0, 3\}$, $H + 1 = H + 4 = \{1, 4\}$, and $H + 2 = H + 5 = \{2, 5\}$,

5. $G = \mathbb{C}^\times$, the group of non-zero complex numbers under multiplication, $S^1 = \{z,\ |z| = 1\}$, the unit circle. The cosets are circles. There are uncountably many distinct cosets, one for each positive real number (radius of a circle).

Proposition 4.5 *The following are equivalent for $g, k \in G$:*

(i) $k \in H + g$;

(ii) $H + g = H + k$;

(iii) $k - g \in H$.

PROOF: Clearly $H + g = H + k \Rightarrow k \in H + g$, so $(ii) \Rightarrow (i)$.

If $k \in H + g$, then $k = h + g$ for some fixed $h \in H$, so $g = k - h$. Let $f \in H + g$. Then, for some $h_1 \in H$, we have $f = h_1 + g = h_1 + k - h \in H + k$, so $Hg \subseteq Hk$. Similarly, if $f \in H + k$, then for some $h_1 \in H$, we have $f = h_1 + k = h_1 + h + g \in H + g$, so $H + k \subseteq H + g$. Thus $H + g = H + k$, and we have proved that $(i) \Rightarrow (ii)$.

If $k \in H + g$, then, as above, $k = h + g$, so $k - g = h \in H$ and $(i) \Rightarrow (iii)$.

Finally, if $k - g \in H$, then putting $h = k - g$, we have $h + g = k$, so $k \in H + g$, proving $(iii) \Rightarrow (i)$. □

Corollary 4.6 *Two right cosets $H + g_1$ and $H + g_2$ of H in G are either equal or disjoint.*

PROOF: If $H + g_1$ and $H + g_2$ are not disjoint, then there exists an element $k \in (H + g_1) \cap (H + g_2)$, but then $H + g_1 = H + k = H + g_2$ by the proposition. □

Corollary 4.7 *The cosets of H in G partition G.*

Proposition 4.8 *If H is finite, then all right cosets have exactly $|H|$ elements.*

PROOF: Since $h_1 + g = h_2 + g \Rightarrow h_1 = h_2$ by the cancellation law, it follows that the map $\phi : H \to H + g$ defined by $\phi(h) = h + g$ is a bijection, and the result follows. □

Corollary 4.7 and Proposition 4.8 together imply:

Theorem 4.9 (Lagrange's Theorem) *Let G be a finite (abelian) group and H a subgroup of G. Then the order of H divides the order of G.*

Definition. The number of distinct right cosets of H in G is called the *index* of H in G and is written as $|G : H|$.

If G is finite, then we clearly have $|G : H| = |G|/|H|$. But, from the example $G = \mathbb{Z}$, $H = 5\mathbb{Z}$ above, we see that $|G : H|$ can be finite even when G and H are infinite.

Proposition 4.10 *Let G be a finite (abelian) group. Then for any $g \in G$, the order $|g|$ of g divides the order $|G|$ of G.*

PROOF: Let $|g| = n$. We saw in Example 2 above that the integer multiples $\{\, mg \mid m \in \mathbb{Z} \,\}$ of g form a subgroup H of G. By minimality of n, the distinct elements of H are $\{0, g, 2g, \ldots, (n-1)g\}$, so $|H| = n$ and the result follows from Lagrange's Theorem. □

As an application, we can now immediately classify all finite (abelian) groups whose order is prime.

Proposition 4.11 *Let G be a (abelian) group having prime order p. Then G is cyclic; that is, $G \cong \mathbb{Z}_p$.*

PROOF: Let $g \in G$ with $0 \neq g$. Then $|g| > 1$, but $|g|$ divides p by Proposition 4.10, so $|g| = p$. But then G must consist entirely of the integer multiples mg $(0 \leq m < p)$ of g, so G is cyclic. □

Definition. If A and B are subsets of a group G, then we define their sum $A + B = \{ a + b \mid a \in A, b \in B \}$.

Lemma 4.12 *If H is a subgroup of the abelian group G and $H + g$, $H + h$ are cosets of H in G, then $(H + g) + (H + h) = H + (g + h)$.*

PROOF: Since G is abelian, this follows directly from commutativity and associativity. □

Theorem 4.13 *Let H be a subgroup of an abelian group G. Then the set G/H of cosets $H + g$ of H in G forms a group under addition of subsets.*

PROOF: We have just seen that $(H + g) + (H + h) = H + (g + h)$, so we have closure, and associativity follows easily from associativity of G. Since $(H + 0) + (H + g) = H + g$ for all $g \in G$, $H = H + 0$ is an identity element, and since $(H - g) + (H + g) = H - g + g = H$, $H - g$ is an inverse to $H + g$ for all cosets $H + g$. Thus the four group axioms are satisfied and G/H is a group. □

Definition. The group G/H is called the *quotient group* (or the *factor group*) of G by H.

Notice that if G is finite, then $|G/H| = |G : H| = |G|/|H|$. So, although the quotient group seems a rather complicated object at first sight, it is actually a smaller group than G.

Examples. 1. Let $G = \mathbb{Z}$ and $H = m\mathbb{Z}$ for some $m > 0$. Then there are exactly m distinct cosets, H, $H + 1$, …, $H + (m - 1)$. If we add together k copies of $H + 1$, then we get $H + k$. So G/H is cyclic of order m and with generator $H + 1$. So by Proposition 4.2, $\mathbb{Z}/m\mathbb{Z} \cong \mathbb{Z}_m$.

2. $G = \mathbb{R}$ and $H = \mathbb{Z}$. The quotient group G/H is isomorphic to the circle subgroup S^1 of the multiplicative group $\mathbb{C}^\times$. One writes an explicit isomorphism $\phi : G/H \to S^1$ by $\phi(x + \mathbb{Z}) = e^{2\pi x i}$.

3. $G = \mathbb{Q}$ and $H = \mathbb{Z}$. The quotient group G/H features in one of the previous exams. It has been required to show that this group is infinite, not finitely generated and that every element of G/H has finite order.

4. The quotient groups play important role in Analysis: they are used to define Lebesgue spaces. Let $p \geq 1$ be a real number, $U \subseteq \mathbb{R}$ an interval. Consider the vector space V of all measurable functions $f : U \to \mathbb{R}$ such that $\int_U |f(x)|^p dx < \infty$. It follows from Minkowski's inequality that this is a vector space and consequently an abelian group. It contains a vector subspace W of *negligible functions*, that is, $f(x)$ satisfying $\int_U |f(x)|^p dx = 0$. The quotient group V/W is actually a vector space called Lebesgue's space and denoted $L^p(U, \mathbb{R})$.

4.3 Homomorphisms and the first isomorphism theorem

Definition. Let G and H be groups. A *homomorphism* ϕ from G to H is a map $\phi : G \to H$ such that $\phi(g_1 + g_2) = \phi(g_1) + \phi(g_2)$ for all $g_1, g_2 \in G$.

Homomorphisms correspond to linear transformations between vector spaces.

Note that an isomorphism is just a bijective homomorphism. There are two other types of 'morphism' that are worth mentioning at this stage.

A homomorphism ϕ is injective if it is an injection; that is, if $\phi(g_1) = \phi(g_2) \Rightarrow g_1 = g_2$. A homomorphism ϕ is surjective if it is a surjection; that is, if $\text{im}(\phi) = H$. Sometimes, a surjective homomorphism is called *epimorphism* while an injective homomorphism is called *monomorphism* but we will not use this terminology in these lectures.

Lemma 4.14 *Let $\phi : G \to H$ be a homomorphism. Then $\phi(0_G) = 0_H$ and $\phi(g) = -\phi(g)$ for all $g \in G$.*

PROOF: Exercise. (Similar to results for linear transformations.) □

Example. Let G be any group, and let $n \in \mathbb{Z}$. Then $\phi : G \to G$ defined by $\phi(g) = ng$ for all $g \in G$ is a homomorphism.

Kernels and images are defined as for linear transformations of vector spaces.

Definition. Let $\phi : G \to H$ be a homomorphism. Then the *kernel* $\ker(\phi)$ of ϕ is defined to be the set of elements of G that map onto 0_H; that is,

$$\ker(\phi) = \{ g \mid g \in G,\ \phi(g) = 0_H \}.$$

Note that by Lemma 4.14 above, $\ker(\phi)$ always contains 0_G.

Proposition 4.15 *Let $\phi : G \to H$ be a homomorphism. Then ϕ is injective if and only if $\ker(\phi) = \{0_G\}$.*

PROOF: Since $0_G \in \ker(\phi)$, if ϕ is injective then we must have $\ker(\phi) = \{0_G\}$. Conversely, suppose that $\ker(\phi) = \{0_G\}$, and let $g_1, g_2 \in G$ with $\phi(g_1) = \phi(g_2)$. Then $0_H = \phi(g_1) - \phi(g_2) = \phi(g_1 - g_2)$ (by Lemma 4.14), so $g_1 - g_2 \in \ker(\phi)$ and hence $g_1 - g_2 = 0_G$ and $g_1 = g_2$. So ϕ is injective. □

Theorem 4.16 *(i) Let $\phi : G \to H$ be a homomorphism. Then $\ker(\phi)$ is a subgroup of G and $\text{im}(\phi)$ is a subgroup of H.*
(ii) Let H be a subgroup of a group G. Then the map $\phi : G \to G/H$ defined by $\phi(g) = H + g$ is a surjective homomorphism with kernel H.

PROOF: (i) is straightforward using Proposition 4.4. For (ii), it is easy to check that ϕ is surjective, and $\phi(g) = 0_{G/H} \Leftrightarrow H + g = H + 0_G \Leftrightarrow g \in H$, so $\ker(\phi) = H$. □

The following lemma explains a connection between quotients and homomorphisms. It clarifies the trickiest point in the proof of the forthcoming *First Isomorphism Theorem*.

Lemma 4.17 *Let $\phi : G \to H$ be a homomorphism with a kernel K, A a subgroup of G. The homomorphism ϕ determines a homomorphism $\overline{\phi} : G/A \to H$ via $\overline{\phi}(A + g) = \phi(g)$ for all $g \in G$ if and only if $A \subseteq K$.*

PROOF: We need to check that the map $\overline{\phi}$ is *well-defined*, i.e. whenever $A + g = A + h$ with $g \neq h$, we need to ensure that $\phi(g) = \phi(h)$. In this case $g = a + h$ for some $a \in A$. Hence, $\overline{\phi}$ is well-defined if and only if $\phi(g) = \phi(a) + \phi(h) = \phi(h)$ for all $g, h \in G$, $a \in A$ if and only if $\phi(a) = 0$ for all $a \in A$ if and only if $A \subseteq K$.

Observe also that, once $\overline{\phi}$ is well-defined, it is trivially a homomorphism, since so is ϕ:

$$\overline{\phi}(A + h) + \overline{\phi}(A + g) = \phi(h) + \phi(g) = \phi(h + g) = \overline{\phi}(A + h + g) = \overline{\phi}((A + h) + (A + g)).$$

□

If one denotes the set of all homomorphism from G to H by $\hom(G, H)$ there is an elegant way to reformulate Lemma 4.17. The composition with the quotient map $\psi : G \to G/A$ defines a bijection

$$\hom(G/A, H) \to \{\alpha \in \hom(G, H) \mid \alpha(A) = \{0\}\}, \quad \phi \mapsto \phi \circ \psi.$$

Theorem 4.18 (First Isomorphism Theorem) *Let $\phi : G \to H$ be a homomorphism with kernel K. Then $G/K \cong \mathrm{im}(\phi)$. More precisely, there is an isomorphism $\overline{\phi} : G/K \to \mathrm{im}(\phi)$ defined by $\overline{\phi}(K + g) = \phi(g)$ for all $g \in G$.*

PROOF: The map $\overline{\phi}$ is a well-defined homomorphism by Lemma 4.17. Clearly, $\mathrm{im}(\overline{\phi}) = \mathrm{im}(\phi)$. Finally,

$$\overline{\phi}(K + g) = 0_H \iff \phi(g) = 0_H \iff g \in K \iff K + g = K + 0 = 0_{G/K}.$$

By Proposition 4.15, $\overline{\phi}$ is injective. Thus $\overline{\phi} : G/K \to \mathrm{im}(\phi)$ is an isomorphism. □

One can associate two quotient groups to a homomorphism $\phi : G \to H$. The cokernel of ϕ is $\mathrm{Coker}(\phi) = H/\mathrm{im}(\phi)$ and the coimage of ϕ is $\mathrm{Coim}(\phi) = G/\ker(\phi)$. In short, the first isomorphism theorem states that the natural homomorphism from the coimage to the image is an isomorphism.

We shall be using this theorem later, when we prove the main theorem on finitely generated abelian group in Subsection 4.7. The crucial observation is that any finitely generated abelian group is the cokernel of a homomorphism between two finitely generated free abelian groups, which we will discuss in the next section.

4.4 Free abelian groups

Definition. The direct sum $\mathbb{Z}^n$ of n copies of $\mathbb{Z}$ is known as a (finitely generated) *free abelian group* of rank n.

More generally, a finitely generated abelian group is called free abelian if it is isomorphic to $\mathbb{Z}^n$ for some $n \geq 0$.

(The free abelian group $\mathbb{Z}^0$ of rank 0 is defined to be the trivial group $\{0\}$ containing the single element 0.)

The groups $\mathbb{Z}^n$ have many properties in common with vector spaces such as $\mathbb{R}^n$, but we must expect some differences, because $\mathbb{Z}$ is not a field.

We can define the standard basis of $\mathbb{Z}^n$ exactly as for $\mathbb{R}^n$; that is, $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, where $\mathbf{x}_i$ has 1 in its i-th component and 0 in the other components. This has the same properties as a basis of a vector space; i.e. it is linearly independent and spans $\mathbb{Z}^n$.

Definition. Elements $x_1, \ldots, x_n$ of an abelian group G are called *linearly independent* if, for $\alpha_1, \ldots, \alpha_n \in \mathbb{Z}$, $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0_G$ implies $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0_{\mathbb{Z}}$.

Definition. Elements $x_1, \ldots, x_n$ form a *free basis* of the abelian group G if and only if they are linearly independent and generate (span) G.

Now consider elements $x_1, \ldots, x_n$ of an abelian group G. It is possible to extend the assignment $\phi(\mathbf{x}_i) = x_i$ to a group homomorphism $\phi : \mathbb{Z}^n \to G$. As a function we define $\phi((a_1, a_2, \ldots, a_n)^T) = \sum_{i=1}^{n} a_i x_i$. We leave the proof of the following result as an exercise.

Proposition 4.19 *(i) The function ϕ is a group homomorphism.*
(ii) Elements x_i are linearly independent if and only if ϕ is injective.
(iii) Elements x_i span G if and only if ϕ is surjective.

(iv) Elements x_i form a basis of G if and only if ϕ is an isomorphism.

Note that this proposition makes a perfect sense for vector spaces. Also note that the last statement implies that $x_1, \dots, x_n$ is a free basis of G if and only if every element $g \in G$ has a unique expression $g = \alpha_1 x_1 + \cdots + \alpha_n x_n$ with $\alpha_i \in \mathbb{Z}$, very much like for vector spaces.

Before Proposition 4.19 we were trying to extend the assignment $\phi(\mathbf{x}_i) = x_i$ to a group homomorphism $\phi : \mathbb{Z}^n \to G$. Note that the extension we wrote is unique. This is the key to the next corollary. The details of the proof are left to the reader.

Corollary 4.20 (Universal property of the free abelian group) *Let G be a free group with basis $x_1, \dots, x_n$. Let H be a group and $a_1, \dots, a_n \in H$. Then there exists a unique group homomorphism $\phi : G \to H$ such that $\phi(x_i) = a_i$ for all i.*

As for finite dimensional vector spaces, it turns out that any two free bases of a free abelian group have the same size, but this has to be proved. It will follow directly from the next theorem.

Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ be the standard free basis of $\mathbb{Z}^n$, and let $\mathbf{y}_1, \dots, \mathbf{y}_m$ be another free basis. As in Linear Algebra, we can define the associated change of basis matrix P (with original basis $\{\mathbf{x}_i\}$ and new basis $\{\mathbf{y}_i\}$), where the columns of P are $\mathbf{y}_i^{\mathrm{T}}$; that is, they express $\mathbf{y}_i$ in terms of $\mathbf{x}_i$. For example, if $n = m = 2$, $\mathbf{y}_1 = (2\ 7)$, $\mathbf{y}_2 = (1\ 4)$, then $P = \begin{pmatrix} 2 & 1 \\ 7 & 4 \end{pmatrix}$. In general, $P = (\rho_{ij})$ is an $n \times m$ matrix with $\mathbf{y}_j = \sum_{i=1}^n \rho_{ij} \mathbf{x}_i$ for $1 \le j \le m$.

Theorem 4.21 *Let $\mathbf{y}_1, \dots, \mathbf{y}_m \in \mathbb{Z}^n$ with $\mathbf{y}_j = \sum_{i=1}^n \rho_{ij} \mathbf{x}_i$ for $1 \le j \le m$. Then the following are equivalent:*

(i) $\mathbf{y}_1, \dots, \mathbf{y}_m$ is a free basis of $\mathbb{Z}^n$;
(ii) $n = m$ and P is an invertible matrix such that P^{-1} has entries in $\mathbb{Z}$;
(iii) $n = m$ and $\det(P) = \pm 1$.

(A matrix $P \in \mathbb{Z}^{n,n}$ with $\det(P) = \pm 1$ is called *unimodular*.)

PROOF: (i) $\Rightarrow$ (ii). If $\mathbf{y}_1, \dots, \mathbf{y}_m$ is a free basis of $\mathbb{Z}^n$ then it spans $\mathbb{Z}^n$, so there is an $m \times n$ matrix $T = (\tau_{ij})$ with $\mathbf{x}_k = \sum_{j=1}^m \tau_{jk} \mathbf{y}_j$ for $1 \le k \le n$. Hence

$$\mathbf{x}_k = \sum_{j=1}^m \tau_{jk} \mathbf{y}_j = \sum_{j=1}^m \tau_{jk} \sum_{i=1}^n \rho_{ij} \mathbf{x}_i = \sum_{i=1}^n (\sum_{j=1}^m \rho_{ij} \tau_{jk}) \mathbf{x}_i,$$

and, since $\mathbf{x}_1, \dots, \mathbf{x}_n$ is a free basis, this implies that $\sum_{j=1}^m \rho_{ij} \tau_{jk} = 1$ when $i = k$ and 0 when $i \ne k$. In other words $PT = I_n$, and similarly $TP = I_m$, so P and T are inverse matrices. But we can think of P and T as inverse matrices over the field $\mathbb{Q}$, so it follows from First Year Linear Algebra that $m = n$, and $T = P^{-1}$ has entries in $\mathbb{Z}$.

(ii) $\Rightarrow$ (i). If $T = P^{-1}$ has entries in $\mathbb{Z}$ then, again thinking of them as matrices over the field $\mathbb{Q}$, $\operatorname{rank}(P) = n$, so the columns of P are linearly independent over $\mathbb{Q}$ and hence also over $\mathbb{Z}$. Since the columns of P are just the column vectors representing $\mathbf{y}_1, \dots, \mathbf{y}_m$, this tells us that $\mathbf{y}_1, \dots, \mathbf{y}_m$ are linearly independent.

Using $PT = I_n$, for $1 \le k \le n$ we have

$$\sum_{j=1}^m \tau_{jk} \mathbf{y}_j = \sum_{j=1}^m \tau_{jk} \sum_{i=1}^n \rho_{ij} \mathbf{x}_i = \sum_{i=1}^n (\sum_{j=1}^m \rho_{ij} \tau_{jk}) \mathbf{x}_i = \mathbf{x}_k,$$

because $\sum_{j=1}^{m} \rho_{ij}\tau_{jk}$ is equal to 1 when $i = k$ and 0 when $i \neq k$. Since $\mathbf{x}_1, \ldots, \mathbf{x}_n$ spans $\mathbb{Z}^n$, and we can express each $\mathbf{x}_k$ as a linear combination of $\mathbf{y}_1, \ldots, \mathbf{y}_m$, it follows that $\mathbf{y}_1, \ldots, \mathbf{y}_m$ span $\mathbb{Z}^n$ and hence form a free basis of $\mathbb{Z}^n$.

(ii) $\Rightarrow$ (iii). If $T = P^{-1}$ has entries in $\mathbb{Z}$, then $\det(PT) = \det(P)\det(T) = \det(I_n) = 1$, and since $\det(P), \det(T) \in \mathbb{Z}$, this implies $\det(P) = \pm 1$.

(iii) $\Rightarrow$ (ii). From First year Linear Algebra, $P^{-1} = \frac{1}{\det(P)}\mathrm{adj}(P)$, so $\det(P) = \pm 1$ implies that P^{-1} has entries in $\mathbb{Z}$. □

Examples. If $n = 2$ and $\mathbf{y}_1 = (2\ 7)$, $\mathbf{y}_2 = (1\ 4)$, then $\det(P) = 8 - 7 = 1$, so $\mathbf{y}_1, \mathbf{y}_2$ is a free basis of $\mathbb{Z}^2$.

But, if $\mathbf{y}_1 = (1\ 0)$, $\mathbf{y}_2 = (0\ 2)$, then $\det(P) = 2$, so $\mathbf{y}_1, \mathbf{y}_2$ is not a free basis of $\mathbb{Z}^2$. Recall that in Linear Algebra over a field, any set of n linearly independent vectors in a vector space V of dimension n form a basis of V. This example shows that this result is not true in $\mathbb{Z}^n$, because $\mathbf{y}_1$ and $\mathbf{y}_2$ are linearly independent but do not span $\mathbb{Z}^2$.

But as in Linear Algebra, for $\mathbf{v} \in \mathbb{Z}^n$, if $\mathbf{x}(= \mathbf{v}^{\mathrm{T}})$ and $\mathbf{y}$ are the column vectors representing $\mathbf{v}$ using free bases $\mathbf{x}_1, \ldots, \mathbf{x}_n$ and $\mathbf{y}_1, \ldots, \mathbf{y}_n$, respectively, then we have $\mathbf{x} = P\mathbf{y}$, so $\mathbf{y} = P^{-1}\mathbf{x}$.

4.5 Unimodular elementary row and column operations and the Smith normal form for integral matrices

We interrupt our discussion of finitely generated abelian groups at this stage to investigate how the row and column reduction process of Linear Algebra can be adapted to matrices over $\mathbb{Z}$. Recall from MA106 that we can use elementary row and column operations to reduce an $m \times n$ matrix of rank r over a field K to a matrix $B = (\beta_{ij})$ with $\beta_{ii} = 1$ for $1 \leq i \leq r$ and $\beta_{ij} = 0$ otherwise. We called this the *Smith Normal Form* of the matrix. We can do something similar over $\mathbb{Z}$, but the non-zero elements β_{ii} will not necessarily all be equal to 1.

The reason that we disallowed $\lambda = 0$ for the row and column operations (R3) and (C3) (multiply a row or column by a scalar λ) was that we wanted all of our elementary operations to be reversible. When performed over $\mathbb{Z}$, (R1), (C1), (R2) and (C2) are reversible, but (R3) and (C3) are reversible only when $\lambda = \pm 1$. So, if A is an $m \times n$ matrix over $\mathbb{Z}$, then we define the three types of *unimodular* elementary row operations as follows:

(UR1): Replace some row $\mathbf{r}_i$ of A by $\mathbf{r}_i + t\mathbf{r}_j$, where $j \neq i$ and $t \in \mathbb{Z}$;

(UR2): Interchange two rows of A;

(UR3): Replace some row $\mathbf{r}_i$ of A by $-\mathbf{r}_i$.

The unimodular column operations (UC1), (UC2), (UC3) are defined similarly. Recall from MA106 that performing elementary row or column operations on a matrix A corresponds to multiplying A on the left or right, respectively, by an elementary matrix. These elementary matrices all have determinant ± 1 (1 for (UR1) and -1 for (UR2) and (UR3)), so are unimodular matrices over $\mathbb{Z}$.

Theorem 4.22 (Smith Normal Form) *Let A be an $m \times n$ matrix over $\mathbb{Z}$ with rank r. Then, by using a sequence of unimodular elementary row and column operations, we can reduce A to a matrix $B = (\beta_{ij})$ with $\beta_{ii} = d_i$ for $1 \leq i \leq r$ and $\beta_{ij} = 0$ otherwise, and where the integers d_i satisfy $d_i > 0$ for $1 \leq i \leq r$, and $d_i | d_{i+1}$ for $1 \leq i < r$. Subject to these conditions, the d_i are uniquely determined by the matrix A.*

Proof: We shall not prove the uniqueness part here. The fact that the number of non-zero β_{ii} is the rank of A follows from the fact that unimodular row and column operations do not change the rank. We use induction on $m + n$. The base case is $m = n = 1$, where there is nothing to prove. Also if A is the zero matrix then there is nothing to prove, so assume not.

Let d be the smallest entry with $d > 0$ in any matrix $C = (\gamma_{ij})$ that we can obtain from A by using unimodular elementary row and column operations. By using (R2) and (C2), we can move d to position $(1,1)$ and hence assume that $\gamma_{11} = d$. If d does not divide γ_{1j} for some $j > 0$, then we can write $\gamma_{1j} = qd + r$ with $q, r \in \mathbb{Z}$ and $0 < r < d$, and then replacing the j-th column $\mathbf{c}_j$ of C by $\mathbf{c}_j - q\mathbf{c}_1$ results in the entry r in position $(1,j)$, contrary to the choice of d. Hence $d|\gamma_{1j}$ for $2 \le j \le n$ and similarly $d|\gamma_{i1}$ for $2 \le i \le m$.

Now, if $\gamma_{1j} = qd$, then replacing $\mathbf{c}_j$ of C by $\mathbf{c}_j - q\mathbf{c}_1$ results in entry 0 position $(1,j)$. So we can assume that $\gamma_{1j} = 0$ for $2 \le j \le n$ and $\gamma_{i1} = 0$ for $2 \le i \le m$. If $m = 1$ or $n = 1$, then we are done. Otherwise, we have $C = (d) \oplus C'$ for some $(m-1) \times (n-1)$ matrix C'. By inductive hypothesis, the result of the theorem applies to C', so by applying unimodular row and column operations to C which do not involve the first row or column, we can reduce C to $D = (\delta_{ij})$, which satisfies $\delta_{11} = d$, $\delta_{ii} = d_i > 0$ for $2 \le i \le r$, and $\delta_{ij} = 0$ otherwise, where $d_i | d_{i+1}$ for $2 \le i < r$. To complete the proof, we still have to show that $d|d_2$. If not, then adding row 2 to row 1 results in d_2 in position (1,2) not divisible by d, and we obtain a contradiction as before. □

In the following two examples we determine the Smith Normal Forms of the matrix A.

Example 1. Let A be the matrix $A = \begin{pmatrix} 42 & 21 \\ -35 & -14 \end{pmatrix}$.

The general strategy is to reduce the size of entries in the first row and column, until the (1,1)-entry divides all other entries in the first row and column. Then we can clear all of these other entries.

Matrix	**Operation**	**Matrix**	**Operation**
$\begin{pmatrix} 42 & 21 \\ -35 & -14 \end{pmatrix}$	$\mathbf{c}_1 \to \mathbf{c}_1 - 2\mathbf{c}_2$	$\begin{pmatrix} 0 & 21 \\ -7 & -14 \end{pmatrix}$	$\mathbf{r}_2 \to -\mathbf{r}_2$ $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$
$\begin{pmatrix} 7 & 14 \\ 0 & 21 \end{pmatrix}$	$\mathbf{c}_2 \to \mathbf{c}_2 - 2\mathbf{c}_1$	$\begin{pmatrix} 7 & 0 \\ 0 & 21 \end{pmatrix}$	

Example 2. Let A be the matrix $A = \begin{pmatrix} -18 & -18 & -18 & 90 \\ 54 & 12 & 45 & 48 \\ 9 & -6 & 6 & 63 \\ 18 & 6 & 15 & 12 \end{pmatrix}$.

Matrix	**Operation**	**Matrix**	**Operation**
$\begin{pmatrix} -18 & -18 & -18 & 90 \\ 54 & 12 & 45 & 48 \\ 9 & -6 & 6 & 63 \\ 18 & 6 & 15 & 12 \end{pmatrix}$	$\mathbf{c}_1 \to \mathbf{c}_1 - \mathbf{c}_3$	$\begin{pmatrix} 0 & -18 & -18 & 90 \\ 9 & 12 & 45 & 48 \\ 3 & -6 & 6 & 63 \\ 3 & 6 & 15 & 12 \end{pmatrix}$	$\mathbf{r}_1 \leftrightarrow \mathbf{r}_4$
$\begin{pmatrix} 3 & 6 & 15 & 12 \\ 9 & 12 & 45 & 48 \\ 3 & -6 & 6 & 63 \\ 0 & -18 & -18 & 90 \end{pmatrix}$	$\mathbf{r}_2 \to \mathbf{r}_2 - 3\mathbf{r}_1$ $\mathbf{r}_3 \to \mathbf{r}_3 - \mathbf{r}_1$	$\begin{pmatrix} 3 & 6 & 15 & 12 \\ 0 & -6 & 0 & 12 \\ 0 & -12 & -9 & 51 \\ 0 & -18 & -18 & 90 \end{pmatrix}$	$\mathbf{c}_2 \to \mathbf{c}_2 - 2\mathbf{c}_1$ $\mathbf{c}_3 \to \mathbf{c}_3 - 5\mathbf{c}_1$ $\mathbf{c}_4 \to \mathbf{c}_4 - 4\mathbf{c}_1$
$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & -6 & 0 & 12 \\ 0 & -12 & -9 & 51 \\ 0 & -18 & -18 & 90 \end{pmatrix}$	$\mathbf{c}_2 \to -\mathbf{c}_2$ $\mathbf{c}_2 \to \mathbf{c}_2 + \mathbf{c}_3$	$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 6 & 0 & 12 \\ 0 & 3 & -9 & 51 \\ 0 & 0 & -18 & 90 \end{pmatrix}$	$\mathbf{r}_2 \leftrightarrow \mathbf{r}_3$
$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & -9 & 51 \\ 0 & 6 & 0 & 12 \\ 0 & 0 & -18 & 90 \end{pmatrix}$	$\mathbf{r}_3 \to \mathbf{r}_3 - 2\mathbf{r}_2$	$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & -9 & 51 \\ 0 & 0 & 18 & -90 \\ 0 & 0 & -18 & 90 \end{pmatrix}$	$\mathbf{c}_3 \to \mathbf{c}_3 + 3\mathbf{c}_2$ $\mathbf{c}_4 \to \mathbf{c}_4 - 17\mathbf{c}_2$
$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 18 & -90 \\ 0 & 0 & -18 & 90 \end{pmatrix}$	$\mathbf{c}_4 \to \mathbf{c}_4 + 5\mathbf{c}_3$ $\mathbf{r}_4 \to \mathbf{r}_4 + \mathbf{r}_3$	$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 18 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$	

Note: There is also a generalisation to integer matrices of the row reduced normal form from Linear Algebra, where only row operations are allowed. This is known as the *Hermite Normal Form* and is more complicated. It will appear on an exercise sheet.

4.6 Subgroups of free abelian groups

Proposition 4.23 *Any subgroup of a finitely generated abelian group is finitely generated.*

Proof: Let $K < G$ with G an abelian group generated by $x_1, \ldots, x_n$. We shall prove by induction on n that K can be generated by at most n elements. If $n = 1$ then G is cyclic. Write $G = \{nx | n \in \mathbb{Z}\}$. Let m be the smallest positive number such that $mx \in K$. If such a number does not exist then $K = \{0\}$. Otherwise, $K \supseteq \{nmx | n \in \mathbb{Z}\}$. The opposite inclusion follows using division with a remainder: write $t = qm + r$ with $0 \le r < m$. Then $tx \in K$ if and only if $rx = (t - mq)x \in K$ if and only if $r = 0$ due to minimality of m. In both cases K is cyclic.

Suppose $n > 1$, and let H be the subgroup of G generated by $x_1, \ldots, x_{n-1}$. By induction, $K \cap H$ is generated by $y_1, \ldots, y_{m-1}$, say, with $m \le n$. If $K \le H$, then $K = K \cap H$ and we are done, so suppose not.

Then there exist elements of the form $h + tx_n \in K$ with $h \in H$ and $t \neq 0$. Since $-(h + tx_n) \in K$, we can assume that $t > 0$. Choose such an element $y_m = h + tx_n \in K$ with t minimal subject to $t > 0$. We claim that K is generated by $y_1, \ldots, y_m$, which will complete the proof. Let $k \in K$. Then $k = h' + ux_n$ with $h' \in H$ and $u \in \mathbb{Z}$. If t does not divide u then we can write $u = tq + r$ with $q, r \in \mathbb{Z}$ and $0 < r < t$, and then $k - qy_m = (h' - qh) + rx_n \in K$, contrary to the choice of t. So $t | u$ and hence $u = tq$ and $k - qy_m \in K \cap H$. But $K \cap H$ is generated by $y_1, \ldots, y_{m-1}$, so we are done. □

Now let H be a subgroup of the free abelian group $\mathbb{Z}^n$, and suppose that H is generated by $\mathbf{v}_1, \ldots, \mathbf{v}_m$. Then H can be represented by an $n \times m$ matrix A in which the columns are $\mathbf{v}_1^T, \ldots, \mathbf{v}_m^T$.

Example 3. If $n = 3$ and H is generated by $\mathbf{v}_1 = (1\ 3\ \ -1)$ and $\mathbf{v}_2 = (2\ 0\ 1)$, then $A = \begin{pmatrix} 1 & 2 \\ 3 & 0 \\ -1 & 1 \end{pmatrix}$.

As we saw above, if we use a different free basis $\mathbf{y}_1, \ldots, \mathbf{y}_n$ of $\mathbb{Z}^n$ with basis change matrix P, then each column $\mathbf{v}_j^T$ of A is replaced by $P^{-1}\mathbf{v}_j^T$, and hence A itself is replaced by $P^{-1}A$.

So in Example 3, if we use the basis $\mathbf{y}_1 = (0\ \ -1\ 0)$, $\mathbf{y}_2 = (1\ 0\ 1)$, $\mathbf{y}_3 = (1\ 1\ 0)$ of $\mathbb{Z}^n$, then

$$P = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix}, \quad P^{-1}A = \begin{pmatrix} -1 & 1 \\ -1 & 1 \\ 2 & 1 \end{pmatrix}.$$

For example, the first column $(-1\ -1\ 2)^{\mathrm{T}}$ of $P^{-1}A$ represents $-\mathbf{y}_1 - \mathbf{y}_2 + 2\mathbf{y}_3 = (1\ 3\ -1) = \mathbf{v}_1$.

In particular, if we perform a unimodular elementary row operation on A, then the resulting matrix represents the same subgroup H of $\mathbb{Z}^n$ but using a different free basis of $\mathbb{Z}^n$.

We can clearly replace a generator $\mathbf{v}_i$ of H by $\mathbf{v}_i + r\mathbf{v}_j$ for $r \in \mathbb{Z}$ without changing the subgroup H that is generated. We can also interchange two of the generators or replace one of the generators $\mathbf{v}_i$ by $-\mathbf{v}_i$ without changing H. In other words, performing a unimodular elementary column operation on A amounts to changing the generating set for H, so again the resulting matrix still represents the same subgroup H of $\mathbb{Z}^n$.

Summing up, we have:

Proposition 4.24 *Suppose that the subgroup H of $\mathbb{Z}^n$ is represented by the matrix $A \in \mathbb{Z}^{n,m}$. If the matrix $B \in \mathbb{Z}^{n,m}$ is obtained by performing a sequence of unimodular row and column operations on A, then B represents the same subgroup H of $\mathbb{Z}^n$ using a (possibly) different free basis of $\mathbb{Z}^n$.*

In particular, by Theorem 4.22, we can transform A to a matrix B in Smith Normal Form. Therefore, if B represents H with the free basis $\mathbf{y}_1, \ldots, \mathbf{y}_n$ of $\mathbb{Z}^n$, then the r non-zero columns of B correspond to the elements $d_1\mathbf{y}_1, d_2\mathbf{y}_2, \ldots, d_r\mathbf{y}_r$ of $\mathbb{Z}^n$. So we have:

Theorem 4.25 *Let H be a subgroup of $\mathbb{Z}^n$. Then there exists a free basis $\mathbf{y}_1, \ldots, \mathbf{y}_n$ of $\mathbb{Z}^n$ such that $H = \langle\, d_1\mathbf{y}_1, d_2\mathbf{y}_2, \ldots, d_r\mathbf{y}_r \,\rangle$, where each $d_i > 0$ and $d_i | d_{i+1}$ for $1 \le i < r$.*

In Example 3, it is straightforward to calculate the Smith Normal Form of A, which is $\begin{pmatrix} 1 & 0 \\ 0 & 3 \\ 0 & 0 \end{pmatrix}$, so $H = \langle \mathbf{y}_1, 3\mathbf{y}_2 \rangle$.

By keeping track of the unimodular row operations carried out, we can, if we need to, find the free basis $\mathbf{y}_1, \ldots, \mathbf{y}_n$ of $\mathbb{Z}^n$. Doing this in Example 3, we get:

Matrix	Operation	New free basis
$\begin{pmatrix} 1 & 2 \\ 3 & 0 \\ -1 & 1 \end{pmatrix}$	$\mathbf{r}_2 \to \mathbf{r}_2 - 3\mathbf{r}_1$ $\mathbf{r}_3 \to \mathbf{r}_3 + \mathbf{r}_1$	$\mathbf{y}_1 = (1\ 3\ -1)$, $\mathbf{y}_2 = (0\ 1\ 0)$, $\mathbf{y}_3 = (0\ 0\ 1)$
$\begin{pmatrix} 1 & 2 \\ 0 & -6 \\ 0 & 3 \end{pmatrix}$	$\mathbf{c}_2 \to \mathbf{c}_2 - 2\mathbf{c}_1$	$\mathbf{y}_1 = (1\ 3\ -1)$, $\mathbf{y}_2 = (0\ 1\ 0)$, $\mathbf{y}_3 = (0\ 0\ 1)$
$\begin{pmatrix} 1 & 0 \\ 0 & -6 \\ 0 & 3 \end{pmatrix}$	$\mathbf{r}_2 \leftrightarrow \mathbf{r}_3$	$\mathbf{y}_1 = (1\ 3\ -1)$, $\mathbf{y}_2 = (0\ 0\ 1)$, $\mathbf{y}_3 = (0\ 1\ 0)$
$\begin{pmatrix} 1 & 0 \\ 0 & 3 \\ 0 & -6 \end{pmatrix}$	$\mathbf{r}_3 \to \mathbf{r}_3 + 2\mathbf{r}_2$	$\mathbf{y}_1 = (1\ 3\ -1)$, $\mathbf{y}_2 = (0\ -2\ 1)$, $\mathbf{y}_3 = (0\ 1\ 0)$
$\begin{pmatrix} 1 & 0 \\ 0 & 3 \\ 0 & 0 \end{pmatrix}$		

4.7 General finitely generated abelian groups

Let G be a finitely generated abelian group. If G has n generators, Proposition 4.19 gives a surjective homomorphism $\phi : \mathbb{Z}^n \to G$. From the First isomorphism Theorem (Theorem 4.18) we deduce that $G \cong \mathbb{Z}^n/K$, where $K = \ker(\phi)$. So we have proved that every finitely generated abelian group is isomorphic to a quotient group of a free abelian group.

From the definition of ϕ, we see that

$$K = \{\, (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n \mid \alpha_1 x_1 + \cdots + \alpha_n x_n = 0_G \,\}.$$

By Theorem 4.23, this subgroup K is generated by finitely many elements $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of $\mathbb{Z}^n$. The notation

$$\langle\, \mathbf{x}_1, \ldots, \mathbf{x}_n \mid \mathbf{v}_1, \ldots, \mathbf{v}_m \,\rangle$$

is often used to denote the quotient group $\mathbb{Z}^n/K$, so we have

$$G \cong \langle\, \mathbf{x}_1, \ldots, \mathbf{x}_n \mid \mathbf{v}_1, \ldots, \mathbf{v}_m \,\rangle.$$

Now we can apply Theorem 4.25 to this subgroup K, and deduce that there is a free basis $\mathbf{y}_1, \ldots, \mathbf{y}_n$ of $\mathbb{Z}^n$ such that $K = \langle\, d_1\mathbf{y}_1, \ldots, d_r\mathbf{y}_r \,\rangle$ for some $r \leq n$, where each $d_i > 0$ and $d_i | d_{i+1}$ for $1 \leq i < r$.

So we also have

$$G \cong \langle\, \mathbf{y}_1, \ldots, \mathbf{y}_n \mid d_1\mathbf{y}_1, \ldots, d_r\mathbf{y}_r \,\rangle,$$

and G has generators $y_1, \ldots, y_n$ with $d_i y_i = 0$ for $1 \leq i \leq r$.

Proposition 4.26 *The group*

$$\langle\, \mathbf{y}_1, \ldots, \mathbf{y}_n \mid d_1\mathbf{y}_1, \ldots, d_r\mathbf{y}_r \,\rangle$$

is isomorphic to the direct sum of cyclic groups

$$\mathbb{Z}_{d_1} \oplus \mathbb{Z}_{d_2} \oplus \cdots \oplus \mathbb{Z}_{d_r} \oplus \mathbb{Z}^{n-r}.$$

PROOF: This is another application of the First Isomorphism Theorem. Let $H = \mathbb{Z}_{d_1} \oplus \mathbb{Z}_{d_2} \oplus \cdots \oplus \mathbb{Z}_{d_r} \oplus \mathbb{Z}^{n-r}$, so H is generated by $y_1, \ldots, y_n$, with $y_1 = (1, 0, \ldots, 0), \ldots, y_n = (0, \ldots, 0, 1)$.

Let $\mathbf{y}_1, \ldots, \mathbf{y}_n$ be the standard free basis of $\mathbb{Z}^n$. By Proposition 4.19, there is a surjective homomorphism ϕ from $\mathbb{Z}^n$ to H for which

$$\phi(\alpha_1\mathbf{y}_1 + \cdots + \alpha_n\mathbf{y}_n) = \alpha_1 y_1 + \cdots + \alpha_n y_n$$

for all $\alpha_1, \ldots, \alpha_n \in \mathbb{Z}$. By Theorem 4.18, we have $H \cong \mathbb{Z}^n/K$, with

$$K = \{\, (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n \mid \alpha_1 y_1 + \cdots + \alpha_n y_n = 0_H \,\}.$$

Now $\alpha_1 y_1 + \cdots + \alpha_n y_n$ is the element $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of H, which is the zero element if and only if α_i is the zero element of $\mathbb{Z}_{d_i}$ for $1 \le i \le r$ and $\alpha_i = 0$ for $r+1 \le i \le n$.

But α_i is the zero element of $\mathbb{Z}_{d_i}$ if and only if $d_i | \alpha_i$, so we have

$$K = \{\, (\alpha_1, \alpha_2, \ldots, \alpha_r, 0, \ldots, 0) \in \mathbb{Z}^n \mid d_i \text{ divides } \alpha_i \text{ for } 1 \le i \le r \,\}$$

which is generated by the elements $d_1\mathbf{y}_1, \ldots, d_r\mathbf{y}_r$. Thus

$$H \cong \mathbb{Z}^n/K = \langle\, \mathbf{y}_1, \ldots, \mathbf{y}_n \mid d_1\mathbf{y}_1, \ldots, d_r\mathbf{y}_r \,\rangle$$

and we are done. □

Putting all of these results together, we get the main theorem of this section.

Theorem 4.27 (Fundamental theorem of finitely generated abelian groups) *If G is a finitely generated abelian group, then G is isomorphic to a direct sum of cyclic groups. More precisely, if G is generated by n elements, then there is an integer r with $0 \le r \le n$, and there are integers $d_1, \ldots, d_r$ with $d_i > 0$ and $d_i | d_{i+1}$ such that*

$$G \cong \mathbb{Z}_{d_1} \oplus \mathbb{Z}_{d_2} \oplus \cdots \oplus \mathbb{Z}_{d_r} \oplus \mathbb{Z}^{n-r}.$$

Thus G is isomorphic to a direct sum of r finite cyclic groups of orders $d_1, \ldots, d_r$, and $n-r$ infinite cyclic groups.

There may be some factors $\mathbb{Z}_1$, the trivial group of order 1. These can be omitted from the direct sum (except in the case when $G \cong \mathbb{Z}_1$ is trivial). It can be deduced from the uniqueness part of Theorem 4.22, which we did not prove, that the numbers in the sequence $d_1, d_2, \ldots, d_r$ that are greater than 1 are uniquely determined by G.

The integer $n - r$ may be 0, which is the case if and only if G is finite. At the other extreme, if all $d_i = 1$, then G is free abelian.

The group G corresponding to Example 1 in Section 4.5 is

$$\langle\, \mathbf{x}_1, \mathbf{x}_2 \mid 42\mathbf{x}_1 - 35\mathbf{x}_2,\ 21\mathbf{x}_1 - 14\mathbf{x}_2 \,\rangle$$

and we have $G \cong \mathbb{Z}_7 \oplus \mathbb{Z}_{21}$, a group of order $7 \times 21 = 147$.

The group defined by Example 2 in Section 4.5 is

$$\left\langle \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4 \;\middle|\; \begin{array}{ll} -18\mathbf{x}_1 + 54\mathbf{x}_2 + 9\mathbf{x}_3 + 18\mathbf{x}_4, & -18\mathbf{x}_1 + 12\mathbf{x}_2 - 6\mathbf{x}_3 + 6\mathbf{x}_4, \\ -18\mathbf{x}_1 + 45\mathbf{x}_2 + 6\mathbf{x}_3 + 15\mathbf{x}_4, & 90\mathbf{x}_1 + 48\mathbf{x}_2 + 63\mathbf{x}_3 + 12\mathbf{x}_4 \end{array} \right\rangle,$$

which is isomorphic to $\mathbb{Z}_3 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_{18} \oplus \mathbb{Z}$, and is an infinite group with a (maximal) finite subgroup of order $3 \times 3 \times 18 = 162$,

The group defined by Example 3 in Section 4.6 is

$$\langle\, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \mid \mathbf{x}_1 + 3\mathbf{x}_2 - \mathbf{x}_3,\ 2\mathbf{x}_1 + \mathbf{x}_3 \,\rangle,$$

and is isomorphic to $\mathbb{Z}_1 \oplus \mathbb{Z}_3 \oplus \mathbb{Z} \cong \mathbb{Z}_3 \oplus \mathbb{Z}$, so it is infinite, with a finite subgroup of order 3.

4.8 Finite abelian groups

In particular, for any finite abelian group G, we have $G \cong \mathbb{Z}_{d_1} \oplus \mathbb{Z}_{d_2} \oplus \cdots \oplus \mathbb{Z}_{d_r}$, where $d_i | d_{i+1}$ for $1 \leq i < r$, and $|G| = d_1 d_2 \cdots d_r$.

Suppose that $d_1, \ldots, d_r$ and $e_1, \ldots, e_s$ are integers satisfying $d_1, e_1 \geq 2$ and $d_i | d_{i+1}$ for $1 \leq i < r$ and $e_i | e_{i+1}$ for $1 \leq i < s$. From the uniqueness part of Theorem 4.22 (which we did not prove), it follows that the groups $\mathbb{Z}_{d_1} \oplus \mathbb{Z}_{d_2} \oplus \cdots \oplus \mathbb{Z}_{d_r}$ and $\mathbb{Z}_{e_1} \oplus \mathbb{Z}_{e_2} \oplus \cdots \oplus \mathbb{Z}_{e_s}$ are isomorphic if and only if $r = s$ and $d_i = e_i$ for $1 \leq i \leq r$.

So the isomorphism classes of finite abelian groups of order $n > 0$ are in one-one correspondence with expressions $n = d_1 d_2 \cdots d_r$ for which $d_i | d_{i+1}$ for $1 \leq i < r$. This enables us to classify isomorphism classes of finite abelian groups.

Examples. 1. $n = 4$. The decompositions are 4 and 2×2, so $G \cong \mathbb{Z}_4$ or $\mathbb{Z}_2 \oplus \mathbb{Z}_2$.

2. $n = 15$. The only decomposition is 15, so $G \cong \mathbb{Z}_{15}$ is necessarily cyclic.

3. $n = 36$. Decompositions are 36, 2×18, 3×12 and 6×6, so $G \cong \mathbb{Z}_{36}$, $\mathbb{Z}_2 \oplus \mathbb{Z}_{18}$, $\mathbb{Z}_3 \oplus \mathbb{Z}_{12}$ and $\mathbb{Z}_6 \oplus \mathbb{Z}_6$.

Although we have not proved in general that groups of the same order but with different decompositions of the type above are not isomorphic, this can always be done in specific examples by looking at the orders of elements.

We saw in an exercise above that if $\phi : G \to H$ is an isomorphism, then $|g| = |\phi(g)|$ for all $g \in G$. It follows that isomorphic groups have the same number of elements of each order.

Note also that, if $g = (g_1, g_2, \ldots, g_n)$ is an element of a direct sum of n groups, then $|g|$ is the least common multiple of the orders $|g_i|$ of the components of g.

Among the four groups of order 36, $G_1 = \mathbb{Z}_{36}$, $G_2 = \mathbb{Z}_2 \oplus \mathbb{Z}_{18}$, $G_3 = \mathbb{Z}_3 \oplus \mathbb{Z}_{12}$ and $G_4 = \mathbb{Z}_6 \oplus \mathbb{Z}_6$, we see that only G_1 contains elements of order 36. Hence G_1 cannot be isomorphic to G_2, G_3 or G_4. Of the three groups G_2, G_3 and G_4, only G_2 contains elements of order 18, so G_2 cannot be isomorphic to G_3 or G_4. Finally, G_3 has elements of order 12 but G_4 does not, so G_3 and G_4 are not isomorphic, and we have now shown that no two of the four groups are isomorphic to each other.

As a slightly harder example, $\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_4$ is not isomorphic to $\mathbb{Z}_4 \oplus \mathbb{Z}_4$, because the former has 7 elements of order 2, whereas the latter has only 3.

4.9 Tensor products

Given two abelian groups A and B, one can form a new abelian group $A \otimes B$, their *tensor product* - do not confuse it with the direct product! We denote by F the free abelian group with the elements of the direct product $A \times B$ as a basis:

$$F = \langle\, A \times B \mid \emptyset \,\rangle.$$

Elements of F are formal finite $\mathbb{Z}$-linear combinations $\sum_i n_i(a_i, b_i)$, $n_i \in \mathbb{Z}$, $a_i \in A$, $b_i \in B$. Let F_0 be the subgroup of F generated by the following elements

$$\begin{array}{ll} (a + a', b) - (a, b) - (a', b) & \qquad n(a, b) - (na, b) \\ (a, b + b') - (a, b) - (a, b') & \qquad n(a, b) - (a, nb) \end{array} \tag{1}$$

for all possible $n \in \mathbb{Z}$, $a, a' \in A$, $b, b' \in B$. The *tensor product* is the quotient group

$$A \otimes B = F/F_0 = \langle A \times B \mid \text{ relations in (1) } \rangle.$$

We have to get used to this definition that may seem strange at a first glance. First, it is easy to materialise certain elements of $A \otimes B$. *Elementary tensors* are

$$a \otimes b = (a, b) + F_0$$

for various $a \in A$, $b \in B$. However, it is important to realise that not all tensors are elementary. Generators for F_0 become relations on elementary tensors,

$$(a+a') \otimes b = a \otimes b + a' \otimes b, \quad n(a \otimes b) = (na) \otimes b,$$
$$a \otimes (b+b') = a \otimes b + a \otimes b', \quad n(a \otimes b) = a \otimes (nb),$$

so a general element of $A \otimes B$ is a sum of elementary tensors $\sum_i a_i \otimes b_i$, but usually it is not an elementary tensor itself.

Exercise. Show that $a \otimes 0 = 0 \otimes b = 0$ for all $a \in A$, $b \in B$.

If the set $\{\mathbf{a}_i\}$ spans A and the set $\{\mathbf{b}_j\}$ spans B, then the set $\{\mathbf{a}_i \otimes \mathbf{b}_j\}$ spans $A \otimes B$. Indeed, given $\sum_k a_k \otimes b_k \in A \otimes B$, we can express all $a_k = \sum_i n_{ki}\mathbf{a}_i$, and all $b_k = \sum_j m_{kj}\mathbf{b}_j$. Then

$$\sum_k a_k \otimes b_k = \sum_k (\sum_i n_{ki}\mathbf{a}_i) \otimes (\sum_j m_{kj}\mathbf{b}_j) = \sum_{k,i,j} n_{ki}m_{kj}\, \mathbf{a}_i \otimes \mathbf{b}_j.$$

In fact, even a more subtle statement holds.

Exercise. Suppose that A, B are free abelian groups. Let $\mathbf{a}_i$ be a free basis of A and let $\mathbf{b}_j$ be a basis of B. Show that the tensor product $A \otimes B$ is a free abelian group with a free basis consisting of the elementary tensors $\mathbf{a}_i \otimes \mathbf{b}_j$.

However, for general groups tensor products could behave in quite an unpredictable way. For instance, $\mathbb{Z}_2 \otimes \mathbb{Z}_3 = 0$. Indeed,

$$1_{\mathbb{Z}_2} \otimes 1_{\mathbb{Z}_3} = 3 \cdot 1_{\mathbb{Z}_2} \otimes 1_{\mathbb{Z}_3} = 1_{\mathbb{Z}_2} \otimes 3 \cdot 1_{\mathbb{Z}_3} = 0.$$

To help sorting out zero from nonzero elements in tensor products we need to understand a connection between tensor products and bilinear maps. Let A, B, and C be abelian groups.

Definition. A function $\omega : A \times B \to C$ is a bilinear map if

$$\omega(a+a', b) = \omega(a,b) + \omega(a',b) \qquad n\omega(a,b) = \omega(na,b)$$
$$\omega(a, b+b') = \omega(a,b) + \omega(a,b') \qquad n\omega(a,b) = \omega(a,nb)$$

for all possible $n \in \mathbb{Z}$, $a, a' \in A$, $b, b' \in B$. Let $\mathrm{Bil}(A \times B, C)$ be the set of all bilinear maps from $A \times B$ to C.

Lemma 4.28 (Universal property of tensor product) *The function*

$$\theta : A \times B \to A \otimes B, \quad \theta(a,b) = a \otimes b$$

is a bilinear map. This bilinear map is universal, that is the composition with θ defines a bijection

$$\hom(A \otimes B, C) \to \mathrm{Bil}(A \times B, C), \quad \phi \mapsto \phi \circ \theta.$$

PROOF: The function θ is a bilinear map: the four properties of a bilinear map easily follow from the corresponding generators of F_0. For instance, $\theta(a+a', b) = \theta(a,b) + \theta(a',b)$ because $(a+a', b) - (a,b) - (a',b) \in F_0$.

Let Fun denote the set of functions between two sets. Recall that F denotes the free abelian group with basis indexed by $A \times B$. By the universal property of free abelian groups (Corollary 4.20), we have a bijection

$$\hom(F, C) \to Fun(A \times B, C).$$

Bilinear maps correspond to functions vanishing on F_0, or equivalently to linear maps from F/F_0 (Lemma 4.17). $\square$

In the following section we will need a criterion for elements of $\mathbb{R}\otimes S^1$ to be nonzero. The circle group S^1 is a group under multiplication, creating certain confusion for tensor products. To avoid this confusion we identify the multiplicative group S^1 with the additive group $\mathbb{R}/2\pi\mathbb{Z}$ via the natural isomorphism $e^{xi} \mapsto x + 2\pi\mathbb{Z}$.

Proposition 4.29 *Let $a \otimes (x + 2\pi\mathbb{Z}) \in \mathbb{R} \otimes \mathbb{R}/2\pi\mathbb{Z}$ where $a, x \in \mathbb{R}$. Then $a \otimes (x + 2\pi\mathbb{Z}) = 0$ if and only if $a = 0$ or $x \in \pi\mathbb{Q}$.*

PROOF: If $a = 0$, then $a \otimes (x + 2\pi\mathbb{Z}) = 0$. If $a \neq 0$ and $x = \frac{n}{m}\pi$ with $m, n \in \mathbb{Z}$, then

$$a \otimes (x + 2\pi\mathbb{Z}) = 2m\frac{a}{2m} \otimes (x + 2\pi\mathbb{Z}) = \frac{a}{2m} \otimes 2m(x + 2\pi\mathbb{Z}) = a \otimes (2n\pi + 2\pi\mathbb{Z}) = a \otimes 0 = 0.$$

In the opposite direction, let us consider $a \otimes (x + 2\pi\mathbb{Z})$ with $a \neq 0$ and $x/\pi \notin \mathbb{Q}$. It suffices to construct a bilinear map $\phi : \mathbb{R} \times \mathbb{R}/2\pi\mathbb{Z} \to A$ to some group A such that $\phi(a, x + 2\pi\mathbb{Z}) \neq 0$. By Lemma 4.28 this gives a homomorphism $\widetilde{\phi} : \mathbb{R} \otimes \mathbb{R}/2\pi\mathbb{Z} \to A$ with $\widetilde{\phi}(a \otimes (x + 2\pi\mathbb{Z})) = \phi(a, x + 2\pi\mathbb{Z}) \neq 0$. Hence, $a \otimes (x + 2\pi\mathbb{Z}) \neq 0$.

Let us consider $\mathbb{R}$ as a vector space over $\mathbb{Q}$. The subgroup $\pi\mathbb{Q}$ of $\mathbb{R}$ is a vector subspace, hence the quotient group $A = \mathbb{R}/\pi\mathbb{Q}$ is also a vector space over $\mathbb{Q}$. Since $2\pi\mathbb{Z} \subset \pi\mathbb{Q}$, we have a homomorphism

$$\beta : \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{R}/\pi\mathbb{Q}, \qquad \beta(z + 2\pi\mathbb{Z}) = z + \pi\mathbb{Q}.$$

Since $x/\pi \notin \mathbb{Q}$, it follows that $\beta(x + 2\pi\mathbb{Z}) \neq 0$. Choose a basis e_i of $\mathbb{R}$ over $\mathbb{Q}$ such that $e_1 = a$. Let $e^i : \mathbb{R} \to \mathbb{Q}$ be the linear function[22] computing the i-th coordinate in this basis:

$$e^i(\sum_j x_j e_j) = x_i.$$

The required bilinear map is defined using multiplication by a scalar in $A = \mathbb{R}/\pi\mathbb{Q}$ by the rule $\phi(b, z+2\pi\mathbb{Z}) = e^1(b)\beta(z+2\pi\mathbb{Z})$. Clearly, $\phi(a, x+2\pi\mathbb{Z}) = e^1(e_1)\beta(x+2\pi\mathbb{Z}) = 1\cdot\beta(x+2\pi\mathbb{Z}) = x + \pi\mathbb{Q} \neq 0$. □

Exercise. $\mathbb{Z}_n \otimes \mathbb{Z}_m \cong \mathbb{Z}_{\gcd(n,m)}$.

4.10 Hilbert's Third problem

All the hard work we have done is going to pay off now. We will understand a solution of the third Hilbert problem. In 1900 Hilbert formulated 23 problems that, in his view, would influence Mathematics of the 20th century. The third problem was solved first, in the same year 1900 by Dehn, which is quite remarkable as the problem was missing from Hilbert's lecture and appeared in print only in 1902, two years after its solution.

In his third problem Hilbert asks whether two 3D polytopes of the same volume are *scissor congruent*. Recall that M and N are congruent if there is a motion that moves M to N. The polytopes M and N are scissor congruent if it is possible to cut M into pieces M_i (cutting along planes) and N into pieces N_i such that individual pieces M_i and N_i are congruent for each i.

Let us consider a "scissor group" generated by all n-dimensional polytopes P

$$\mathbb{P}_n = \langle\, P \mid M - N,\, A - B - C \,\rangle$$

with these relations for each pair M, N of congruent polytopes and each cut $A = B \cup C$ of a polytope by a hyperplane. For a polytope M, let us denote by $[M] \in \mathbb{P}_n$ its class in the scissor

[22] commonly known as a covector

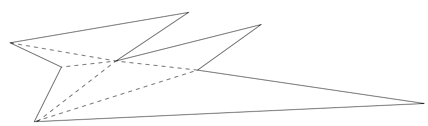
group. Clearly, M and N are scissor congruent if and only if $[M] = [N]$. By Lemma 4.17, n-dimensional volume is a homomorphism

$$\nu_n : \mathbb{P}_n \to \mathbb{R}, \quad \nu_n([M]) = volume(M).$$

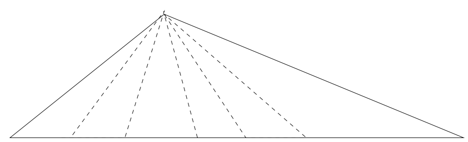
The 3rd Hilbert problem is whether ν_3 is injective.

Theorem 4.30 *ν_2 is injective.*

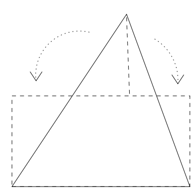
PROOF: For a polygon M, there are triangles $T_1, T_2, \ldots, T_n$ such that $[M] = [T_1] + [T_2] + \cdots + [T_n]$. It follows from triangulation of M illustrated on the next picture.



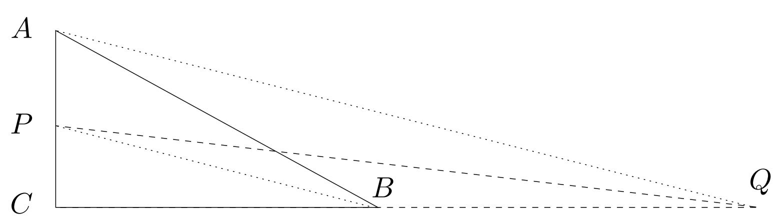
It suffices to show that if two triangles T and T' have the same area, then $[T] = [T']$. Indeed, using it, one can reshape triangles to $T_1', T_2', \ldots, T_n'$ so that they add up to a triangle T:



Then $[M] = [T_1] + [T_2] + \cdots + [T_n] = [T_1'] + [T_2'] + \cdots + [T_n'] = [T]$ and we supposedly know that two triangles of the same area are scissors equivalent. The following picture shows that a triangle with base b and height h is equivalent to the rectangle with sides b and $h/2$.



In particular, any triangle is equivalent to a right-angled triangle. The last picture shows that two right-angled triangles of the same area are scissors congruent.



The equal area triangles are CAB and CPQ. This means that $|CA||CB| = |CP||CQ|$. Hence, $|CA|/|CQ| = |CP|/|CB|$ and the triangles CPB and CAQ are similar. In particular, the edges AQ and PB are parallel, thus the triangles APB and QPB share the same base and height and, consequently, are scissors congruent. Finally,

$$[CAB] = [CPB] + [APB] = [CPB] + [QPB] = [CPQ]$$

as required. □

Observe that ν_n is surjective, hence ν_2 is an isomorphism and $\mathbb{P}_2 \cong \mathbb{R}$. However, ν_3 is not injective, disproving the 3rd Hilbert problem.

Theorem 4.31 *Let T be a regular tetrahedron, C a cube, both of unit volume. Then $[T] \neq [C] \in \mathbb{P}_3$.*

PROOF: Let M be a polytope with the of edges I. For each edge i, let h_i be its length h_i, $\alpha_i \in \mathbb{R}/2\pi\mathbb{Z}$ the angle near this edge. The *Dehn invariant* of M is

$$\delta(M) = \sum_i h_i \otimes \alpha_i \in \mathbb{R} \otimes \mathbb{R}/2\pi\mathbb{Z}.$$

Using Lemma 4.17, we conclude that δ is a well-defined homomorphism $\delta : \mathbb{P}_3 \to \mathbb{R}\otimes\mathbb{R}/2\pi\mathbb{Z}$. Indeed, δ defines a homomorphism from the free group F generated by all polytopes. Keeping in mind $\mathbb{P}_3 = F/F_0$, we need to check that δ vanishes on generators of F_0. Clearly, $\delta(M-N) = 0$ if M and N are congruent. It is slightly more subtle to see that $\delta(A - B - C) = 0$ if $A = B \cup C$ is a cut. One can collect the terms in 4 types of groups, with the zero sum in each group.

Survivor: an edge length h with angle α survives completely in B or C. This contributes $h \otimes \alpha - h \otimes \alpha = 0$.

Edge cut: an edge length h with angle α is cut into edges of lengths h_B in B and h_C in C. This contributes $h \otimes \alpha - h_B \otimes \alpha - h_C \otimes \alpha = (h - h_B - h_C) \otimes \alpha = 0 \otimes \alpha = 0$.

Angle cut: an edge length h with angle α has its angle cut into angles α_B in B and α_C in C. This contributes $h \otimes \alpha - h \otimes \alpha_B - h \otimes \alpha_C = h \otimes (\alpha - \alpha_B - \alpha_C) = h \otimes 0 = 0$.

New edge: a new edge of length h is created. If its angle in B is α, then its angle in C is $\pi - \alpha$. This contributes $-h \otimes \alpha - h \otimes (\pi - \alpha) = -h \otimes \pi = 0$, by Proposition 4.29.

Finally, using Proposition 4.29,

$$\delta([C]) = 12(1 \otimes \frac{\pi}{4}) = 12 \otimes \frac{\pi}{4} = 0, \text{ while } \delta([T]) = 6\frac{\sqrt{2}}{\sqrt[3]{3}} \otimes \arccos\frac{1}{3} \neq 0,$$

by Lemma 4.32. Hence, $[C] \neq [T]$. □

Lemma 4.32 $\arccos(1/3)/\pi \notin \mathbb{Q}$.

PROOF: Let $\arccos(1/3) = q\pi$. We consider a sequence $x_n = \cos(2^n q\pi)$. If q is rational, then this sequence admits only finitely many values. On the other hand, $x_0 = 1/3$ and

$$x_{n+1} = \cos(2 \cdot 2^n q\pi) = 2\cos^2(2^n q\pi) - 1 = 2x_n^2 - 1.$$

Computing several first terms,

$$x_1 = \frac{-7}{9},\ x_2 = \frac{17}{81},\ x_3 = \frac{-5983}{3^8},\ x_4 = \frac{28545857}{3^{16}}, \ldots$$

as can be easily shown the denominators grow indefinitely. Contradiction. □

Now it is natural to ask what exactly the group $\mathbb{P}_3$ is. It was proved later (in 1965) that the joint homomorphism $(\nu_3, \delta) : \mathbb{P}_3 \to \mathbb{R} \oplus (\mathbb{R} \otimes \mathbb{R}/2\pi\mathbb{Z})$ is injective. It is not surjective and the image can be explicitly described, but we won't do it here.

4.11 Possible topics for the second year essays

If you would like to write an essay taking something further from this course, here are some suggestions. Ask me if you want more information.

(i) Bilinear and quadratic forms over fields of characteristic 2 (i.e. where $1 + 1 = 0$). You can do hermitian forms too.

(ii) Grassmann algebras, determinants and tensors.

(iii) Matrix Exponents and Baker-Campbell-Hausdorff formula.

(iv) Abelian groups and public key cryptography (be careful not to repeat whatever is covered in Algebra 2).

(v) Lattices (abelian groups with bilinear forms), the E_8 lattice and the Leech lattice.
(vi) Abelian group law on an elliptic curve.
(vii) Groups $\mathbb{P}_n$ for other n, including a precise description of $\mathbb{P}_3$.

The End