1
Linear spaces

In this chapter we recall briefly some salient facts about linear spaces and linear maps. Proofs for the most part are omitted.

Maps

Let X and Y be sets and $f : X \to Y$ a map. Then, for each $x \in X$ an element $f(x) \in Y$ is defined, the subset of Y consisting of all such elements being called the *image* of f, denoted by $\operatorname{im} f$. More generally $f : X \rightarrowtail Y$ will denote a map of an unspecified subset of X to Y, X being called the *source* of the map and the subset of X consisting of those points $x \in X$ for which $f(x)$ is defined being called the *domain* of f, denoted by $\operatorname{dom} f$. In either case the set Y is the *target* of f.

Given a map $f : X \rightarrowtail Y$ and a point $y \in Y$, the subset $f^{-1}\{y\}$ of X consisting of those points $x \in X$ such that $f(x) = y$ is called the *fibre* of f over y, this being non-null if and only if $y \in \operatorname{im} f$. The set of non-null fibres of f is called the *coimage* of f and the map

$$\operatorname{dom} f \to \operatorname{coim} f ; x \mapsto f^{-1}\{f(x)\}$$

the *partition* of $\operatorname{dom} f$ induced by f. The fibres of a map f are sometimes called the *level sets* or the *contours* of f, especially when the target of f is the field of real numbers $\mathbf{R}$.

The *composite* gf of maps $f : X \rightarrowtail Y$ and $g : Y \rightarrowtail Z$ (read 'g following f') is the map $X \rightarrowtail Z ; x \mapsto g(f(x))$, with $\operatorname{dom} gf = f^{-1}(\operatorname{dom} g)$.

Proposition 1.1 *For any maps $f : W \to X$, $g : X \to Y$ and $h : Y \to Z$*

$$h(gf) = hgf = (hg)f.$$

Proof We traverse the *rebracketing pentagon* as follows:– for any $w \in W$

$$((h(gf))(w) = h((gf)(w)) = h(g(f(w))) = (hg)(f(w)) = ((hg)f)(w).$$

□

For any set X the *identity map* $X \to X : x \mapsto x$ will be denoted by 1_X, maps $f : X \to Y$ and $g : Y \to X$ such that $gf = 1_X$ and $fg = 1_Y$ being said to be *inverses* of each other, with $g = f^{-1}$ and $f = g^{-1}$. Given invertible maps $f : X \to Y$ and $g : Y \to Z$ then $(gf)^{-1} = f^{-1}g^{-1}$.

To any map $f : X \to Y$ there is associated an equation $f(x) = y$. The map f is said to be *surjective* or a *surjection* if, for each $y \in Y$, there is some $x \in X$ such that $f(x) = y$. It is said to be *injective* or an *injection* if, for each $y \in Y$, there is at most one element $x \in X$, though possibly none, such that $f(x) = y$. The map fails to be surjective if there exists an element $y \in Y$ such that the equation $f(x) = y$ has no solution $x \in X$, and fails to be injective if there exist $x, x' \in X$ such that $f(x') = f(x)$. A map that is both injective and surjective is said to be *bijective* or a *bijection*. A map is bijective if and only if it is invertible.

If maps $f : X \to Y$ and $g : Y \to X$ are such that $fg = 1_Y$ then, by Exercise 1.1, f is surjective and g is injective. The injection g is said to be a *section* of the surjection f. It selects for each $y \in Y$ a *single* $x \in X$ such that $f(x) = y$. It is assumed that any surjection $f : X \to Y$ has a section $g : Y \to X$, this assumption being known as the *axiom of choice*.

Linear spaces and maps

A *linear space* (of vectors), X, over a field (of scalars), $\mathbf{K}$, is an additive abelian group X, with zero element O (or 0) and furnished with a scalar multiplication $\mathbf{K} \times X \to X; (x, \lambda) \mapsto x\lambda = \lambda x$ satisfying both distributive laws, with $\lambda(\mu x) = (\lambda\mu)x$, for any $\lambda, \mu \in \mathbf{K}$ and any $x \in X$. Moreover, $1x = x$, for any $x \in X$, implying that $0x = O$ and $(-1)x = -x$. Also, for any $\lambda \in \mathbf{K}$, $\lambda O = O$. For us the field $\mathbf{K}$ will normally be either the real field $\mathbf{R}$ or the complex field $\mathbf{C}$ and the linear spaces will normally be *finite-dimensional*, a *basis* for such a space X being a finite set of vectors that are linearly independent and that span X. The number of vectors in any basis is independent of the basis and is called the *dimension* of X.

A *linear map* is a map between linear spaces that respects the linear structures; that is $f : X \to Y$ between linear spaces X and Y is *linear* if, for any $a, b \in X$, $\lambda, \mu \in \mathbf{K}$, $f(\lambda a + \mu b) = \lambda f(a) + \mu f(b)$. Such a map $f : X \to Y$ is uniquely determined by the action of f on any basis for X, and any assignment of f on the elements of a basis for X extends to a linear map of the whole of X to Y.

It is easily verified that the composite of any two composable linear maps is linear, while it follows as a corollary of Exercise 1.2 that the inverse of a linear bijection is linear.

A *linear subspace* of a linear space X is a subset W that acquires a linear structure by the restriction to W of the linear structure for X.

The *kernel* of a linear map $f : X \to Y$ is the set $\{x \in X : f(x) = 0\}$, written $\ker f$.

Proposition 1.2 *For any linear map $f : X \to Y$, $\ker f$ is a linear subspace of X and $\operatorname{im} f$ is a linear subspace of Y.*

Proposition 1.3 *A linear map f is injective if and only if $\ker f = \{0\}$.*

The *rank* of a linear map between finite-dimensional linear spaces is the dimension of the image of the map, this image being a linear subspace of the target space. The *kernel rank* or *nullity* of the map is the dimension of its kernel. The rank of the linear map $f : X \to Y$ will be denoted by $\operatorname{rk} f$ and the kernel rank by $\operatorname{kr} f$.

Proposition 1.4 *Let $f : X \to Y$ be a linear map, X and Y being finite-dimensional linear spaces. Then $\operatorname{rk} f + \operatorname{kr} f = \dim X$.*

In the case that X is a finite-dimensional linear space a linear map $f : X \to X$ is injective if and only if it is surjective, it then being an *automorphism* or self-isomorphism of X. More generally if X and Y are linear spaces of the same finite dimension then any linear injection $X \to Y$ is an isomorphism.

A linear space X is said to be the *direct sum* $X_0 \oplus X_1$ of linear subspaces X_0 and X_1 if $X_0 \cap X_1 = \{0\}$ and $X_0 + X_1 = X$, each of the subspaces then being a *linear complement* of the other. Then $\dim X_0 + \dim X_1 = \dim X$. Associated to any direct sum decomposition $X = X_0 \oplus X_1$ there are *projection maps* $X \to X_0$ with kernel X_1 and $X \to X_1$ with kernel X_0.

The *direct product* of linear spaces X and Y is the Cartesian product $X \times Y$, with the sum $(X \times Y)^2 \to X \times Y\,; ((x, y), (x', y')) \mapsto (x + x', y + y')$ and scalar product $\mathbf{K} \times (X \times Y) \to X \times Y\,; (\lambda, (x, y)) \mapsto (\lambda x, \lambda y)$, the linear space $\mathbf{K}^n$ being the n-fold direct power of the field $\mathbf{K}$.

The choice of a basis for an n-dimensional $\mathbf{K}$-linear space induces an isomorphism with the linear space $\mathbf{K}^n$, any linear map $\mathbf{K}^n \to \mathbf{K}^m$ being represented by its *matrix*, an array of real numbers with m rows and n

columns, whose columns are the images of the vectors of the standard ordered basis for $\mathbf{K}^n$, vectors in this context being represented by column matrices.

A non-zero kernel vector of a linear map $\mathbf{K}^n \to \mathbf{K}^m$ 'is' a linear dependence relation between the columns of the matrix of the map.

The *transpose* of a matrix with m rows and n columns is the matrix with n rows and m columns whose ith row is the ith column of the original matrix, for each i. Transposition will be denoted by τ, the transpose of a matrix a being denoted by a^τ.

A map $\beta : X \times Y \to Z$ is said to be *bilinear* if for any $a \in X$ and $b \in Y$ the maps $X \to Z; (x, b) \mapsto \beta(x, b)$ and $Y \to Z; y \mapsto \beta(a, y)$ are both linear. Scalar multiplication is an example of a bilinear map.

The set $L(X, Y)$ of linear maps between linear spaces X and Y, of dimensions n and m say, has a natural linear structure of dimension mn. In particular the linear space $L(X, \mathbf{K})$ of linear maps from X to $\mathbf{K}$, also of dimension n, is called the *dual* of X and will be denoted by X^L. The map

$$X \to (X^L)^L;\ x \mapsto \epsilon_x,$$

where $\epsilon_x(f) = f(x)$, is easily proved to be injective and so is an isomorphism.

The *dual* of a linear map $f : X \to Y$ between finite-dimensional linear spaces X and Y is the linear map

$$f^L : Y^L \to X^L;\ \omega \mapsto \omega f,$$

where ωf denotes the composite of the map $\omega : Y \to \mathbf{R}$ following the map $f : X \to Y$. Clearly, for composable linear maps $f : X \to Y$ and $g : W \to X$ we have $(fg)^L = g^L f^L$.

Proposition 1.5 *Let α and β be elements of the dual space X^L of a finite-dimensional real or complex linear space X such that* $\ker \alpha = \ker \beta$. *Then there exists a non-zero scalar λ such that $\beta = \lambda \alpha$.*

Proof Either $\ker \alpha = \ker \beta = X$, in which case $\alpha = \beta = 0$ and λ can be any non-zero element of $\mathbf{R}$, or both α and β are surjective. Then any element of X is of the form $a + \mu b$, where $\alpha(a) = 0$ and $\alpha(b) = 1$, and $\beta(a + \mu b) = \mu\,\beta(b) = \alpha(a + \mu b)\lambda$, where $\lambda = \beta(b)$; that is, $\beta = \lambda\,\alpha$. □

The *dual annihilator* $W^@$ of a linear subspace W of a finite-dimensional linear space X is the kernel of the map $\iota^L : X^L \to W^L$ dual to the inclusion map $\iota : W \to X$. Its dimension is equal to $\dim X - \dim W$.

A linear map $f : X \to X$ is said to be an *endomorphism* of the linear space X and the linear space $L(X, X)$ of all such endomorphisms of X is also denoted by $\operatorname{End} X$. As previously mentioned, a linear isomorphism $f : X \to X$ is said to be an *automorphism* of the linear space X and the set of all such automorphisms of X is denoted either by $GL(X)$ ('GL' standing for 'general linear') or by $\operatorname{Aut} X$. The linear space of all $n \times n$ matrices representing the elements of $L(\mathbf{K}^n, \mathbf{K}^n)$ will be denoted by $\mathbf{K}(n)$.

A *linear involution* of a linear space X is a linear map $t : X \to X$ such that $t^2 = 1_X$.

$^2\mathbf{K}$-*modules and maps*

Let Λ be a commutative and associative ring with unit element. Then a Λ-*module* X is a linear space over the ring Λ, the terminology *linear space* being reserved mainly for the case that Λ is a field.

Consider the case that Λ is the *double field* $^2\mathbf{K}$ consisting of the $\mathbf{K}$-linear space $\mathbf{K}^2$ assigned the product $(a, b)(c, d) = (ac, bd)$. A direct sum decomposition $X_0 \oplus X_1$ of a $\mathbf{K}$-linear space X may be regarded as a $^2\mathbf{K}$-module structure for X by setting

$$(\lambda, \mu)x = \lambda x_0 + \mu x_1, \quad \text{for all } x \in X \text{ and } (\lambda, \mu) \in {}^2\mathbf{K}.$$

Conversely, any $^2\mathbf{K}$-module structure for X determines a direct sum decomposition $X_0 \oplus X_1$ of X as a $\mathbf{K}$-linear space in which $X_0 = (1, 0)X$ $(= \{(1, 0)x : x \in X\})$ and $X_1 = (0, 1)X$.

Proposition 1.6 *Let $t : X \to X$ be a linear involution of the $\mathbf{K}$-linear space X. Then a $^2\mathbf{K}$-module structure, and therefore a direct sum decomposition, is defined for X by setting, for any $x \in X$,*

$$(1, 0)x = \frac{1}{2}(x + t(x)) \text{ and } (0, 1)x = \frac{1}{2}(x - t(x)).$$

$^2\mathbf{K}$-*module maps* and $^2\mathbf{K}$-*submodules* are defined in the obvious ways.

In working with a $^2\mathbf{K}$-module map $t : X \to Y$ it is often convenient to represent X and Y each as the *product* of its components and then to use notations associated with maps between products, as, for example, in the next proposition.

Proposition 1.7 *Let $t : X \to Y$ be a $^2\mathbf{K}$-module map. Then t is of the form $\begin{pmatrix} a_0 & 0 \\ 0 & a_1 \end{pmatrix}$, where $a_0 \in L(X_0, Y_0)$ and $a_1 \in L(X_1, Y_1)$. Conversely any map of this form is a $^2\mathbf{K}$-module map.*

A $^2\mathbf{K}$-module X such that $X_0 = (1, 0)X$ and $X_1 = (0, 1)X$ are isomorphic will be called a $^2\mathbf{K}$-*linear space* and a $^2\mathbf{K}$-module map $X \to Y$ between $^2\mathbf{K}$-linear spaces X and Y will be called a $^2\mathbf{K}$-*linear map*.

Affine spaces and maps

An *affine space* is a linear space with its origin deleted – it acquires a unique linear structure so soon as a point is chosen as origin, and the transfer from any one linear structure to any other is by a translation. An *affine map* is a map $f : X \to Y$ between affine spaces X and Y that becomes linear so soon as a point a of X is chosen as origin for X and $f(a)$ is chosen as origin for Y. An *affine map* between *linear spaces* is the sum of a linear map and a constant map. An *affine subspace* of a linear space is a translate or parallel of a linear subspace.

Determinants

We assume that the reader is familiar with the basic properties of determinants. Briefly, to any element a of $\mathbf{K}(n)$ there is a unique real number, the *determinant* of a, $\det a$, such that

(i) if any column of a matrix a is multiplied by a scalar λ then the determinant is multiplied by λ,
(ii) if any column of a matrix a is added to another then the determinant remains unaltered,
(iii) the determinant of the identity is 1.

The map is defined, for all $a \in \mathbf{K}(n)$, by the formula

$$\det a = \sum_{\pi \in \mathbf{n}!} \operatorname{sgn} \pi \prod_{j \in \mathbf{n}} a_{\pi(j),j},$$

where $\mathbf{n!}$ denotes the set of permutations of the set $\mathbf{n}$ of all natural numbers m such that $0 \leq m < n$. Moreover,

(iv) for any $a, b \in \mathbf{K}(n)$, $\det b\,a = \det b \det a$,
(v) for any invertible $a \in \mathbf{K}(n)$, $\det a^{-1} = (\det a)^{-1}$,
(vi) for any $a \in \mathbf{K}(n)$, a is invertible if and only if $\det a$ is invertible, that is, if and only if $\det a = 0$.

Any linear isomorphism $a : \mathbf{K}^n \to X$ induces a map

$$\operatorname{End} X = L(X, X) \to \mathbf{K};\ f \mapsto \det(a^{-1} f a)$$

called the *determinant* on End X and also denoted by det. This map is easily seen to be independent of the isomorphism a.

Proposition 1.8 *Any invertible matrix $a \in \mathbf{K}(n)$ is reducible by a series of column operations to a matrix with all entries on the main diagonal equal to* 1 *except for one which is equal to* det a, *and all entries off the main diagonal equal to zero.*

Note that each of the column operations may be performed by multiplying the matrix on the right by a matrix all of whose entries on the main diagonal are equal to 1 and all entries off the main diagonal except one are equal to zero.

Linear groups

For any linear space X the set Aut X has a natural group structure. The group Aut $\mathbf{K}^n$ is usually denoted by $GL(n;\mathbf{K})$ and called the *general linear group of degree n*. The subgroup of $GL(n;\mathbf{K})$ consisting of all automorphisms of $\mathbf{K}^n$ of determinant 1 is called the *special linear group of degree n*, denoted by $SL(n;\mathbf{K})$.

For any *real* linear space X there is a map $\zeta : \operatorname{End} X \to \{-1, 0, 1\}$, taking the value 1 if the determinant is positive, the value -1 if the determinant is negative and the value 0 if the determinant is zero. Automorphisms for which the value of ζ is equal to 1 are said to be *orientation-preserving*, while those for which the value is equal to -1 are said to be *orientation-reversing*. For any finite-dimensional linear space X the restriction of ζ to Aut X is a group isomorphism with the multiplicative group $S^0 = \{\pm 1\}$. The orientation-preserving automorphisms of X form a subgroup of Aut X which we shall denote by $\operatorname{Aut}^+ X$.

The question of orientation does not arise for complex linear spaces since there is no notion of a *positive* complex number.

Exercises

1.1 Let $f : X \to Y$ and $g : Y \to X$ be such that $fg = 1_Y$. Prove that f is surjective and that g is injective.

1.2 Let W, X and Y be linear spaces and let $t : X \to Y$ and $u : W \to X$ be maps whose composite $tu : W \to Y$ is linear.

Then

(a) if t is a linear injection, u is linear,
(b) if u is a linear surjection, t is linear.

1.3 Let X and Y be **K**-linear spaces. Prove that

$$X \times Y = (X \times \{0\}) \oplus (\{0\} \times Y).$$

1.4 Let **K** be either the field **R** or the field **C** and consider the product $\mathbf{K}^{2n} \times \mathbf{K}^{2n} \to \mathbf{K}$; $(x, y) \mapsto x \wedge y$ defined by

$$x \wedge y = \sum_{i \in \mathbf{n}} (x_i y_{n+i} - x_{n+i} y_i).$$

Verify that the product is bilinear and that, for every $x, y \in \mathbf{K}^{2n}$, $y \wedge x = -x \wedge y$.

Now define $\theta : \mathbf{K}(2n) \to \mathbf{K}$ by the formula

$$\theta(a) = \frac{1}{n!2^n} \sum_{\pi \in \mathbf{2n}!} \operatorname{sgn} \pi \prod_{i \in \mathbf{n}} a_{\pi(i)} \wedge a_{\pi(n+i)},$$

where $\mathbf{2n}!$ denotes the set of permutations of all the natural numbers m such that $0 \le m < 2n$.

Verify that this is an alternating $2n$-linear map on the columns of the matrix a, with $\theta({}^{2n}1) = 1$, where ${}^{2n}1$ denotes the unit $2n \times 2n$ matrix, and therefore that $\theta(a) = \det a$.

This exercise will be of use in Proposition 6.11.

2

Real and complex algebras

A *linear algebra* over the field of real numbers $\mathbf{R}$ is, by definition, a linear space A over $\mathbf{R}$ together with a bilinear map $A^2 \to A$, the *algebra product*.

Examples include $\mathbf{R}$ itself, the field of *complex numbers*, $\mathbf{C}$, consisting of the linear space $\mathbf{R}^2$ with the product $(a,b)(c,d) = (ac - bd, ad + bc)$, the *double field* ${}^2\mathbf{R}$ consisting of the linear space $\mathbf{R}^2$ with the product $(a,b)(c,d) = (ac, bd)$, and the *full matrix algebra* $\mathbf{R}(n)$ of all $n \times n$ matrices with real entries, with matrix multiplication as the product.

An algebra A may, or may not, have a unit element, and the product need be neither commutative nor associative, though it is usual to mention explicitly any failure of associativity. The unit element, if it exists, will normally be denoted by $1_{(A)}$ or simply, where no confusion need arise, by 1, the map $\mathbf{R} \to A; \lambda \mapsto \lambda 1_{(A)}$ being injective. (The notation 1_A is reserved for the identity map on A.)

All the above examples are associative and have a unit element, and all are commutative, with the exception of the matrix algebra $\mathbf{R}(n)$, with $n > 1$. The double field ${}^2\mathbf{R}$ is often identified with the subalgebra of $\mathbf{R}(2)$ consisting of the diagonal 2×2 matrices, the unit element being denoted by ^{21}. Likewise, for any n the n-fold power ${}^n\mathbf{R}$ of $\mathbf{R}$ may be identified with the subalgebra of the algebra $\mathbf{R}(n)$ consisting of the diagonal $n \times n$ matrices, the unit matrix in $\mathbf{R}(n)$ similarly being denoted by ${}^n1$.

Examples of non-associative algebras include the Cayley algebra and Lie algebras, discussed in later chapters.

Concepts defined in the obvious ways include not only *subalgebras* but *algebra maps*, *algebra-reversing maps*, and in particular *algebra isomorphisms* and *algebra anti-isomorphisms*, the latter for example being a linear isomorphism f of one algebra A to another B that reverses the order of multiplication, that is, for all $x, y \in A$, and all $\lambda, \mu \in \mathbf{R}$,

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y), \qquad \text{and} \qquad f(xy) = f(y)f(x).$$

The *centre* of a real algebra A is the set of all those elements of A that commute with each element of A. For example the centre of $\mathbf{R}(n)$ consists of all real multiples of the identity ${}^n1$. The centre of an algebra A is a subalgebra of A.

Analogous definitions hold for linear algebras not only over any field, in particular the field $\mathbf{C}$ of complex numbers, but also over the double fields ${}^2\mathbf{R}$ and ${}^2\mathbf{C}$. The part of a ${}^2\mathbf{R}$-linear space is played by an $\mathbf{R}$-linear space A with a prescribed direct sum decomposition $A_0 \oplus A_1$, where A_0 and A_1 are isomorphic linear subspaces of A (so, in the case that A is finite-dimensional, $\dim A_0 = \dim A_1$) with scalar multiplication defined by

$$((\lambda, \mu), (x_0 + x_1)) \mapsto (\lambda x_0 + \mu x_1).$$

For example, the elements of ${}^2\mathbf{R}(2)$ may be represented as matrices either of the form $\begin{pmatrix} a_0 & 0 & c_0 & 0 \\ 0 & a_1 & 0 & c_1 \\ b_0 & 0 & d_0 & 0 \\ 0 & b_1 & 0 & d_1 \end{pmatrix}$ or of the form $\begin{pmatrix} a_0 & c_0 & 0 & 0 \\ b_0 & d_0 & 0 & 0 \\ 0 & 0 & a_1 & c_1 \\ 0 & 0 & b_1 & d_1 \end{pmatrix}$.

Our preference is for the second form, where $\mathbf{R}^4 = \mathbf{R}^2 \times \{0\} \oplus \{0\} \times \mathbf{R}^2$ is thought of as $\mathbf{R}^2 \times \mathbf{R}^2$.

For any k with $0 \le k < n$ the kth and $(n+k)$th columns of ${}^2\mathbf{R}(n)$ will be said to be *partners*. Thus in the example the columns $\begin{pmatrix} c_0 \\ d_0 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ c_1 \\ d_1 \end{pmatrix}$ are partners.

Similar remarks apply to the algebra ${}^2\mathbf{C}(n)$.

Any algebra over $\mathbf{C}$, ${}^2\mathbf{R}$ or ${}^2\mathbf{C}$ may also be regarded as an algebra over $\mathbf{R}$, while any algebra over ${}^2\mathbf{C}$ may be regarded as an algebra over $\mathbf{C}$. In particular, as the corollary to the next proposition shows, the real algebra $\mathbf{C}$ is isomorphic to a subalgebra of the algebra $\mathbf{R}(2)$ of 2×2 real matrices.