1   Fields and vector spaces

In this section we revise some algebraic preliminaries and establish notation.

1.1   Division rings and fields

A *division ring*, or *skew field*, is a structure F with two binary operations called *addition* and *multiplication*, satisfying the following conditions:

(a) $(F, +)$ is an abelian group, with identity 0, called the *additive group* of F;

(b) $(F \setminus 0, \cdot)$ is a group, called the *multiplicative group* of F;

(c) left or right multiplication by any fixed element of F is an endomorphism of the additive group of F.

Note that condition (c) expresses the two distributive laws. Note that we must assume both, since one does not follow from the other.

   The identity element of the multiplicative group is called 1.

   A *field* is a division ring whose multiplication is commutative (that is, whose multiplicative group is abelian).

Exercise 1.1 Prove that the commutativity of addition follows from the other axioms for a division ring (that is, we need only assume that $(F, +)$ is a group in (a)).

Exercise 1.2 A *real quaternion* has the form $a + bi + cj + dk$, where $a, b, c, d \in \mathbb{R}$. Addition and multiplication are given by "the usual rules", together with the following rules for multiplication of the elements $1, i, j, k$:

$\cdot$	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

Prove that the set $\mathbb{H}$ of real quaternions is a division ring. (*Hint*: If $q = a + bi + cj + dk$, let $q^* = a - bi - cj - dk$; prove that $qq^* = a^2 + b^2 + c^2 + d^2$.)

Multiplication by zero induces the zero endomorphism of $(F,+)$. Multiplication by any non-zero element induces an automorphism (whose inverse is multiplication by the inverse element). In particular, we see that the automorphism group of $(F,+)$ acts transitively on its non-zero elements. So all non-zero elements have the same order, which is either infinite or a prime p. In the first case, we say that the *characteristic* of F is zero; in the second case, it has *characteristic* p.

The structure of the multiplicative group is not so straightforward. However, the possible finite subgroups can be determined. If F is a field, then any finite subgroup of the multiplicative group is cyclic. To prove this we require *Vandermonde's Theorem*:

Theorem 1.1 *A polynomial equation of degree n over a field has at most n roots.*

Exercise 1.3 Prove Vandermonde's Theorem. (*Hint*: If $f(a) = 0$, then $f(x) = (x-a)g(x)$.)

Theorem 1.2 *A finite subgroup of the multiplicative group of a field is cyclic.*

Proof An element ω of a field F is an *nth root of unity* if $\omega^n = 1$; it is a *primitive nth root of unity* if also $\omega^m \neq 1$ for $0 < m < n$.

Let G be a subgroup of order n in the multiplicative group of the field F. By Lagrange's Theorem, every element of G is an nth root of unity. If G contains a primitive nth root of unity, then it is cyclic, and the number of primitive nth roots is $\phi(n)$, where ϕ is Euler's function. If not, then of course the number of primitive nth roots is zero. The same considerations apply of course to any divisor of n. So, if $\psi(m)$ denotes the number of primitive mth roots of unity in G, then

(a) for each divisor m of n, either $\psi(m) = \phi(m)$ or $\psi(m) = 0$.

Now every element of G has some finite order dividing n; so

(b) $\displaystyle\sum_{m|n} \psi(m) = n$.

Finally, a familiar property of Euler's function yields:

(c) $\displaystyle\sum_{m|n} \phi(m) = n$.

From (a), (b) and (c) we conclude that $\psi(m) = \phi(m)$ for all divisors m of n. In particular, $\psi(n) = \phi(n) \neq 0$, and G is cyclic. ∎

For division rings, the position is not so simple, since Vandermonde's Theorem fails.

Exercise 1.4 Find all solutions of the equation $x^2 + 1 = 0$ in $\mathbb{H}$.

However, the possibilities can be determined. Let G be a finite subgroup of the multiplicative group of the division ring F. We claim that there is an abelian group A such that G is a group of automorphisms of A acting semiregularly on the non-zero elements. Let B be the subgroup of $(F, +)$ generated by G. Then B is a finitely generated abelian group admitting G acting semiregularly. If F has non-zero characteristic, then B is elementary abelian; take $A = B$. Otherwise, choose a prime p such that, for all $x, g \in G$, the element $(xg - x)p^{-1}$ is not in B, and set $A = B/pB$.

The structure of semiregular automorphism groups of finite groups (a.k.a. *Frobenius complements*) was determined by Zassenhaus. See Passman, *Permutation Groups*, Benjamin, New York, 1968, for a detailed account. In particular, either G is metacyclic, or it has a normal subgroup isomorphic to $\mathrm{SL}(2,3)$ or $\mathrm{SL}(2,5)$. (These are finite groups G having a unique subgroup Z of order 2, such that G/Z is isomorphic to the alternating group A_4 or A_5 respectively. There is a unique such group in each case.)

Exercise 1.5 Identify the division ring $\mathbb{H}$ of real quaternions with the real vector space $\mathbb{R}^4$ with basis $\{1, \mathrm{i}, \mathrm{j}, \mathrm{k}\}$. Let U denote the multiplicative group of *unit quaternions*, those elements $a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k}$ satisfying $a^2 + b^2 + c^2 + d^2 = 1$. Show that conjugation by a unit quaternion is an orthogonal transformation of $\mathbb{R}^4$, fixing the 1-dimensional space spanned by 1 and inducing an orthogonal transformation on the 3-dimensional subspace spanned by $\mathrm{i}, \mathrm{j}, \mathrm{k}$.

Prove that the map from U to the 3-dimensional orthogonal group has kernel ± 1 and image the group of rotations of 3-space (orthogonal transformations with determinant 1).

Hence show that the groups $\mathrm{SL}(2,3)$ and $\mathrm{SL}(2,5)$ are finite subgroups of the multiplicative group of $\mathbb{H}$.

Remark: This construction explains why the groups $\mathrm{SL}(2,3)$ and $\mathrm{SL}(2,5)$ are sometimes called the *binary tetrahedral* and *binary icosahedral* groups. Construct also a *binary octahedral* group of order 48, and show that it is not isomorphic to $\mathrm{GL}(2,3)$ (the group of 2×2 invertible matrices over the integers mod 3), even though both groups have normal subgroups of order 2 whose factor groups are isomorphic to the symmetric group S_4.

1.2 Finite fields

The basic facts about finite fields are summarised in the following two theorems, due to Wedderburn and Galois respectively.

Theorem 1.3 *Every finite division ring is commutative.*

Theorem 1.4 *The number of elements in a finite field is a prime power. Conversely, if q is a prime power, then there is a unique field with q elements, up to isomorphism.*

The unique finite field with a given prime power order q is called the *Galois field* of order q, and denoted by $\mathrm{GF}(q)$ (or sometimes $\mathbb{F}_q$). If q is prime, then $\mathrm{GF}(q)$ is isomorphic to $\mathbb{Z}/q\mathbb{Z}$, the integers mod q.

We now summarise some results about $\mathrm{GF}(q)$.

Theorem 1.5 *Let $q = p^a$, where p is prime and a is a positive integer. Let $F = \mathrm{GF}(q)$.*

(a) F has characteristic p, and its additive group is an elementary abelian p-group.

(b) The multiplicative group of F is cyclic, generated by a primitive $(p^a - 1)$th root of unity (called a primitive element *of F).*

(c) The automorphism group of F is cyclic of order a, generated by the Frobenius automorphism $x \mapsto x^p$.

(d) For every divisor b of a, there is a unique subfield of F of order p^b, consisting of all solutions of $x^{p^b} = x$; and these are all the subfields of F.

Proof Part (a) is obvious since the additive group contains an element of order p, and part (b) follows from Theorem 1.2. Parts (c) and (d) are most easily proved using Galois theory. Let E denote the subfield $\mathbb{Z}/p\mathbb{Z}$ of F. Then the degree of F over E is a. The Frobenius map $\sigma : x \mapsto x^p$ is an E-automorphism of F, and has order a; so F is a Galois extension of E, and σ generates the Galois group. Now subfields of F necessarily contain E; by the Fundamental Theorem of Galois Theory, they are the fixed fields of subgroups of the Galois group $\langle \sigma \rangle$. ∎

For explicit calculation in $F = \mathrm{GF}(p^a)$, it is most convenient to represent it as $E[x]/(f)$, where $E = \mathbb{Z}/p\mathbb{Z}$, $E[x]$ is the polynomial ring over E, and f is the (irreducible) minimum polynomial of a primitive element of F. If α denotes the coset $(f) + x$, then α is a root of f, and hence a primitive element.

Now every element of F can be written uniquely in the form

$$c_0 + c_1\alpha + \cdots + c_{a-1}\alpha^{a-1},$$

where $c_0, c_1, \ldots, c_{a-1} \in E$; addition is straightforward in this representation. Also, every non-zero element of F can be written uniquely in the form α^m, where $0 \le m < p^a - 1$, since α is primitive; multiplication is straightforward in this representation. Using the fact that $f(\alpha) = 0$, it is possible to construct a table matching up the two representations.

Example The polynomial $x^3 + x + 1$ is irreducible over $E = \mathbb{Z}/2\mathbb{Z}$. So the field $F = E(\alpha)$ has eight elements, where α satisfies $\alpha^3 + \alpha + 1 = 0$ over E. We have $\alpha^7 = 1$, and the table of logarithms is as follows:

α^0	1
α^1	α
α^2	α^2
α^3	$\alpha + 1$
α^4	$\alpha^2 + \alpha$
α^5	$\alpha^2 + \alpha + 1$
α^6	$\alpha^2 + 1$

Hence

$$(\alpha^2 + \alpha + 1)(\alpha^2 + 1) = \alpha^5 \cdot \alpha^6 = \alpha^4 = \alpha^2 + \alpha.$$

Exercise 1.6 Show that there are three irreducible polynomials of degree 4 over the field $\mathbb{Z}/2\mathbb{Z}$, of which two are primitive. Hence construct $\mathrm{GF}(16)$ by the method outlined above.

Exercise 1.7 Show that an irreducible polynomial of degree m over $\mathrm{GF}(q)$ has a root in $\mathrm{GF}(q^n)$ if and only if m divides n.

Hence show that the number a_m of irreducible polynomials of degree m over $\mathrm{GF}(q)$ satisfies

$$\sum_{m \mid n} m a_m = q^n.$$

6

Exercise 1.8 Show that, if q is even, then every element of $GF(q)$ is a square; while, if q is odd, then half of the non-zero elements of $GF(q)$ are squares and half are non-squares.

Elements If q is odd, show that -1 is a square in $GF(q)$ if and only if $q \equiv 1 \pmod 4$.

1.3 Vector spaces

A *left vector space* over a division ring F is a unital left F-module. That is, it is an abelian group V, with a anti-homomorphism from F to $\text{End}(V)$ mapping 1 to the identity endomorphism of V.

Writing scalars on the left, we have $(cd)v = c(dv)$ for all $c, d \in F$ and $v \in V$: that is, scalar multiplication by cd is the same as multiplication by d followed by multiplication by c, not vice versa. (The opposite convention would make V a right (rather than left) vector space; scalars would more naturally be written on the right.) The unital condition simply means that $1v = v$ for all $v \in V$.

Note that F is a vector space over itself, using field multiplication for the scalar multiplication.

If F is a division ring, the *opposite* division ring F° has the same underlying set as F and the same addition, with multiplication given by

$$a \circ b = ba.$$

Now a right vector space over F can be regarded as a left vector space over F°.

A *linear transformation $T : V \to W$* between two left F-vector spaces V and W is a vector space homomorphism; that is, a homomorphism of abelian groups which commutes with scalar multiplication. We write linear transformations on the right, so that we have

$$(cv)T = c(vT)$$

for all $c \in F$, $v \in V$. We add linear transformations, or multiply them by scalars, pointwise (as functions), and multiply then by function composition; the results are again linear transformations.

If a linear transformation T is one-to-one and onto, then the inverse map is also a linear transformation; we say that T is *invertible* if this occurs.

Now $\text{Hom}(V, W)$ denotes the set of all linear transformations from V to W. The *dual space* of F is $F^* = \text{Hom}(V, F)$.

Exercise 1.9 Show that V^* is a right vector space over F.

A vector space is *finite-dimensional* if it is finitely generated as F-module. A *basis* is a minimal generating set. Any two bases have the same number of elements; this number is usually called the dimension of the vector space, but in order to avoid confusion with a slightly different geometric notion of dimension, I will call it the *rank* of the vector space. The rank of V is denoted by $\mathrm{rk}(V)$.

Every vector can be expressed uniquely as a linear combination of the vectors in a basis. In particular, a linear combination of basis vectors is zero if and only if all the coefficients are zero. Thus, a vector space of rank n over F is isomorphic to F^n (with coordinatewise addition and scalar multiplication).

I will assume familiarity with standard results of linear algebra about ranks of sums and intersections of subspaces, about ranks of images and kernels of linear transformations, and about the representation of linear transformations by matrices with respect to given bases.

As well as linear transformations, we require the concept of a *semilinear transformation* between F-vector spaces V and W. This can be defined in two ways. It is a map T from V to W satisfying

(a) $(v_1 + v_2)T = v_1 T + v_2 T$ for all $v_1, v_2 \in V$;

(b) $(cv)T = c^\sigma v T$ for all $c \in F$, $v \in V$, where σ is an automorphism of F called the *associated automorphism* of T.

Note that, if T is not identically zero, the associated automorphism is uniquely determined by T.

The second definition is as follows. Given an automorphism σ of F, we extend the action of σ to F^n coordinatewise:

$$(c_1, \ldots, c_n)^\sigma = (c_1^\sigma, \ldots, c_n^\sigma).$$

Hence we have an action of σ on any F-vector space with a given basis. Now a σ-*semilinear transformation* from V to W is the composition of a linear transformation from V to W with the action of σ on W (with respect to some basis).

The fact that the two definitions agree follows from the observations

- the action of σ on F^n is semilinear in the first sense;

- the composition of semilinear transformations is semilinear (and the associated automorphism is the composition of the associated automorphisms of the factors).

8

This immediately shows that a semilinear map in the second sense is semilinear in the first. Conversely, if T is semilinear with associated automorphism σ, then the composition of T with σ^{-1} is linear, so T is σ-semilinear.

Exercise 1.10 Prove the above assertions.

If a semilinear transformation T is one-to-one and onto, then the inverse map is also a semilinear transformation; we say that T is *invertible* if this occurs.

Almost exclusively, I will consider only finite-dimensional vector spaces. To complete the picture, here is the situation in general. In ZFC (Zermelo–Fraenkel set theory with the Axiom of Choice), every vector space has a basis (a set of vectors with the property that every vector has a unique expression as a linear combination of a *finite* set of basis vectors with non-zero coefficients), and any two bases have the same cardinal number of elements. However, without the Axiom of Choice, there may exist a vector space which has no basis.

Note also that there exist division rings F with bimodules V such that V has different ranks when regarded as a left or a right vector space.

1.4 Projective spaces

It is not easy to give a concise definition of a projective space, since projective geometry means several different things: a geometry with points, lines, planes, and so on; a topological manifold with a strange kind of torsion; a lattice with meet, join, and order; an abstract incidence structure; a tool for computer graphics.

Let V be a vector space of rank $n+1$ over a field F. The "objects" of the n-dimensional projective space are the subspaces of V, apart from V itself and the zero subspace $\{0\}$. Each object is assigned a dimension which is one less than its rank, and we use geometric terminology, so that *points*, *lines* and *planes* are the objects of dimension 0, 1 and 2 (that is, rank 1, 2, 3 respectively). A *hyperplane* is an object having codimension 1 (that is, dimension $n-1$, or rank n). Two objects are *incident* if one contains the other. So two objects of the same dimension are incident if and only if they are equal.

The n-dimensional projective space is denoted by $\mathrm{PG}(n,F)$. If F is the Galois field $\mathrm{GF}(q)$, we abbreviate $\mathrm{PG}(n,\mathrm{GF}(q))$ to $\mathrm{PG}(n,q)$. A similar convention will be used for other geometries and groups over finite fields.

A 0-dimensional projective space has no internal structure at all, like an idealised point. A 1-dimensional projective space is just a set of points, one more than the number of elements of F, with (at the moment) no further structure. (If
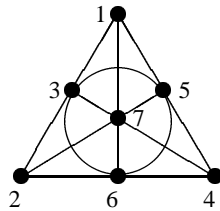
$\{e_1, e_2\}$ is a basis for V, then the points are spanned by the vectors $\lambda e_1 + e_2$ (for $\lambda \in F$) and e_1.)

For $n > 1$, $\mathrm{PG}(n, F)$ contains objects of different dimensions, and the relation of incidence gives it a non-trivial structure.

Instead of our "incidence structure" model, we can represent a projective space as a collection of subsets of a set. Let S be the set of points of $\mathrm{PG}(n, F)$. The *point shadow* of an object U is the set of points incident with U. Now the point shadow of a point P is simply $\{P\}$. Moreover, two objects are incident if and only if the point shadow of one contains that of the other.

The diagram below shows $\mathrm{PG}(2, 2)$. It has seven points, labelled 1, 2, 3, 4, 5, 6, 7; the line shadows are 123, 145, 167, 246, 257, 347 356 (where, for example, 123 is an abbreviation for $\{1, 2, 3\}$).

The correspondence between points and spanning vectors of the rank-1 subspaces can be taken as follows:

1	2	3	4	5	6	7
$(0,0,1)$	$(0,1,0)$	$(0,1,1)$	$(1,0,0)$	$(1,0,1)$	$(1,1,0)$	$(1,1,1)$

The following geometric properties of projective spaces are easily verified from the rank formulae of linear algebra:

(a) Any two distinct points are incident with a unique line.

(b) Two distinct lines contained in a plane are incident with a unique point.

(c) Any three distinct points, or any two distinct collinear lines, are incident with a unique plane.

(d) A line not incident with a given hyperplane meets it in a unique point.

(e) If two distinct points are both incident with some object of the projective space, then the unique line incident with them is also incident with that object.

10

Exercise 1.11 Prove the above assertions.

It is usual to be less formal with the language of incidence, and say "the point P lies on the line L", or "the line L passes through the point P" rather than "the point P and the line L are incident". Similar geometric language will be used without further comment.

An *isomorphism* from a projective space Π_1 to a projective space Π_2 is a map from the objects of Π_1 to the objects of Π_2 which preserves the dimensions of objects and also preserves the relation of incidence between objects. A *collineation* of a projective space Π is an isomorphism from Π to Π.

The important theorem which connects this topic with that of the previous section is the *Fundamental Theorem of Projective Geometry*:

Theorem 1.6 *Any isomorphism of projective spaces of dimension at least two is induced by an invertible semilinear transformation of the underlying vector spaces. In particular, the collineations of $\mathrm{PG}(n,F)$ for $n \geq 2$ are induced by invertible semilinear transformations of the rank-$(n+1)$ vector space over F.*

This theorem will not be proved here, but I make a few comments about the proof. Consider first the case $n = 2$. One shows that the field F can be recovered from the projective plane (that is, the addition and multiplication in F can be defined by geometric constructions involving points and lines). The construction is based on choosing four points of which no three are collinear. Hence any collineation fixing these four points is induced by a field automorphism. Since the group of invertible linear transformations acts transitively on quadruples of points with this property, it follows that any collineation is induced by the composition of a linear transformation and a field automorphism, that is, a semilinear transformation.

For higher-dimensional spaces, we show that the coordinatisations of the planes fit together in a consistent way to coordinatise the whole space.

In the next chapter we study properties of the collineation group of projective spaces. Since we are concerned primarily with groups of matrices, I will normally speak of $\mathrm{PG}(n-1,F)$ as the projective space based on a vector space of rank n, rather than $\mathrm{PG}(n,F)$ based on a vector space of rank $n+1$.

Next we give some numerical information about finite projective spaces.

Theorem 1.7 *(a) The number of points in the projective space $\mathrm{PG}(n-1,q)$ is $(q^n - 1)/(q-1)$.*

11

(b) More generally, the number of $(m-1)$-dimensional subspaces of $\mathrm{PG}(n-1,q)$ is

$$\frac{(q^n-1)(q^n-q)\cdots(q^n-q^{m-1})}{(q^m-1)(q^m-q)\cdots(q^m-q^{m-1})}.$$

(c) The number of $(m-1)$-dimensional subspaces of $\mathrm{PG}(n-1,q)$ containing a given $(l-1)$-dimensional subspace is equal to the number of $(m-l-1)$-dimensional subspaces of $\mathrm{PG}(n-l-1,q)$.

Proof (a) The projective space is based on a vector space of rank n, which contains q^n vectors. One of these is the zero vector, and the remaining q^n-1 each span a subspace of rank 1. Each rank 1 subspace contains $q-1$ non-zero vectors, each of which spans it.

(b) Count the number of linearly independent m-tuples of vectors. The jth vector must lie outside the rank $(j-1)$ subspace spanned by the preceding vectors, so there are q^n-q^{j-1} choices for it. So the number of such m-tuples is the numerator of the fraction. By the same argument (replacing n by m), the number of linearly independent m-tuples which span a given rank m subspace is the denominator of the fraction.

(c) If U is a rank l subspace of the rank m vector space V, then the Second Isomorphism Theorem shows that there is a bijection between rank m subspaces of V containing U, and rank $(m-l)$ subspaces of the rank $(n-l)$ vector space V/U. ∎

The number given by the fraction in part (b) of the theorem is called a *Gaussian coefficient*, written $\begin{bmatrix} n \\ m \end{bmatrix}_q$. Gaussian coefficients have properties resembling those of binomial coefficients, to which they tend as $q \to 1$.

Exercise 1.12   (a) Prove that

$$\begin{bmatrix} n \\ k \end{bmatrix}_q + q^{n-k+1}\begin{bmatrix} n \\ k-1 \end{bmatrix}_q = \begin{bmatrix} n+1 \\ k \end{bmatrix}_q.$$

(b) Prove that for $n \geq 1$,

$$\prod_{i=0}^{n-1}(1+q^i x) = \sum_{k=0}^{n} q^{k(k-1)/2}\begin{bmatrix} n \\ k \end{bmatrix}_q x^k.$$

(This result is known as the *q-binomial theorem*, since it reduces to the binomial theorem as $q \to 1$.)

If we regard a projective space $\mathrm{PG}(n-1,F)$ purely as an incidence structure, the dimensions of its objects are not uniquely determined. This is because there is an additional symmetry known as *duality*. That is, if we regard the hyperplanes as points, and define new dimensions by $\dim^*(U) = n-2-\dim(U)$, we again obtain a projective space, with the same relation of incidence. The reason that it is a projective space is as follows.

Let $V^* = \mathrm{Hom}(V,F)$ be the dual space of V, where V is the underlying vector space of $\mathrm{PG}(n-1,F)$. Recall that V^* is a right vector space over F, or equivalently a left vector space over the opposite field F°. To each subspace U of V, there is a corresponding subspace $U^\dagger$ of V^*, the *annihilator* of U, given by

$$U^\dagger = \{f \in V^* : uf = 0 \text{ for all } u \in U\}.$$

The correspondence $U \mapsto U^\dagger$ is a bijection between the subspaces of V and the subspaces of V^*; we denote the inverse map from subspaces of V^* to subspaces of V also by $\dagger$. It satisfies

(a) $(U^\dagger)^\dagger = U$;

(b) $U_1 \leq U_2$ if and only if $U_1^\dagger \geq U_2^\dagger$;

(c) $\mathrm{rk}(U^\dagger) = n - \mathrm{rk}(U)$.

Thus we have:

Theorem 1.8 *The dual of* $\mathrm{PG}(n-1,F)$ *is the projective space* $\mathrm{PG}(n-1,F^\circ)$. *In particular, if $n \geq 3$, then* $\mathrm{PG}(n-1,F)$ *is isomorphic to its dual if and only if F is isomorphic to its opposite F°.*

Proof   The first assertion follows from our remarks. The second follows from the first by use of the Fundamental Theorem of Projective Geometry.   ∎

Thus, $\mathrm{PG}(n-1,F)$ is self-dual if F is commutative, and for some non-commutative division rings such as $\mathbb{H}$; but there are division rings F for which $F \not\cong F^\circ$.

An isomorphism from F to its opposite is a bijection σ satisfying

$$\begin{aligned}
(a+b)^\sigma &= a^\sigma + b^\sigma, \\
(ab)^\sigma &= b^\sigma a^\sigma,
\end{aligned}$$

for all $a,b \in F$. Such a map is called an *anti-automorphism* of F.

Exercise 1.13 Show that $\mathbb{H} \cong \mathbb{H}^\circ$. (*Hint*: $(a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k})^\sigma = a - b\mathrm{i} - c\mathrm{j} - d\mathrm{k}$.)