1. FINITE-DIMENSIONAL VECTOR SPACES

§1.1. Axioms

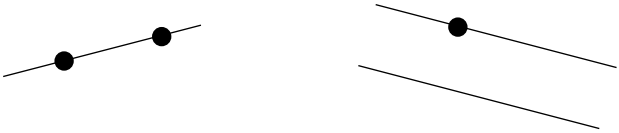
This may be your first encounter with an axiomatic system. So, here's some general discussion about what it means to define a mathematical structure by axioms.

The father of axioms was Euclid. He developed Euclidean Geometry by starting with some axioms. For him, axioms were self-evident truths, such as:

- **Every two distinct points lie on exactly one line.**

and

- **Given a line, and a point that does not lie on that line, there is exactly one line through the given point parallel to the given line.**



Now Euclid was attempting to make a mathematical model of a plane in our real universe and in fact, these axioms are far from self-evident.

Well, they seem reasonable enough for lines drawn on a sheet of paper, or in the sand. But what if there are many lines through the same pair of points? What if there is a line through a given point almost parallel but not one

exactly parallel. If ‘parallel’ means never intersect, how can we tell that the line we draw satisfying the assumptions in the second scenario, might not meet the first line if both lines were extended a billion miles.

In fact, in the nineteenth century, some geometers played with the second of these axioms and constructed a couple of non-Euclidean geometries, Elliptic Geometry and Hyperbolic Geometry. Although Euclidean Geometry, in its 3-dimensional version, is a quite accurate model for our spatial world, some cosmologists believe that one or other of these non-Euclidean geometries is more accurate for the whole cosmos.

Axioms, in the sense that we are using the term here, arose towards the end of the eighteenth century. Early in that century, the French mathematician Galois invented the concept of a group, and developed a considerable amount of the theory of groups.

Now Galois was concerned with polynomials, and their zeros (values of x that make the polynomial $P(x)$ equal to zero). For him the group consisted of rearrangements of the zeros of the polynomial.

Then somebody realised that the things being permuted didn’t have to have anything to do with polynomials, and so the groups were groups of ‘substitutions’ or, as we would say, rearrangements of any collection of numbers.

A considerable amount of the theory of Permutation Groups was developed. But somebody once noticed that everything could be proved from just four axioms. Group Theory was the first branch of mathematics to become abstract.

The modern concept of a group is that it consists of a set of ‘things’. They could be numbers, or polynomials, or matrices, or permutations, or even the four different ways of rotating a mattress.

There is assumed to be an undefined operation, called ‘multiplication’ that combines any two of these things to produce one of them.

For any group we need to know what the elements are and how $x * y$ is defined. If the elements of the group are numbers we might define $x * y$ to be ordinary multiplication. Or it could be ordinary addition. There is a group of real numbers where $x * y$ is defined by:

$$x * y = x + y + xy.$$

If the elements of the group are permutations we may define the product to be the result of performing one permutation followed by another (or even, for $x * x$, the operation of performing x twice.

If there is any danger of confusing $x * y$ with ordinary multiplication we need to continue using this rather cumbersome notation. But otherwise, we write it as xy.

Now, as I said, for a specific example of a group we need a definition of the set of things, as well as a definition of multiplication. But if we are developing the

theory we can consider the set to consist of undefined 'things' and the multiplication to be undefined. However we need to make certain assumptions in order to prove theorems. These we call the group axioms. Far from being 'self-evident truths' they are part of the definition of a group.

A group is a set G, containing a specific element, denoted by 1, a function $x \to x^{-1}$ from G to G, and a binary operation xy, such that the following 4 axioms hold:

THE GROUP AXIOMS

(1) (Closure under multiplication): $xy \in G$.
(2) (Associativity under multiplication): $x(yz) = (xy)z$.
(3) (Identity under multiplication): $1x = x = x1$.
(4) (Inverses under multiplication): $xx^{-1} = 1 = x^{-1}x$.

The whole of Group Theory can be developed using only these 4 axioms. Any theorem we prove will hold for any group, no matter what the elements are or what the operation is – provided these axioms hold.

NOTES:

(1) Axiom (1) is really redundant, because the fact that $x * y \in G$ is part of the definition of a binary operation. However it is included for emphasis.
(2) The Associative Law is important because, without it, the expression xyz would be ambiguous. Moreover

powers of x, where x^n denotes $x * x * \ldots * x$, with n factors, would be ambiguous and would depend on the order in which the operations are carried out.
Is $x^4 = ((x * x) * x) * x$ or $(x * (x * x)) * x$ or
$x * ((x * x) * x)$ or $x * (x * (x * x))$?
Without the associative law these might all be different.

(3) If G is a group of numbers, don't assume that 1 represents the number 1. In some groups it might be 0 – in others it could be − 1. If there is danger of confusion we write the identity as e. And by 'e' I don't mean the special exponential number.

(4) We never write x^{-1} as $\frac{1}{x}$ because fractions can be ambiguous. Does $\frac{y}{x}$ mean $x^{-1}y$, or yx^{-1}? Without the commutative law these could be different.

There's a 5th axiom that can be added, which gives rise to the definition of a commutative group.

(5) (Commutativity under multiplication): $xy = yx$.

A **commutative group** is one that satisfies all 5 axioms. Commutative groups are more usually called **abelian** groups, after the Norwegian mathematician Abel. The theory of abelian groups is a subset of Group Theory as a whole.

If you want to explore the rich world of Group Theory you can find an account in my notes on Group Theory.

The axiomatic approach proved so successful that it has crept into most branches of mathematics. In abstract algebra we have groups, semigroups, rings, fields, vector spaces. Topology begins with the axioms for a topological space. Set Theory starts with the ZF axioms. Different geometries usually begin with a set of axioms. Calculus has gone abstract at advanced levels. There we call the study Analysis, and we might begin a course in Analysis with the axioms for a Hilbert Space.

§1.2. Fields

By now you'll have acquired a fair knowledge of matrices. These are a concrete embodiment of something rather more abstract. Sometimes it's easier to use matrices, but at other times the abstract approach allows us more freedom.

A **field** F, is a set, together with two operations, addition (written $x + y$) and multiplication (written xy) such that the following properties (called the field axioms) hold. In addition there are two special elements of F, 0 and 1, which are not equal. Also, for all $x \in$ F

there are elements $-x$ and, unless $x = 0$, x^{-1}. All of these properties are required to hold for all $x, y, z \in F$.

FIELD AXIOMS

(1) (Closure under addition): $x + y \in F$.

(2) (Associativity under addition):

$$x + (y + z) = (x + y) + z.$$

(3) (Commutativity under addition): $x + y = y + x$.

(4) (Identity under addition): $x + 0 = x = 0 + x$

(5) (Inverses under addition): $x + (-x) = 0 = (-x) + x$.

(6) (Closure under multiplication): $xy \in F$.

(7) (Associativity under multiplication): $x(yz) = (xy)z$.

(8) (Commutativity under multiplication): $xy = yx$.

(9) (Identity under multiplication): $1x = x = x1$.

(10) (Inverses under multiplication): $xx^{-1} = 1 = x^{-1}x$.

(11) (Distributivity):

$$x(y + z) = xy + xz \text{ and } (x + y)z = xz + yz.$$

NOTES:

(1) We insist that 0 and 1 are distinct. So the smallest possible field has 2 elements.

(2) The existence of a multiplicative inverse only extends to non-zero elements. If you allowed 0^{-1} you would be able to prove that all elements were equal, which doesn't give any useful examples.

(3) Several axioms have two equalities, one of which is redundant because of one of the commutative laws. For

example we say $xx^{-1} = 1$, as well as $1 = x^{-1}x$. The reason for this is so that we can consider the axioms independently. There are algebraic systems where the commutative law for multiplication doesn't hold, yet both halves of the inverse law for multiplication do.

(4) Axioms (6) – (10) mimic axioms (1) – (5), though there is a subtle difference between axioms (5) and (10) in so far as axiom (10) only holds for non-zero x.
(5) Axioms (1) – (5) are the axioms for a commutative group. So are axioms (6) – (10), but with different notation. Together they say that a field is a commutative group under addition and the non-zero elements are a commutative group under multiplication.

But it is the distributive law that binds the two structures together it. Without it we would have two completely separate structures that happen to live in the same body, like Dr Jekyll and Mr Hyde.

(6) There are other laws that we insist on for all fields, but they are not listed as axioms because they are consequences of these 11 axioms. Such a law is the one that says that $0x = 0 = x0$.

Example 1: The following are examples of fields:

$\mathbb{Q}$, the set of rational numbers.
$\mathbb{R}$, the set of real numbers.
$\mathbb{C}$, the set of complex numbers.
$\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$.

The associative, commutative and distributive laws hold throughout the system of complex numbers, so the only axioms that need to be checked are the closure, identity and inverse laws. For $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ these are obvious. Let us check them for $\mathbb{Q}[\sqrt{2}]$.

$(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2}$, so $\mathbb{Q}[\sqrt{2}]$ is closed under addition.
$(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2}$, so $\mathbb{Q}[\sqrt{2}]$ is closed under addition.
$\mathbb{Q}[\sqrt{2}]$ clearly contain both 0 and 1 so the identity laws hold.

$$-(a + b\sqrt{2}) = (-a) + (-b)\sqrt{2} \text{ and}$$

$$\frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{a^2 - 2b^2} = \frac{a}{a^2 - 2b^2} + \left(\frac{-b}{a^2 - 2b^2}\right)\sqrt{2}$$

so $\mathbb{Q}[\sqrt{2}]$ is closed under both additive and multiplicative inverses.

There are certain properties of field that are consequences of these field axioms. For example there's no axiom that says that $0x = 0$. However if we consider that $0x = (0 + 0)x = 0x + 0x$ we reach this conclusion (with the help of the additive identity axiom and the distributive law).

The cancellation law that we constantly use in basic algebra is not one of the 11 axioms but is a consequence of them. If $ab = 0$ then $a = 0$ or $b = 0$. For if $a \neq 0$ then a^{-1} exists and if $ab = 0$ then $0 = a^{-1}(ab) = (a^{-1}a)b = 1b = b$.

Theorem 1: If p is prime then $\mathbb{Z}_p = \{0, 1, 2, \dots, p-1\}$ is a field under addition and multiplication modulo p (where we add and multiply normally but then take the remainder on dividing by p).

Proof: All the field axioms are obvious except for inverses under multiplication.

Suppose x is a non-zero element of $\mathbb{Z}_p$. Regarding x as an integer this means that x is not divisible by p. In other words, x is coprime to p. This means that $ax + bp = 1$ for some integers a, b. Now, interpreting this modulo p this becomes $ax = 1$, so x has an inverse under multiplication. ✋☺

Example 2: In $\mathbb{Z}_{13}$ the following table gives the inverses under addition and multiplication for the non-zero elements.

x	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
–x	12	11	10	9	8	7	6	5	4	3	2	1
$\mathbf{x^{-1}}$	1	7	9	10	8	11	2	5	3	4	6	12

Integers that are 1 plus a multiple of 13 include
14, 27, 40, 53, 66, 79 and 92.

So, for example, $10 \times 4 = 40 = 1 \bmod 17$ and so 10 and 4 are inverses of one another under multiplication.

Then, since $7 = -6$, $7^{-1} = -6^{-1} = -11 = 2$.

If p is not prime, however, $\mathbb{Z}_p$ is not a field because, if $p = ab$ for some integers a, b where $1 < a, b < p$, then modulo p we would have $ab = 0$ while $a \neq 0$ and $b \neq 0$.

So the only systems of integers modulo p that give fields are those where p is prime. But these are not the only finite fields. For every prime power p^n there exists a field with p^n elements (and for no other sizes). Tthe field of order p^n (there is only one) is not $\mathbb{Z}_{p^n}$, unless $n = 1$.

Example 3: The following are the addition and multiplication tables for the field with 4 elements:

+	0	1	2	3
0	0	1	2	3
1	1	0	3	2
2	2	3	0	1
3	3	2	1	0

×	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	3	1
3	0	3	1	2

Now let us look at a few properties that we insist on for fields, but which are consequences of the 11 axioms and therefore don't have to be included among the axioms.

Theorem 2: If $x + y = x + z$ then $y = z$.
Proof: Suppose that $x + y = x + z$.
$\therefore -x + (x + y) = -x + (x + z)$
$\therefore (-x + x) + y = (-x + x) + z$ by Axiom (2)
$\therefore 0 + y = 0 + z$ by axiom (5)
$\therefore y = z$ by axiom (4)

Theorem 3: $0x = 0 = x0$ for all x in a field.
Proof: $0x + 0 = 0x = (0 + 0)x$, by Axiom (4)
$= 0x + 0x$ by Axiom (11)
$\therefore 0 = 0x$ by Theorem 2. Similarly $x0 = 0$.

One of the deep mysteries, when you first learnt algebra, was probably why $(-x)(-y) = xy$. I remember my teacher fobbing me off with the explanation that "two negatives make a positive – after all, if you're not unkind, you're kind."

This seemed to satisfy me then. I didn't reply, "well if two negatives make a positive why isn't −1 plus −1 equal to plus 2?"

But any explanation along the lines of double negatives is totally invalid. I suppose I was grateful that my teacher didn't try to explain it properly. Here's a proper proof.

Theorem 4: $(-x)(-y) = xy$ for all x, y in a field.

Proof: $[xy + x(-y)] + (-x)(-y) = xy + [x(-y) + (-x)(-y)]$ by Axiom (2)

$\therefore\ [x(y + (-y)] + (-x)(-y) = xy + [x + (-x)](-y)$ by Axiom (11)

$\therefore\ x0 + (-x)(-y) = xy + 0(-y)$ by Axiom (5)

$\therefore\ 0 + (-x)(-y) = xy + 0$ by Theorem (3)

$\therefore\ (-x)(-y) = xy$ by Axiom (4)

§1.3. Vector Spaces

A **vector space over a field** F is a set V, together with a specific element $0 \in V$, and a function $v \to -v$ from V to V, such that the following 10 axioms hold.

VECTOR SPACE AXIOMS

(1) (Closure under addition): $u + v \in F$.

(2) (Associativity under addition):

$$x + (v + w) = (u + v) + w.$$

(3) (Commutativity under addition): $u + v = v + \underline{u}$.

(4) (Identity under addition): $v + 0 = v = 0 + v$

(5) (Inverses under addition): $v + (-v) = 0 = (-v) + v$.

(6) (Closure under scalar multiplication): $\lambda v \in V$.

(7) (Associativity under scalar multiplication):

$$\lambda(\mu v) = (\lambda\mu)v.$$

(8) (Identity under scalar multiplication): $1v = v$.

(9) (Inverses under multiplication): $xx^{-1} = 1 = x^{-1}x$.

(10) (Distributivity):

$$\lambda(u + v) = \lambda u + \lambda v \text{ and } (\lambda + \mu)v = \lambda v + \mu v.$$

NOTES:

(1) You may wonder why we haven't printed the vectors u, v, w in bold type. That is done when you first learn about vectors so that you can see clearly which are vectors and which are scalars. So we wrote $\lambda\mathbf{v}$ to emphasise that λ is a scalar and **v** is a vector. However there's nothing in the set of axioms for a vector space that says that vectors and scalars are different things. There are examples where the field F is a subset of the vector space V and so the elements of F are both vectors and scalars.

(2) There is some similarity between the vector space axioms and the field axioms. In fact the first 5 axioms are

the same in both cases. Both fields and vector spaces are abelian groups under addition.

When it comes to multiplication we only define the product of a scalar (element of F) with a vector (element of V). Moreover we always write the scalar before the vector, so there is no commutative law for scalar multiplication.

Example 4: $\mathbb{C}$ is a vector space over $\mathbb{R}$. Here the real numbers are scalars, and the complex numbers are the vectors. But the real number are also scalars.

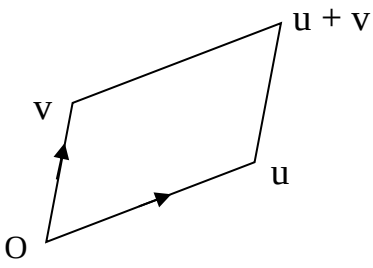
This situation occurs whenever you have a subfield of a larger field. If you go through the 10 axioms, in the situation where F is a subfield of V, you will find that the 10 vector space axioms are direct consequences of the field axioms.

In the area of mathematics that studies fields (Galois Theory) the theory of vector spaces plays an important role.

Example 5: For any field F, F^n is the set of all n-tuples $(x_1, x_2, \ldots, x_n)$ where each $x_i \in F$ and where addition and scalar multiplication are defined in the usual way. It's easily seen to be a vector space over F.

Example 6: Consider 3-dimensional Euclidean space. This is a vector space over $\mathbb{R}$ as follows. The vectors are the directed line segments from the origin to a point. The

scalars are the real numbers. Addition is defined by completing a parallelogram.



Multiplying a non-zero vector v by a positive scalar λ produces a vector with the same direction as v but with λ times the length. Multiplying by a negative scalar magnifies the length and reverses the direction. And, of course $0v = 0 = \lambda 0$ for all vectors v and scalars λ.

Clearly this is essentially the same as $\mathbb{R}^3$. We will make the concept of 'essentially the same' precise in a later chapter.

Example 7: For any field F, F^∞ is the set of all infinite sequences $(x_1, x_2, \ldots)$ with each $x_i \in F$ and with addition and multiplication defined in the obvious way.

Example 8: For any field F, F[x] is the set of all polynomials $a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ where each $x_i \in F$ and n is any non-negative integer. Addition and scalar multiplication are defined in the obvious way. F[x] is clearly a vector space over F. Note that we could write the polynomial as an infinite sequence $(a_0, a_1, \ldots)$, but F[x]

differs from F^{∞} in that here all the components from some point on are zero.

Example 9: For any field F, $M_n(F)$ is the set of all $n \times n$ matrices with the usual addition and scalar multiplication. This is clearly a vector space over F.

Example 10: Let $V = \wp\{1, 2, 3\}$, that is the set of all subsets of {1, 2, 3, 4, 5, 6, 7, 8, 9, 10} and let the field of scalars be $F = \mathbb{Z}_2 = \{0, 1\}$ with addition and multiplication mod 2.

Define $S + T = S \cup T - S \cap T$.

Then $\{1, 3, 7\} + \{3, 5, 7, 8\} = \{1, 3, 5, 7, 8\} - \{3, 7\}$
$= \{1, 5, 8\}$.

For all subsets S, define $0S = \varnothing = \{\ \}$, the empty set and $1S = S$.

This strange vector space has $2^{10} = 1024$ elements. The zero vector is $\varnothing$ and $S + S = \varnothing$ for all subsets S.

Up to this point the vector spaces have had recognisable components. In the next example this is not the case.

Example 11: Diff($\mathbb{R}$) is the set of all differentiable functions from the reals to the reals. The sum of two differentiable functions $f(x)$ and $g(x)$ is the differentiable function $f(x) + g(x)$ and the scalar multiple of the differentiable function $f(x)$ by the scalar λ is the

differentiable function $\lambda f(x)$. Thus the closure laws hold. The remaining axioms are just as obvious.
The function $f(x) = x^2$ belongs to Diff($\mathbb{R}$). But what are its components?

Spaces of functions are very important in the deeper study of analysis.

§1.4. Subspaces

A **subspace** of a vector space is a non-empty subset that is closed under addition and scalar multiplication. The fact that $-v = (-1)v$ guarantees that a subspace is also closed under inverses. Note too that if V is any vector space $\{0\}$ is a subspace of V, as is V itself.

If U is a subspace of V we write $\mathbf{U \leq V}$. Every vector space is a subspace of itself, but if we want to emphasise that U is not the same as V we can write $\mathbf{U < V}$ and say that U is a **proper subspace** of V.

Examples 12:
(1) $\{(x, y, z) \in \mathbb{R}^3 \mid z = x + y\}$ is a subspace of $\mathbb{R}^3$.
(2) $\{(x, y, z) \in \mathbb{R}^3 \mid 3x + 2y + 5z = 0 \text{ and } 7x - y + 2z = 0\}$ is a subspace of $\mathbb{R}^3$.
(3) A plane that passes through the origin is a subspace of 3-dimensional space.
(4) The set of differentiable functions from $\mathbb{R}$ to $\mathbb{R}$ is a subspace of the set of continuous functions from $\mathbb{R}$ to $\mathbb{R}$, which in turn is a subspace of the set of *all* functions from $\mathbb{R}$ to $\mathbb{R}$.

(5) The set of convergent sequences is a subspace of $\mathbb{R}^\infty$.
(6) The set of diagonal $n \times n$ matrices is a subspace of the space of all $n \times n$ matrices.
(7) $\mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}$.
(8) In Example 3, $\mathbb{Z}_2$ is a subspace of the field with 4 elements.
(9) For any vector space V $\{0\} \leq V$.

There are 10 axioms for a vector space, but axioms (3), (4), (7), (8), (9) and (10) will automatically be inherited by any subset. If the subset is non-empty we can dispense with (5) and (6) as well.

Theorem 5: If a non-empty subset is closed under addition and scalar multiplication then it is a subspace.
Proof: Closure under inverses follows from closure under scalar multiplication and the fact that $-v = (-1)v$. Closure under zero follows from the fact that $0v = 0$, **provided the subset is non-empty**. ☝☺

But note that the empty set is (vacuously) closed under addition, inverses and scalar multiplication, but it is not a subspace.

If U and V are subspaces of W there are two other subspaces that can be formed from them (though in special cases these may coincide with U or V). The **intersection** $U \cap V$ is a subspace, as is the **sum**, U + V, which is defined to be $\{u + v \mid u \in U, v \in V\}$.

Theorem 6: If U, V are subspaces of W (a vector space over F) then so are:

(1) $U \cap V$ and
(2) $U + V$.

Proof:
(1) Since $0 \in U \cap V$, it is non-empty.
Let $u, v \in U \cap V$ and $\lambda \in F$.
Then $u, v \in U$ and λv, for all scalars λ, belong to U.
Similarly they both belong to V and so belong to $U \cap V$.

(2) Since $0 = 0 + 0 \in U + W$, it is non-empty.
Let $w_1 = u_1 + v_1$ and $w_2 = u_2 + v_2$ belong to $U + V$, with $u_1, u_2 \in U$ and $v_1, v_2 \in V$.
Then $w_1 + w_2 = (u_1 + v_1) + (u_2 + v_2)$
$= (u_1 + u_2) + (v_1 + v_2) \in U + V$.
Let $\lambda \in F$ and $w = u + v$ where $u \in U$ and $v \in V$.
Then $\lambda w = \lambda(u + v) = \lambda u + \lambda v \in U + V$. ✋☺

Example 13: If U, V are distinct lines through the origin (in 3-dimensional space) $U \cap V$ is $\{0\}$ and $U + V$ is the plane that passes through both lines. If U, V are distinct planes through the origin then $U \cap V$ is the line where the planes intersect and $U + V$ is the 3-dimensional space.

§1.5. Bases

A **linear combination** of $v_1, v_2, \ldots, v_n \in V$ is any vector of the form $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n$.

The **span** of $v_1, v_2, \ldots, v_n$ is the subspace $\langle \mathbf{v_1, v_2, \ldots, v_n} \rangle$, which is the set of all linear combinations of them. The space spanned by the empty set is defined to be $\{0\}$. So the span of a finite set is the smallest subspace that contains it. If this is the space V we say that $v_1, v_2, \ldots, v_n$ **span** V.

Examples 14:
(1) $\langle v \rangle$ is the set of all scalar multiples of v;
(2) (1, 1, 0), (1, 0, 0) and (0, 1, 1) span $\mathbb{R}^3$ since
$(x, y, z) = (y - z)(1, 1, 0) + (x + z - y)(1, 0, 0) + z(0, 1, 1)$.
(3) (1, 1, 0), (1, 2, 1) and (0, 1, 1) span the plane:

$$\{(x, y, z) \mid x + z = y\},$$

not the whole of $\mathbb{R}^3$.

Clearly $\langle v_1, v_2, \ldots, v_n \rangle = \langle v_1 \rangle + \langle v_2 \rangle + \ldots + \langle v_n \rangle$.

Theorem 7: If $u \in \langle v_1, v_2, \ldots, v_n \rangle$ then $\langle u, v_1, v_2, \ldots, v_n \rangle = \langle v_1, v_2, \ldots, v_n \rangle$.
Proof: Suppose $u \in \langle v_1, v_2, \ldots v_n \rangle$ and let $u = x_1v_1 + x_2v_2 + \ldots + x_nv_n$.
Clearly $\langle v_1, v_2, \ldots, v_n \rangle \leq \langle u, v_1, v_2, \ldots, v_n \rangle$.
Since $\lambda u + \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = (\lambda x_1 + \lambda_1)v_1 + (\lambda x_2 + \lambda_2)v_2 + \ldots + (\lambda x_n + \lambda_n)v_n$ the inequality holds in reverse. ✋☺

The vectors $v_1, v_2, \ldots, v_n$ are defined to be **linearly independent** if

$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0$ implies that $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$.

If they're not linearly independent they are said to be **linearly dependent**.

Example 14: In 3-dimensional space two vectors are linearly dependent if they are in the same line. Three vectors are linearly dependent if they are in the same plane.

Example 15: Are (1, 1, 1, 4), (1, 4, 7, 13), (3, 1, 8, 6), (4, 2, 6, 10) linearly independent?

Solution: Suppose $\lambda_1(1, 1, 1, 4) + \lambda_2(1, 4, 7, 13) + \lambda_3(3, 1, 8, 6) + \lambda_4(4, 2, 6, 10) = (0, 0, 0, 0)$

We solve the resulting system of equations by reducing the coefficient matrix to echelon form.

$$\begin{pmatrix} 1 & 1 & 3 & 4 \\ 1 & 4 & 1 & 2 \\ 1 & 7 & 8 & 6 \\ 4 & 13 & 6 & 10 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 3 & 4 \\ 0 & 3 & -2 & -2 \\ 0 & 6 & 5 & 2 \\ 0 & 9 & -6 & -6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 3 & 4 \\ 0 & 3 & -2 & -2 \\ 0 & 0 & 9 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$ so there is a non-zero solution and hence the vectors are linearly dependent. Solving the system of equations we obtain the non-zero solution

$\lambda_4 = 9$, $\lambda_3 = -6$, $\lambda_2 = 2$, $\lambda_1 = -20$.

So $2(1, 4, 7, 13) + 9(4, 2, 6, 10) = 6(3, 1, 8, 6) + 20(1, 1, 1, 4)$.

A **basis** for a finite-dimensional vector space V is a linearly independent set that spans V.

Example 16:
The set {(1, 0, 0, … , 0, 0), (0, 1, 0, … , 0, 0), …
… , (0, 0, 0, …, 0, 1)
is a basis for F^n, where F is any field. This is called the **standard basis**. We often denote it by $\{e_1, e_2, \dots, e_n\}$.

A **minimal spanning set** for V is a set of vectors that spans V and has smallest size of any spanning set.

Theorem 8: If V has a spanning set of size *m* and a linearly independent set of size *n* then $m \geq n$.
Proof: Let A be a spanning set for V and let B be a linearly independent subset of V.
Suppose #A = *m* and #B = *n*. Let $b \in B$.
Since A spans V, b is a linear combination of the vectors in A.
Now some of the vectors in A might also be in B.
But since B is linearly independent the coefficient of some element $v \in A - B$ must be non-zero. Then v can be expressed as a linear combination of $A - \{v\} + \{b\}$ and so this set spans V. In other words, we can replace v by b and the resulting set still spans V. Continuing in this way we can transfer all the vectors from B into A, displacing an equal number of vectors. Hence $m \geq n$. ✋☺

Corollary: All bases for a finite-dimensional vector space have the same number of elements.
Proof: If two bases have *m* and *n* elements respectively then $m \leq n$ and $n \leq m$.

The unique number of vectors in a basis of a finite-dimensional vector space V is called the **dimension** of V and is denoted by **dim(V)**. A set of vectors in V whose size exceeds dim(V) is always linearly dependent. A set of vectors whose size is less than dim(V) cannot span V.

Examples 17:

(1) dim $F^n = n$ because the vectors (1, 0, 0, … , 0), (0, 1, 0, … , 0), … (0, 0, 0, … , 0, 1) form a basis (called the **standard basis**).

(2) The dimension of three dimensional Euclidean space is 3, of course!

(3) $\dim M_n(F) = n^2$. The standard basis consists of the matrices E_{ij} which have a 1 in the *i-j* position and 0's elsewhere.

(4) dim $\mathbb{C}$ (as a vector space over $\mathbb{R}$) is 2, with $\{1, i\}$ as an obvious basis.

(5) The space Diff($\mathbb{R}$) is infinite dimensional.

(6) The dimension of the zero subspace is 0.

§1.6. Sums and Direct Sums

Recall that the **sum** of two subspaces U, V is
$U + V = \{u + v \mid u \in U, v \in V\}$
and their intersection is
$U \cap V = \{v \mid v \in U \text{ and } v \in V\}$.

The sum U + V is called a **direct sum** whenever
$U \cap V = 0$.
(Here we are writing the zero subspace as 0, instead of {0}).

If W is the direct sum of U and V we write $W = U \oplus V$.

Example 18: In $\mathbb{R}^3$, if U, V are two distinct planes through the origin then $U + V = \mathbb{R}^3$. The sum is not direct, however, because $U \cap V$ will be a line through the origin. On the other hand if U is a plane through the origin and V is a line through the origin that doesn't lie in the plane, then $U + V = \mathbb{R}^3$ as before, but this time $U \cap V = 0$. We can therefore write $\mathbb{R}^3 = U \oplus V$.

Theorem 7: If $W = U \oplus V$ then every element of W can be expressed uniquely as u + v for $u \in U$ and $v \in V$.
Proof: The only part that isn't immediately obvious is the directness of the sum.
Suppose $u_1 + v_1 = u_2 + v_2$ with $u_1, u_2 \in U$ and $v_1, v_2 \in V$.
Then $u_1 - u_2 = v_2 - v_1 \in U \cap V = 0$. Hence $u_1 = u_2$ and $v_1 = v_2$. ☺

The converse also holds. If every element of W can be expressed uniquely as u + v for $u \in U$ and $v \in V$ then

$$W = U \oplus V.$$

The dimension of $U \cap V$ can be expressed in terms of the dimensions of U, V and U + V.

Theorem 8: If U, V are subspaces of W then
$\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V)$.

Proof: Let $m = \dim(U)$, $n = \dim(V)$ and $r = \dim(U \cap V)$.

Take a basis $w_1, w_2, \ldots, w_r$ for $U \cap V$.

Extend this to a basis $w_1, w_2, \ldots, w_r, u_1, u_2, \ldots, u_{m-r}$ for U.

Now extend the basis for $U \cap V$ to a basis:

$$w_1, w_2, \ldots, w_r, v_1, v_2, \ldots, v_{n-r} \text{ for V.}$$

We shall show that $v_1, \ldots, v_r, u_1, \ldots, u_{m-r}, v_1, \ldots, v_{n-r}$ is a basis for U + V.

They span V:

Let $u + v \in U + V$ where $u \in U$ and $v \in V$.

Then $u = \alpha_1 w_1 + \ldots + \alpha_r w_r + \beta_1 u_1 + \ldots + \beta_{m-r} u_{m-r}$ for some α_i's and β_i's $\in F$

and $v = \gamma_1 w_1 + \ldots + \gamma_r w_r + \delta_1 v_1 + \ldots + \delta_{n-r} v_{n-r}$ for some γ_i's and δ_i's $\in F$.

Hence $u + v = [\alpha_1 w_1 + \ldots + \alpha_r w_r + \beta_1 u_1 + \ldots + \beta_{m-r} u_{m-r}]$
$+ [\gamma_1 w_1 + \ldots + \gamma_r w_r + \delta_1 v_1 + \ldots + \delta_{n-r} v_{n-r}]$
$= (\alpha_1 + \gamma_1) w_1 + \ldots + (\alpha_r + \gamma_r) w_r + \beta_1 u_1 + \ldots$
$\ldots + \beta_{m-r} u_{m-r} + \delta_1 v_1 + \ldots + \delta_{n-r} v_{n-r}.$

They are linearly independent:

Suppose $\alpha_1 w_1 + \ldots + \alpha_r w_r + \beta_1 u_1 + \ldots + \beta_{m-r} u_{m-r}$

$$+ \delta_1 v_1 + \ldots + \delta_{n-r} v_{n-r} = 0 \quad \ldots\ldots\ldots (*)$$

Then $\delta_1 v_1 + \ldots + \delta_{n-r} v_{n-r} = -\alpha_1 w_1 - \ldots - \alpha_r w_r$

$$- \beta_1 u_1 - \ldots - \beta_{m-r} u_{m-r} \in U \cap V = 0.$$

Hence $\delta_1 v_1 + \ldots + \delta_{n-r} v_{n-r} = 0$ and

$\alpha_1 w_1 + \ldots + \alpha_r w_r + \beta_1 u_1 + \ldots + \beta_{m-r} u_{m-r} = 0.$

Since $v_1 \ldots, v_{n-r}$ are linearly independent (they are a basis for V) it follows that

$$\delta_1 = \ldots = \delta_{n-r} = = 0.$$

Since $w_1, \ldots, w_r, u_1 \ldots, u_{m-r}$ are linearly independent (they are a basis for U) it follows that

$$\alpha_1 = \ldots = \alpha_r = \beta_1, \ldots, \beta_{m-r} = 0.$$

Hence $\dim(U + V) = r + (m - r) + (n - r)$

$$= m + n - r. \text{ 👋☺}$$

We can generalise a sum to a sum of any finite number of terms.

The sum of $U_1, \ldots, U_k$ is $U_1 + \ldots + U_k$, the set of all vectors of the form $u_1 + \ldots + u_k$ where $u_r \in U_r$ for each r.

The above sum is a **direct sum** if $u_1 + \ldots + u_k = 0$, with each $u_r \in U_r$, implies that each $u_r = 0$.

EXERCISES FOR CHAPTER 1

Exercise 1: Prove that the set of all real symmetric matrices is a vector space over $\mathbb{R}$.

Exercise 2: Prove that for functions of a real variable $a(x)$, $b(x)$, $c(x)$ the solutions to the differential equation

$$a(x)\frac{d^2y}{dx^2} + b(x)\frac{dy}{dx} + c(x)y = 0$$

form a subspace of the space of all differentiable functions of x.

Exercise 3: Prove that the set of bounded sequences of real numbers is a vector space over $\mathbb{R}$.

Exercise 4: Show that the set S of all real sequences (a_n) where $\lim_{n\to\infty} a_n = \pm\infty$ is NOT a vector space.

Exercise 5: If $U = \{(x, y, z) \mid 3x - 2y + 5z = 0\}$
and $V = \{(x, y, z) \mid x - y + 3z = 0\}$
find $U \cap V$ and $U + V$.

Exercise 6: Show that (1, 2, 3), (4, 5, 6), (7, 8, 9) are linearly dependent.

Exercise 7: If u = (5, –2, 3) and v = (1, 4, –2) show that (13, –14, 13) $\in \langle u, v \rangle$.

Exercise 8: Show that {(5, 4, 2), (1, 2, 3), (0, 2, 1)} are linearly independent.

Exercise 9: Is {(1, 5, 7), (2, –1, 3), (6, 2, 8), (0, 5, 1)} linearly dependent or independent.

Exercise 10: Find the dimension of the space of all 3×3 symmetric matrices.

Exercise 11: Suppose U, V are subspaces of $\mathbb{R}^7$ with dimensions 4 and 5 respectively.
Suppose too that $U + V = \mathbb{R}^7$.
Find $\dim(U \cap V)$.

SOLUTIONS FOR CHAPTER 1

Exercise 1: Closure under +: Suppose A, B are real symmetric matrices. Then $A^T = A$ and $B^T = B$.
Hence $(A + B)^T = A^T + B^T = A + B$, so A + B is symmetric.
Closure under scalar ×: For any scalar k, $(kA)^T = kA^T = kA$ so the set is closed under scalar multiplication.
Hence the set is a subspace.

Exercise 2: Closure under +: Suppose $f(x)$, $g(x)$ are solutions to the differential equation.
Then $a(x)\frac{d^2f(x)}{dx^2} + b(x)\frac{df(x)}{dx} + c(x)f(x) = 0$ and

$$a(x)\frac{d^2g(x)}{dx^2} + b(x)\frac{dg(x)}{dx} + c(x)g(x) = 0.$$

Hence:

$$a(x)\frac{d^2(f(x)+g(x))}{dx^2} + b(x)\frac{d(f(x)+g(x))}{dx} + c(x)(f(x)+g(x))$$

$$= a(x)\left(\frac{d^2f(x)}{dx^2} + \frac{d^2g(x)}{dx^2}\right) + b(x)\left(\frac{df(x)}{dx} + \frac{dg(x)}{dx}\right) + c(x)f(x) + c(x)g(x)$$

$$= a(x)\frac{d^2f(x)}{dx^2} + b(x)\frac{df(x)}{dx} + c(x)f(x) + a(x)\frac{d^2g(x)}{dx^2} + b(x)\frac{dg(x)}{dx} + c(x)g(x)$$

$$= 0.$$

Closure under scalar multiplication: For any real number k

$$a(x)\frac{d^2(kf(x))}{dx^2} + b(x)\frac{d(kf(x))}{dx} + c(x)kf(x)$$

$$= ka(x)\left(\frac{d^2f(x)}{dx} + b(x)\frac{df(x)}{dx} + c(x)f(x)\right)$$

$$= 0$$

Exercise 3:

Closure under +:

Suppose (a_n) and (b_n) are bounded sequences.

Then, there exist K, L such that for all n

$$|a_n| \le K \text{ and } |b_n| \le L.$$

Now $|a_n + b_n| \le |a_n| + |b_n| \le K + L$ for all n.

Hence the sequence $(a_n) + (b_n) = (a_n + b_n)$ is bounded.

Closure under scalar multiplication: For any real number k, $|ka_n| = |k|.|a_n| \leq |k|.K$.
Therefore $k(a_n) = (ka_n)$ is bounded.
Hence the set is a subspace.

Exercise 4: It isn't closed under addition.
If $a_n = n$ and $b_n = -n$ then both (a_n) and (b_n) belong to S.
But $(a_n) + (b_n)$ does not.

Exercise 5:
$U \cap V = \{(x, y, z) \mid 3x - 2y + 5z = 0 \text{ and } x - y + 3z = 0\}$.
We solve the homogeneous system $\begin{pmatrix} 1 & -1 & 3 \\ 3 & -2 & 5 \end{pmatrix}$.

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

Let $z = k$. Then $y = 4k$ and $x = k$.
So $U \cap V = \{k(1, 4, 1) \mid k \in \mathbb{R}\}$.
Now $\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V)$
$= 2 + 2 - 1 = 3$ so $U + V = \mathbb{R}^3$.

Exercise 6: $(1, 2, 3) + (7, 8, 9) = 2(4, 5, 6)$.

Exercise 7:
Suppose $(13, -14, 13) = k(5, -2, 3) + h(1, 4, -2)$.
We attempt to solve the system $\begin{cases} 5k + h = 13 \\ -2k + 4h = -14 \\ 3k - 2h = 13 \end{cases}$

$$\left(\begin{array}{cc|c} 5 & 1 & 13 \\ -2 & 4 & -14 \\ 3 & -2 & 13 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 9 & -15 \\ -2 & 4 & -14 \\ 3 & -2 & 13 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 9 & -15 \\ 0 & 22 & -44 \\ 0 & -29 & 58 \end{array}\right)$$

$$\to \left(\begin{array}{cc|c} 1 & 9 & -15 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{array}\right).$$

So $h = -2$, $k = -15 - 9(-2) = 3$.
So $(13, -14, 13) = 3(5, -2, 3) - 2(1, 4, -2) \in \langle u, v \rangle$.

Exercise 8:
Suppose $a(5, 4, 2) + b(1, 2, 3) + c(0, 2, 1) = (0, 0, 0)$.

We solve the homogeneous system $\begin{cases} 5a + b = 0 \\ 4a + 2b + 2c = 0 \\ 2a + 3b + c = 0 \end{cases}$.

$$\begin{pmatrix} 5 & 1 & 0 \\ 4 & 2 & 2 \\ 2 & 3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & -2 \\ 4 & 2 & 2 \\ 2 & 3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & -2 \\ 0 & 6 & 10 \\ 0 & 5 & 5 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 6 & 10 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{pmatrix}.$$

$\therefore$ $a = b = c = 0$ and so the vectors are linearly independent.

Alternatively we can evaluate the determinant

$$\begin{vmatrix} 5 & 1 & 0 \\ 4 & 2 & 2 \\ 2 & 3 & 1 \end{vmatrix} = 5(2-6) - (4-4) = -20.$$

Since this is non-zero the vectors are linearly independent.

WARNING: This second method only works for n vectors in an n-dimensional vector space.

Exercise 9: Linearly dependent. Whenever you have more vectors than the dimension of the vector space from which they come, they must be linearly dependent.

Exercise 10:

A 3×3 symmetric matrix has the form $\begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix}$.

Clearly

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$ is a basis so the dimension is 6.

Exercise 11:

$\dim(U \cap V) = \dim(U) + \dim(V) - \dim(U + V)$

$= 4 + 5 - 7 = 2.$