Further Mathematical Methods (Linear Algebra) 2002

Lecture 1: Vector Spaces

In this hand-out we are going to revise some of the ideas which we encountered in MA100 about vector spaces. The Learning Objectives associated with this hand-out are given at the end.

1.1 Logic and Sets

Before we talk about vector spaces, we shall take a moment to examine the meaning of some pieces of notation. Firstly, most of the symbols that we meet will be familiar (or defined *in situ*), however three of them may be new to you. These are: '$\forall$', '$\exists$' and 'iff' which mean 'for all', 'there exists' and 'if and only if' respectively. I am sure that context will make the meaning of the first two clear, however the third one is a bit different. The words 'if and only if' (or, the symbol 'iff') represent a logical connective (called the *biconditional*) and you may have seen it before written as '$\Leftrightarrow$'.

It is useful to take a moment to see what this means, although what we say here will be stunningly brief and probably more than you need to know about logic. You should be aware that some mathematical theorems take the form 'if p, then q' (this is sometimes written as 'p implies q' or '$p \Rightarrow q$') where p and q are two propositions. If a theorem does take this form, its *converse* is written as 'if q, then p'. However, not every theorem of this form has a converse that is true — some do, some don't. But, *if* both the theorem *and* its converse are true, *then* we can say that 'p if and only if q' or 'p iff q' (this is sometimes written as 'p implies and is implied by q' or, as we noted before, '$p \Leftrightarrow q$'). Notice that when you come to prove theorems that have the form 'p iff q', it is necessary to prove both 'if p, then q' *and* 'if q, then p'. I shall refer to these two parts of the proof as 'RTL' (meaning 'Right-To-Left') and 'LTR' (meaning 'Left-To-Right') respectively.

Incidentally, another useful logical term that we will use is *contrapositive*. We say that the contrapositive of the theorem 'if p then q' is the expression 'if not-q then not-p' (here 'not' represents the logical operator called *negation*). It should be noted that (*unlike* the converse), the contrapositive is *logically equivalent* to the theorem we started with. You probably have some awareness of these ideas, but I just wanted to bring them out into the open without going into any of the details.

A *set* is a collection of objects, called *elements* or *members*, that can be viewed as an object itself. For example, the set $\{2n|n \in \mathbb{N}\}$ is the set of all even numbers. A set can be specified by either listing all of its elements between 'curly' brackets, say $\{2, 4, 6, \ldots\}$, or by giving a membership criterion, say 'the set of all numbers of the form $2n$ with $n \in \mathbb{N}$'. We denote the fact that an object is an element of a set by using the symbol '$\in$' and the fact that an object is not an element of a set by using the symbol '$\notin$'. For example, $2 \in \{2, 4, 6, \ldots\}$, but $3 \notin \{2, 4, 6, \ldots\}$. Lastly we say that two sets are equal if and only if they have the same elements, i.e.

$$A = B \text{ if and only if } (x \in A \Longleftrightarrow x \in B),$$

for all elements x and any two sets A and B.

Further, we can say that a set A is a *subset* of a set B, denoted by $A \subseteq B$, if all of the elements in A are also in B, i.e.

$$\text{If } (x \in A \Longrightarrow x \in B), \text{ then } A \subseteq B.$$

for all elements x. For example, we know that the set of all even numbers is a subset of the set of all natural numbers, i.e. $\{2, 4, 6, \ldots\} \subseteq \{1, 2, 3, 4, 5, 6, \ldots\}$. An immediate consequence of this is that two sets, A and B, are equal if A is a subset of B and *vice versa*, i.e.

$$A = B \text{ if and only if } (A \subseteq B \text{ and } B \subseteq A).$$

(Can you see why?) This result gives us a useful way of showing that two sets are equal, namely that two sets A and B are equal if we can show that A is a subset of B and B is a subset of A.[1] For example, we can show that the two sets $\{2, 4, 6, \ldots\}$ and $\{2n|n \in \mathbb{N}\}$ are equal using this method, i.e.

[1]This method for establishing the equality of two sets is sometimes called a 'double inclusion proof' since '$A \subseteq B$' is sometimes read as 'A is *included* in B'.

- For any element $m \in \{2, 4, 6, \ldots\}$, we can find an $n \in \mathbb{N}$ such that $m = 2n$ and so $m \in \{2n|n \in \mathbb{N}\}$. Thus, $\{2, 4, 6, \ldots\} \subseteq \{2n|n \in \mathbb{N}\}$.
- For any element $m \in \{2n|n \in \mathbb{N}\}$ we have $m = 2n$ and so $m \in \{2, 4, 6, \ldots\}$. Thus, $\{2n|n \in \mathbb{N}\} \subseteq \{2, 4, 6, \ldots\}$.

and so the two sets are equal, as one should expect since both of them are representations of the set of all even numbers.

1.2 What is a Vector Space?

A vector space is a mathematical structure which consists of a non-empty set of *objects* which are [unsurprisingly] called *vectors* and two mathematical operations called *vector addition* (denoted by '+') and *scalar multiplication* (denoted by '·'). The former rule associates an object $\mathbf{u} + \mathbf{v}$, called the *vector sum of* $\mathbf{u}$ *and* $\mathbf{v}$, with each pair of objects $\mathbf{u}$ and $\mathbf{v}$ in V; whilst the latter rule associates an object $\alpha \cdot \mathbf{u}$, called the *scalar multiple of* $\mathbf{u}$ *by* α, with each object $\mathbf{u}$ in V and any scalar α. The definition below tells us what we should expect:

Definition 1.1 *Let V be a non-empty set of objects, on which two operations,* vector addition *(i.e. '+') and* scalar multiplication *(i.e. '·'), are defined such that:*

(A)	$\forall \mathbf{u}, \mathbf{v} \in V,\ \mathbf{u} + \mathbf{v} \in V$	*Vector Addition*
(M)	$\forall \mathbf{u} \in V$ *and* $\forall \alpha \in K,\ \alpha \cdot \mathbf{u} \in V$	*Scalar Multiplication*

where K, the set of scalars, is a given field.[2] *If the following axioms are satisfied, then we call V a* vector space*, and we call the objects in V* vectors*:*

(AC)	$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$	*Commutativity of (A)*
(AA)	$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$	*Associativity of (A)*
(A0)	$\exists \mathbf{0} \in V$ *such that* $\forall \mathbf{u} \in V,\ \mathbf{0} + \mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{u}$	*Identity under (A)*
(AI)	$\exists (-\mathbf{u}) \in V$ *such that* $(-\mathbf{u}) + \mathbf{u} = \mathbf{u} + (-\mathbf{u}) = \mathbf{0}$	*Inverse under (A)*
(MD1)	$\alpha \cdot (\mathbf{u} + \mathbf{v}) = \alpha \cdot \mathbf{u} + \alpha \cdot \mathbf{v}$	*Distributivity of (M)*
(MD2)	$(\alpha + \beta) \cdot \mathbf{u} = \alpha \cdot \mathbf{u} + \beta \cdot \mathbf{u}$	*Distributivity of (M)*
(MA)	$(\alpha\beta) \cdot \mathbf{u} = \alpha \cdot (\beta \cdot \mathbf{u})$	*Associativity of (M)*
(M1)	$1 \cdot \mathbf{u} = \mathbf{u}$	*Identity under (M)*

for all objects $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, and for all scalars $\alpha, \beta \in K$.[3]

It is, perhaps, useful to have an understanding of what all of these axioms do, and we shall briefly examine the function of each one below.

The first two axioms, namely (A) and (M), demand that the vector space is *closed* under the two operations which are defined on it. This basically means that performing these operations on any objects that are in the vector space will lead to another object that is also in the vector space, i.e. you can't get an object that is not in the vector space by using them. The 'A' and 'M' axioms then tell us what properties the operations — vector addition and scalar multiplication respectively — must have. The commutativity, associativity and distributivity axioms — (AC), (AA), (MD1), (MD2) and (MA) respectively — should be familiar as they just tell us about the 'order' in which we can legally perform the operations. The remaining three axioms may be slightly less familiar, and so we shall consider each in turn. (A0) dictates that the vector space must contain an element, called $\mathbf{0}$, which serves as the identity under vector addition, that is, when it is added to another vector, that vector remains unchanged. (AI) specifies that every element in the vector space must have

[2] I do not really want to go into the question of what can count as a set of scalars, except to note that it is a *field*, which is another mathematical structure (and one that I really do not want to discuss). But, note that two 'common' fields are the set of real numbers, $\mathbb{R}$ and the set of complex numbers, $\mathbb{C}$. Thus, when we use the word 'scalar', it will be assumed that the set of scalars is either $\mathbb{R}$ or $\mathbb{C}$ as we know how real and complex numbers behave.

[3] We could have used an alternative form of the axioms (A0) and (AI), namely demanding only that $\mathbf{u} + \mathbf{0} = \mathbf{u}$ and $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ hold respectively. That these forms are equivalent follows at once from the axiom (AC).

an additive inverse, which is also in the vector space, and when a vector and its inverse are added together, they 'cancel,' leaving us with the additive identity.[4] Lastly, we have (M1), which tells us that the set of scalars being used must contain a multiplicative identity, i.e. an element which, when multiplied by a vector, leaves that vector unchanged.[5]

However, this is not a course in pure maths and so, there is no need for us to dwell on why we have selected these axioms. But, in passing, we note that vector spaces also have the following properties:

Theorem 1.2 *If V is a vector space, then*

1. *The additive identity, namely $\mathbf{0}$, is unique; and further, $0 \cdot \mathbf{u} = \mathbf{0}$ and $\alpha \cdot \mathbf{0} = \mathbf{0}$.*
2. *The additive inverse of a vector $\mathbf{u} \in V$, namely $-\mathbf{u}$, is such that $(-1) \cdot \mathbf{u} = -\mathbf{u}$.*
3. *If $\alpha \cdot \mathbf{u} = \mathbf{0}$, then $\alpha = 0$ or $\mathbf{u} = \mathbf{0}$.*

for all $\mathbf{u} \in V$ and all scalars, α.

Proof: Assume that V is a vector space, and that $\mathbf{u} \in V$ and α are an arbitrary vector and scalar respectively.

(1): We want to show that the additive identity, $\mathbf{0} \in V$ where

$$\mathbf{u} + \mathbf{0} = \mathbf{u} \tag{1.1}$$

is unique. To do this, we assume the existence of *any* other additive identity $\mathbf{0}' \in V$, which must be such that

$$\mathbf{0}' + \mathbf{u} = \mathbf{u} \tag{1.2}$$

Now, putting $\mathbf{u} = \mathbf{0}'$ in Equation 1.1 gives

$$\mathbf{0}' + \mathbf{0} = \mathbf{0}',$$

and putting $\mathbf{u} = \mathbf{0}$ in Equation 1.2 gives

$$\mathbf{0}' + \mathbf{0} = \mathbf{0},$$

and equating these we get

$$\mathbf{0} = \mathbf{0}'.$$

Thus, we have shown that the additive identity is unique (as required).

Further, we can show that $0 \cdot \mathbf{u} = \mathbf{0}$ because, using (MD2) we have

$$(0 + 0) \cdot \mathbf{u} = 0 \cdot \mathbf{u} + 0 \cdot \mathbf{u}$$

and as the number zero has the property that $0 + 0 = 0$, this means that

$$0 \cdot \mathbf{u} = 0 \cdot \mathbf{u} + 0 \cdot \mathbf{u}$$

Now, by (AI), there exists an additive inverse, $(-0 \cdot \mathbf{u}) \in V$ such that

$$(-0 \cdot \mathbf{u}) + 0 \cdot \mathbf{u} = 0 \cdot \mathbf{u} + (-0 \cdot \mathbf{u}) = \mathbf{0}$$

and adding this to both sides of the previous expression, we get

$$0 \cdot \mathbf{u} + (-0 \cdot \mathbf{u}) = [0 \cdot \mathbf{u} + 0 \cdot \mathbf{u}] + (-0 \cdot \mathbf{u})$$

[4] We know that this exists due to (A0).

[5] For example, if our scalars were real numbers, this role would be played by the number 1. This is the only multiplicative inverse that we will need and so I have chosen to denote this special object by '1'.

Then, applying (AA), i.e.

$$0 \cdot \mathbf{u} + (-0 \cdot \mathbf{u}) = 0 \cdot \mathbf{u} + [0 \cdot \mathbf{u} + (-0 \cdot \mathbf{u})]$$

and using the fact that by (AI), $0 \cdot \mathbf{u} + (-0 \cdot \mathbf{u}) = \mathbf{0}$, we find that

$$\mathbf{0} = 0 \cdot \mathbf{u} + \mathbf{0}$$

Finally, applying (A0), with $0 \cdot \mathbf{u}$ replacing $\mathbf{u}$, i.e. $0 \cdot \mathbf{u} + \mathbf{0} = 0 \cdot \mathbf{u}$, we have

$$\mathbf{0} = 0 \cdot \mathbf{u}$$

(as required).

And, further still, $\alpha \cdot \mathbf{0} = \mathbf{0}$ because, using (A0) with $\mathbf{u} = \mathbf{0}$ we get

$$\mathbf{0} + \mathbf{0} = \mathbf{0}$$

and then, by (A) we can multiply both sides by a scalar, α say, i.e.

$$\alpha \cdot (\mathbf{0} + \mathbf{0}) = \alpha \cdot \mathbf{0}$$

which by (MD2) becomes

$$\alpha \cdot \mathbf{0} + \alpha \cdot \mathbf{0} = \alpha \cdot \mathbf{0}.$$

Now, by (AI), there exists an additive inverse, $(-\alpha \cdot \mathbf{0}) \in V$, and adding this to both sides, we get

$$(-\alpha \cdot \mathbf{0}) + (\alpha \cdot \mathbf{0} + \alpha \cdot \mathbf{0}) = (-\alpha \cdot \mathbf{0}) + \alpha \cdot \mathbf{0}.$$

which, after applying (AA) on the left-hand side and using the fact that by (AI), $(-\alpha \cdot \mathbf{0}) + \alpha \cdot \mathbf{0} = \mathbf{0}$, becomes

$$\mathbf{0} + \alpha \cdot \mathbf{0} = \mathbf{0}.$$

Then, using (A0) with $\mathbf{u} = \alpha \cdot \mathbf{0}$, i.e. $\mathbf{0} + \alpha \cdot \mathbf{0} = \alpha \cdot \mathbf{0}$, this gives

$$\alpha \cdot \mathbf{0} = \mathbf{0}$$

(as required).

(2) and (3): See the problems. ♠

This proof has given us an idea of how to prove theorems about vector spaces. It is important to notice that although what we have justified looks obvious, the proof of (1) can only proceed because every step was justified by an appeal to the axioms. Once theorems [or any part of them] have been proved, we can then use these theorems and the axioms in further proofs. This should be borne in mind when giving proofs in this course.

1.3 Some Common Vector Spaces.

We shall now introduce the three vector spaces that will be used in this course. Among other things, this will help us to clarify what the axioms in Definition 1.1 commit us to.[6] At this point, we shall also drop the convention that a '·' denotes scalar multiplication (although, occasionally, we shall use it to make things clearer). Thus, from now an expression like '$\alpha\mathbf{u}$' will be synonymous with '$\alpha \cdot \mathbf{u}$'.

[6]We should really *prove* that the following sets are vector spaces instead of just 'noting' that the axioms will be satisfied.

1.3.1 Euclidean n-space, $\mathbb{R}^n$.

This is a vector space which we are all familiar with, and in the general, n-dimensional case, consists of the set of all n-tuples of real numbers. These n-tuples can be interpreted in two ways: as the points, $(x_1, x_2, \ldots, x_n)$ in an n-dimensional space, or as the [position] *vectors*, $\mathbf{x} = [x_1, x_2, \ldots, x_n]^t$, which represent these points.[7] Of course, the latter view is the more useful here, and using the familiar rules of vector manipulation we can define vector addition:

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

and scalar multiplication:

$$\alpha\mathbf{x} = \alpha \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix}$$

where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ as in this *real* space, we take $x_1, \ldots, x_n, y_1, \ldots, y_n$ and the scalars to be elements of the set $\mathbb{R}$.

To be sure that this is indeed a vector space, we note that we have a vector which plays the role of the additive identity — namely, the *null vector* $\mathbf{0} = [0, 0, \ldots, 0]^t$, which is the position vector of the origin. We also have an additive inverse for each element $\mathbf{x} \in \mathbb{R}^n$, which is given by $-\mathbf{x} = [-x_1, -x_2, \ldots, -x_n]^t$; as well as a multiplicative identity which is, as mentioned before, the number 1. Lastly, the rules specified above guarantee that the rest of the axioms hold and so $\mathbb{R}^n$ is indeed a vector space.

For example: The vectors $[1, 0, 2]^t$ and $[1, 2, 0]^t$ are both in the vector space $\mathbb{R}^3$. Using the rules defined above, we can manipulate these vectors using the operations of vector addition and scalar multiplication. For instance,

$$\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Also, notice that the inverse of the vector $[1, 0, 2]^t$ is given by $[-1, 0, -2]^t$ and clearly satisfies the equality

$$\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

as required by (AI). Here, as mentioned above, the role of the additive identity is played by the null vector $\mathbf{0} = [0, 0, 0]^t$ and this has the property that, for instance,

$$\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

as required by (A0). ♣

[7] We shall use the convention that the position vectors in $\mathbb{R}^n$ are *column* vectors. This means that when we write them as *row* vectors we should stress that we are really looking at the *transpose* of the position vector.

1.3.2 Complex n-space, $\mathbb{C}^n$.

This is *almost* identical to $\mathbb{R}^n$ although now, in Definition 1.1 we take scalars which are complex numbers, i.e. $\alpha \in \mathbb{C}$ (notice that the results of Theorem 1.2 also continue to hold). Consequently, a general vector in $\mathbb{C}^n$ will be an entity that is complex, and so is not defined in any real space like $\mathbb{R}^n$. This might make this extension seem uninteresting, but as we shall see later, differences do arise.

For example: The vectors $[1, 0, 1+i]^t$ and $[i, 2, 0]^t$ are both in the vector space $\mathbb{C}^3$ (where i is the square root of -1, remember?). Using the rules defined above, we can manipulate these vectors using the operations of vector addition and scalar multiplication. For instance,

$$\begin{bmatrix} 1 \\ 0 \\ 1+i \end{bmatrix} + \begin{bmatrix} i \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1+i \\ 2 \\ 1+i \end{bmatrix} = (1+i) \begin{bmatrix} 1 \\ 1-i \\ 1 \end{bmatrix}.$$

where we have used the fact that

$$\frac{2}{1+i} = \frac{2(1-i)}{(1+i)(1-i)} = \frac{2(1-i)}{1+1} = 1 - i$$

Also, notice that the inverse of the vector $[i, 2, 0]^t$ is given by $[-i, -2, 0]^t$ and clearly satisfies the equality

$$\begin{bmatrix} i \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} -i \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} -i \\ -2 \\ 0 \end{bmatrix} + \begin{bmatrix} i \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

as required by (AI). Here, as mentioned above, the role of the additive identity is [again] played by the null vector $\mathbf{0} = [0, 0, 0]^t$ and this has the property that, for instance,

$$\begin{bmatrix} i \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} i \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} i \\ 2 \\ 0 \end{bmatrix}$$

as required by (A0). ♣

1.3.3 Real-function space, $\mathbb{F}^{\mathbb{R}}$.

This space consists of the set of all real-valued functions which are defined on the entire real line, i.e. $\mathbb{R}$.[8] Consequently, in this vector space, the functions *themselves* are the vectors and so we shall write $\mathbf{f}(x)$ instead of $f(x)$, say, where $f : \mathbb{R} \to \mathbb{R}$. Then, using the familiar rules for manipulating functions we can define [*point-wise*] vector addition:

$$[\mathbf{f} + \mathbf{g}](x) = f(x) + g(x), \text{ that is, } \mathbf{f} + \mathbf{g} : x \to f(x) + g(x), \ \forall x \in \mathbb{R},$$

and [*point-wise*] scalar multiplication:

$$[\alpha\mathbf{f}](x) = \alpha f(x), \text{ that is, } \alpha\mathbf{f} : x \to \alpha f(x), \ \forall x \in \mathbb{R},$$

where $\mathbf{f}, \mathbf{g}$ are the vectors in $\mathbb{F}^{\mathbb{R}}$ representing the functions $f, g : \mathbb{R} \to \mathbb{R}$ and $\alpha \in \mathbb{R}$.[9] It may seem a bit weird to think of functions as vectors in a vector space, especially as the functions themselves are dependent on the variable x, so perhaps a picture, like Figure 1.1, will help.

[8] There are many ways of restricting this space, but many of these will not concern us. One way is to restrict the functions to those that are differentiable (for all $x \in \mathbb{R}$) a certain number of times. For example, we could specify that we are working with the set of all real-valued functions that can be differentiated m times. Or, we could restrict the functions to those that are defined over a certain interval. For example, we could use the set of real-valued functions that are defined over the closed interval $[a, b]$, and this is denoted by $\mathbb{F}^{[a,b]}$. Alternatively, we could extend the set to include complex-valued functions and get a complex-function space. This notion of a 'restriction' of a vector space which gives us another vector space is essentially the idea of a vector subspace, see Section 1.4.

[9] Notice that we have put the vector sum, $\mathbf{f} + \mathbf{g}$, and the scalar multiple, $\alpha\mathbf{f}$, in square brackets to indicate that it is this function as a whole that is being evaluated at a certain point x. This is the essence of *point-wise* operations.

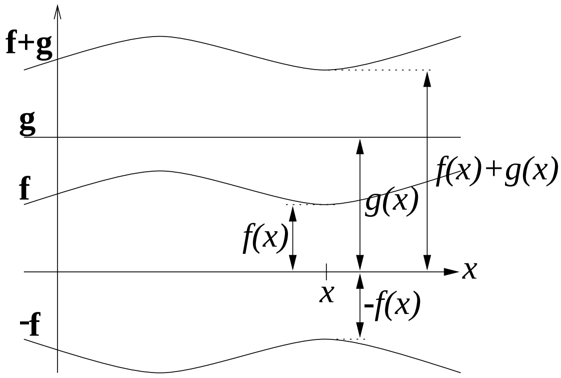

Figure 1.1: *Here,* $\mathbf{f}$ *and* $\mathbf{g}$ *are vectors in* $\mathbb{F}^{\mathbb{R}}$. *Notice that the vector sum* $\mathbf{f}+\mathbf{g}$ *and the scalar multiple* $-1 \cdot \mathbf{f}$ *(or,* $-\mathbf{f}$*, see Theorem 1.2) are obtained by first calculating* $f(x)$ *and* $g(x)$ *at each point* x*, and then these values are used to find* $f(x)+g(x)$ *and* $-f(x)$ *at each point. These are sometimes called point-wise* operations.

To be sure that this is indeed a vector space, we note that we have a vector which plays the role of the additive identity — namely, the zero function, $\mathbf{0} : x \to 0$ for all $x \in \mathbb{R}$ (i.e. the function which gives us zero for all values of the variable x) and we can denote this by $\mathbf{0}(x)$, or just $\mathbf{0}$. We also have an additive inverse for each element $\mathbf{f} \in \mathbb{F}^{\mathbb{R}}$, which is given by $-\mathbf{f} : x \to -f(x)$ for all $x \in \mathbb{R}$. Lastly, we have a multiplicative identity, which is, as before, the number 1; and the rules specified above guarantee that the rest of the axioms hold and so $\mathbb{F}^{\mathbb{R}}$ is indeed a vector space.

For example: The vectors $\mathbf{2}$ and $\mathbf{6x^2}$ are both in the vector space $\mathbb{F}^{\mathbb{R}}$ where they represent the functions

$$\mathbf{2} : x \to 2 \text{ and } \mathbf{6x^2} : x \to 6x^2$$

respectively for all $x \in \mathbb{R}$. Using the rules defined above, we can manipulate these vectors using the operations of vector addition and scalar multiplication. For instance, as $2 + 6x^2 = 2(1 + 3x^2)$ for all $x \in \mathbb{R}$,

$$\mathbf{2} + \mathbf{6x^2} = 2(\mathbf{1} + \mathbf{3x^2})$$

where $\mathbf{1} + \mathbf{3x^2} : x \to 1 + 3x^2$ for all $x \in \mathbb{R}$. Also, notice that the inverse of the vector $\mathbf{6x^2}$ is given by $-\mathbf{6x^2}$ and clearly, because $6x^2 + (-6x^2) = (-6x^2) + 6x^2 = 0$ for all $x \in \mathbb{R}$, this satisfies the equality

$$\mathbf{6x^2} + (-\mathbf{6x^2}) = (-\mathbf{6x^2}) + \mathbf{6x^2} = \mathbf{0}$$

as required by (AI). Here, as mentioned above, the role of the additive identity is played by the null vector $\mathbf{0} : x \to 0$ for all $x \in \mathbb{R}$ and, for instance, because $6x^2 + 0 = 0 + 6x^2 = 6x^2$ for all $x \in \mathbb{R}$, this has the property that

$$\mathbf{6x^2} + \mathbf{0} = \mathbf{0} + \mathbf{6x^2} = \mathbf{6x^2}$$

as required by (A0). ♣

1.4 Vector Subspaces

Imagine that we have a vector space, which as we saw above will consist of a set of objects, V, and two operations that are defined so as to satisfy the ten vector space axioms. What happens if we take a subset of this set of objects? Or, more specifically, can we find any subset, say, W of V which also has the properties of a vector space? The answer is, of course, yes! We can define a vector *subspace* as a set of objects, $W \subseteq V$, that also satisfy the vector space axioms. However, in the case of vector subspaces, there is no need to test all of the axioms; we just have to check that the first two — namely closure under the two operations — hold. This is because each vector subspace is 'embedded' in a vector space and so it 'inherits' most of the required structure. We can therefore say that:

Definition 1.3 *A non-empty subset of a vector space, V, is a* subspace *of V if it is a vector space under the operations of vector addition and scalar multiplication defined in V.*

and further, we notice that:

Theorem 1.4 *If V is a vector space and $W \subseteq V$ is a non-empty subset, then W is a subspace iff the following two conditions hold:*

$$\begin{array}{lll} \text{(CA)} & \forall \mathbf{u}, \mathbf{v} \in W,\ \mathbf{u} + \mathbf{v} \in W & \textit{Closure under (A).} \\ \text{(CM)} & \forall \mathbf{u} \in W,\ \alpha \mathbf{u} \in W & \textit{Closure under (M).} \end{array}$$

for all scalars, α.

The proof of this theorem is left as an exercise for anyone who is interested.

This theorem gives us a simple way of testing if a subset of the objects in a vector space is a subspace, but sometimes a quicker method may be available (if, for instance, the subset is obviously not a vector space because it violates one of the other vector space axioms). We now turn our attention to this aspect of vector spaces.

1.4.1 When is a subset of a vector space a subspace?

You may have noticed that in Definition 1.3 and Theorem 1.4 we stipulated that the subset of V under consideration must be non-empty. This must obviously be the case as if we took the *empty* subset, $\varnothing$, of V then it would not contain any elements, and so *a fortiori*, it would not contain an additive identity, i.e. $\mathbf{0} \notin \varnothing$. But, of course, this violates axiom (A0) of Definition 1.1 and so $\varnothing$ cannot be a vector space, or indeed, a subspace (by Definition 1.3).

However, if we took the subset of V that contained just the additive identity, i.e. the set $\{\mathbf{0}\}$, then this is a subspace [albeit a trivial one] by Theorem 1.4 because we know that

$$\begin{array}{lll} \text{(CA)} & \mathbf{0} + \mathbf{0} = \mathbf{0} & \text{:by axiom (A0) of Definition 1.1} \\ \text{(CM)} & \alpha \mathbf{0} = \mathbf{0} & \text{:by part (1) of Theorem 1.2} \end{array}$$

for all scalars, α.

Unfortunately, we often need to bring more sophisticated techniques to bear on the problem of deciding whether a subset is a subspace, let us consider these in the context of some examples.

1.4.2 Some Relevant Methods of Proof

As well as doing some examples, we shall also revise some useful methods of proof that may have been forgotten [or never really understood].

To show that a subset is a subspace:

In this case, we can invoke Theorem 1.4, where it is necessary to show that the subset is closed under the operations of vector addition and scalar multiplication. The key here is to show that the closure conditions are satisfied by *all* of the vectors in the subset, and this means performing the proof with *general* vectors and scalars in the subset.[10]

Example: Show that the subset $S_{a,b,c} = \{\mathbf{x} \,|\, ax_1 + bx_2 + cx_3 = 0\}$ (where a, b and c are [given] real numbers) of $\mathbb{R}^3$ is a vector space, i.e. show that $S_{a,b,c}$ is a subspace of $\mathbb{R}^3$.

Solution: To show this, we must consider two *general* vectors $\mathbf{x}, \mathbf{y} \in S_{a,b,c}$, i.e.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \tag{1.3}$$

where the components of these two vectors must satisfy the equations

[10] I stress this point because a common error in this case is to show that the subset is closed for one or two *specific* vectors in the subset. This is clearly insufficient as there could be other vectors in the subset which fail to satisfy the closure conditions. (See the next two footnotes.)

1. $ax_1 + bx_2 + cx_3 = 0$, and

2. $ay_1 + by_2 + cy_3 = 0$

as this is the condition that any vector must satisfy to be in $S_{a,b,c}$. We can now proceed to verify that $S_{a,b,c}$ is a subspace by checking that it is closed under both vector addition and scalar multiplication.

To see that $S_{a,b,c}$ is closed under vector addition, we note that using our rules, we have:

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix} \tag{1.4}$$

and this new vector will also be in $S_{a,b,c}$ if its components also satisfy the condition, i.e. it will be in $S_{a,b,c}$ *provided* that:

$$a(x_1 + y_1) + b(x_2 + y_2) + c(x_3 + y_3) = 0$$

and this is the case because,

$$a(x_1 + y_1) + b(x_2 + y_2) + c(x_3 + y_3) = (ax_1 + bx_2 + cx_3) + (ay_1 + by_2 + cy_3) = 0 + 0 = 0$$

where we have used the two equations — namely, (1) and (2) from above — which we know must be satisfied by the vectors $\mathbf{x}$ and $\mathbf{y}$. Consequently, as we started with two *general* vectors in $S_{a,b,c}$, we have shown that the vector sum of *any* two vectors in $S_{a,b,c}$ will give us another vector in $S_{a,b,c}$, i.e. $S_{a,b,c}$ is closed under vector addition.[11]

Then, to see that $S_{a,b,c}$ is closed under scalar multiplication, we note that using our rules, we have:

$$\alpha\mathbf{x} = \alpha \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \alpha x_3 \end{bmatrix} \tag{1.5}$$

and this new vector will also be in $S_{a,b,c}$ if its components also satisfy the condition, i.e. it will be in $S_{a,b,c}$ *provided* that:

$$a(\alpha x_1) + b(\alpha x_2) + c(\alpha x_3) = 0$$

and this is the case because,

$$a(\alpha x_1) + b(\alpha x_2) + c(\alpha x_3) = \alpha(ax_1 + bx_2 + cx_3) = \alpha \cdot 0 = 0$$

where we have, again, used equation (1) from above (which we know must be satisfied by the vector $\mathbf{x}$). Consequently, as we started with a *general* vector in $S_{a,b,c}$, we have shown that the scalar multiple of *any* vector in $S_{a,b,c}$ with *any* scalar, $\alpha \in \mathbb{R}$ will give us another vector in $S_{a,b,c}$, i.e. $S_{a,b,c}$ is closed under scalar multiplication.[12]

Thus, by Theorem 1.4, we have demonstrated that $S_{a,b,c}$ is a subspace of $\mathbb{R}^3$. (You may care to notice that the additive identity — namely the null vector, $\mathbf{0}$ — is contained within $S_{a,b,c}$ because it represents the trivial solution of the equation that defines the set (i.e. $a0 + b0 + c0$ is clearly 0 as required). Further, this fact is included in our justification because we have shown that $S_{a,b,c}$ is closed under scalar multiplication, and so, if we choose $\alpha = 0$ we get the null vector. Also, we can see that each vector $\mathbf{x} \in S_{a,b,c}$ has an additive inverse in $S_{a,b,c}$ which we can obtain, again

[11] As mentioned in the previous footnote, to prove that a subset of $\mathbb{R}^3$, $S_{1,1,1}$ say, is closed under vector addition it would be insufficient to use a pair of specific vectors. For example, the vectors $[1, -1, 0]^t$ and $[1, 0, -1]^t$ are both in $S_{1,1,1}$ because they both satisfy the equation $x_1 + x_2 + x_3 = 0$. Further, we can see that $[1, -1, 0]^t + [1, 0, -1]^t = [2, -1, -1]^t$ is in $S_{1,1,1}$ as it too satisfies the membership criterion. But, it should be clear that this doesn't show that $S_{1,1,1}$ is a vector space because we have not considered how the infinite number of other vectors in $S_{1,1,1}$ behave under vector addition. We could, possibly, say that we have amassed some 'evidence' for $S_{1,1,1}$ being a vector space (as we have not found a counter-example). But this is maths, and not science, and a 'quasi'-scientific methodology is not of much use here.

[12] Again, we can take the vector $[1, 0, -1]^t$ which is in $S_{1,1,1}$ and multiply it by a specific scalar, say 3, to get $[3, 0, -3]^t$ which is another vector in $S_{1,1,1}$; or we could multiply it by a general scalar, $\alpha \in \mathbb{R}$ say, to get $[\alpha, 0, -\alpha]^t$ which also gives us a vector in $S_{1,1,1}$. However, either way, this is not enough (for the same reasons that were outlined in the previous footnote).

due to the fact that we have shown that $S_{a,b,c}$ is closed under scalar multiplication, by multiplying $\mathbf{x}$ by the scalar $\alpha = -1$. Thus, $S_{a,b,c}$ contains all the things that we expect a vector space to contain!) ♣

In order to firm up your 'intuitions' as to what sorts of subsets are vector spaces, we can see that, geometrically, for any choice of our indices (i.e. a, b and c) we will get a plane going through the origin — for example, see Figure 1.2. Except, of course, when we have $a = b = c = 0$, in which case,

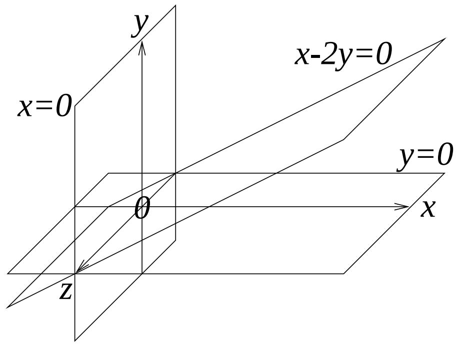

Figure 1.2: *Three subspaces of* $\mathbb{R}^3$*: The planes* $x = 0$, $y = 0$ *and* $x - 2y = 0$ *(corresponding to the subsets* $S_{1,0,0}$, $S_{0,1,0}$ *and* $S_{1,-2,0}$ *respectively).*

the membership criterion reduces to the identity $0 = 0$, which is satisfied for all values of x_1, x_2 and x_3. Thus, in this [special] case, we get the vector space $\mathbb{R}^3$! Further, notice that the requirement that the vector space be closed under vector addition and scalar multiplication means, as mentioned before, that using these operations on vectors in the subspace, will always give us other vectors in the subspace. The reason for this is clear if we look at Figure 1.3. Here the vectors are just the *position* vectors of the points in the plane represented by the subset; and most notably, all of these

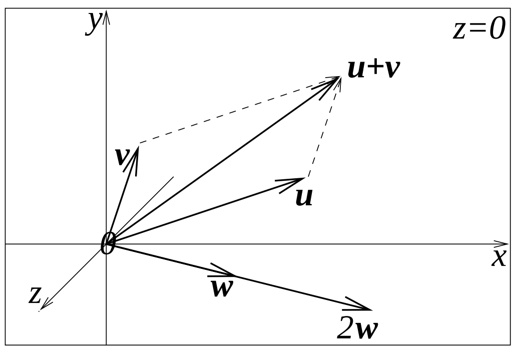

Figure 1.3: *The plane* $z = 0$ *(corresponding to the subset* $S_{0,0,1}$*). Notice that the [position] vectors* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in S_{0,0,1}$ *[corresponding to the points* $(u_1, u_2, 0), (v_1, v_2, 0), (w_1, w_2, 0)$ *in the plane* $z = 0$*] all lie in the plane, as does the vector sum,* $\mathbf{u} + \mathbf{v}$ *and the scalar multiple,* $2\mathbf{w}$ *[which give the position vectors of the points* $(u_1 + v_1, u_2 + v_2, 0)$ *and* $(2w_1, 2w_2, 0)$ *respectively; and, as expected, these points are clearly in the plane* $z = 0$ *too].*

vectors actually *lie* in the plane due to the fact that the plane contains the origin. Indeed, this is the reason *why* the two operations always give us vectors that lie within the subspace.

To show that a subset is not a subspace:

We have two different ways (although mathematically, they are just two different ways of using the same method) of showing that a subset is not a vector space.[13]

[13]Note that a particularly simple way of achieving this end is to show that the subset in question does not contain something which is essential for a vector space. For example, a subset cannot be a vector space, and hence a subspace, if it doesn't contain the additive identity, or the required additive inverses. This will be illustrated in due course.

Method I: General violation of the subspace conditions.

This method is based on showing that one of the conditions in Theorem 1.4 doesn't generally hold for the subset under consideration. Clearly, we only need to show that one of the conditions fails; but here, for the sake of illustration, we shall consider whether one, or both, of the conditions fail to hold.

Example: Show that the subset $S_{a,b,c,r} = \{\mathbf{x} \mid ax_1 + bx_2 + cx_3 = r\}$ (where a, b, c and r are [given] real numbers and $r \neq 0$) of $\mathbb{R}^3$ is not a vector space, i.e. show that $S_{a,b,c,r}$ is not a subspace of $\mathbb{R}^3$.[14]

Solution: To show this, we must consider two *general* vectors $\mathbf{x}, \mathbf{y} \in S_{a,b,c,r}$, as before, but they will be the same as the vectors we considered in the previous example (see Equation 1.3) *except* that their components will now have to satisfy the two equations

1. $ax_1 + bx_2 + cx_3 = r$, and
2. $ay_1 + by_2 + cy_3 = r$

as this is the condition that any vector must satisfy to be in $S_{a,b,c,r}$. We can now proceed to verify that $S_{a,b,c,r}$ is not a subspace by showing that it fails to be closed under vector addition and scalar multiplication.

To see that $S_{a,b,c,r}$ is not closed under vector addition, we notice that Equation 1.4 tells us that the vector sum will be in $S_{a,b,c,r}$ *provided* that:

$$a(x_1 + y_1) + b(x_2 + y_2) + c(x_3 + y_3) = r.$$

But, this is not the case because,

$$a(x_1 + y_1) + b(x_2 + y_2) + c(x_3 + y_3) = (ax_1 + bx_2 + cx_3) + (ay_1 + by_2 + cy_3) = r + r = 2r \neq r,$$ [15]

where we have used the two equations — namely, (1) and (2) from above — which we know must be satisfied by the vectors $\mathbf{x}$ and $\mathbf{y}$. Consequently, as we started with two *general* vectors in $S_{a,b,c,r}$, we have shown that the vector sum of *any* two vectors in $S_{a,b,c,r}$ will *not* give us another vector in $S_{a,b,c,r}$, i.e. $S_{a,b,c,r}$ is *not* closed under vector addition.

Nor is $S_{a,b,c,r}$ closed under scalar multiplication, we notice that Equation 1.5 tells us that the vector sum will be in $S_{a,b,c,r}$ *provided* that:

$$a(\alpha x_1) + b(\alpha x_2) + c(\alpha x_3) = r.$$

But, this is not the case because, for any $\alpha \neq 1$,

$$a(\alpha x_1) + b(\alpha x_2) + c(\alpha x_3) = \alpha(ax_1 + bx_2 + cx_3) = \alpha \cdot r \neq r,$$ [16]

where we have, again, used equation (1) from above which we know must be satisfied by the vector $\mathbf{x}$. Consequently, as we started with a *general* vector in $S_{a,b,c,r}$, we have shown that, in general (i.e. if $\alpha \neq 1$), the scalar multiple of *any* vector in $S_{a,b,c,r}$ with any scalar $\alpha \in \mathbb{R}$ ($\alpha \neq 1$) will *not* give us another vector in $S_{a,b,c,r}$, i.e. $S_{a,b,c,r}$ is *not* closed under scalar multiplication.

Thus, we have shown that $S_{a,b,c,r}$ satisfies neither of the conditions required of a subspace by Theorem 1.4 and therefore, it is not a subspace of $\mathbb{R}^3$. ♣

Method II: Counter-examples to the subspace conditions.

This method is based on showing that one of the conditions in Theorem 1.4 doesn't hold for some *specific* vector[s] in the subset under consideration. Again, we only need to find a counter-example for one of the conditions; but here, for the sake of illustration, we shall look at a case where we can

[14] As per the last footnote, this is obviously not a vector space because it does not contain the additive identity. That is, the null vector, $\mathbf{0} \notin S_{a,b,c,r}$ as $a \cdot 0 + b \cdot 0 + c \cdot 0 = 0 \neq r$.

[15] Unless, of course, $r = 0$. But, we have excluded this case in our definition of $S_{a,b,c,r}$ because, as we saw before, this case *would* give us a subspace.

[16] Unless, of course, $r = 0$, but we discussed this case in the previous footnote.

find counter-examples to both conditions.

Example: Show that the subset $S_{a,b,c,r} = \{\mathbf{x} \mid ax_1 + bx_2 + cx_3 = r\}$ (where a, b, c and r are [given] real numbers and a, b, $r \neq 0$)[17] of $\mathbb{R}^3$ is not a vector space, i.e. show that $S_{a,b,c,r}$ is not a subspace of $\mathbb{R}^3$.

Solution: To find a counter-example to the requirement that the subset should be closed under vector addition, take two vectors, say, $[r/a, 0, 0]^t$ and $[0, r/b, 0]^t$, that are in $S_{a,b,c,r}$. Then notice that their vector sum, i.e. the vector $[r/a, r/b, 0]^t$, is not in $S_{a,b,c,r}$ as we require it to be [because $a(r/a) + b(r/b) + 0 = r + r = 2r \neq r$]. Thus, we have found a counter-example and so $S_{a,b,c,r}$ cannot be a subspace by Theorem 1.4.

Similarly, to find a counter-example to the requirement that the subset should be closed under scalar multiplication, take a vector and a scalar, say, $[r/a, 0, 0]^t$ and the number 3, that are in $S_{a,b,c,r}$ and $\mathbb{R}$ respectively. Then notice that the scalar multiple of this vector with this scalar, i.e. the vector $[3r/a, 0, 0]^t$, is not in $S_{a,b,c,r}$ as we require it to be [because $a(3r/a) + 0 + 0 = 3r \neq r$]. Thus, we have found a counter-example and so $S_{a,b,c,r}$ cannot be a subspace by Theorem 1.4.

Notice however that not every choice of scalar and vector will give a counter-example. Observe that any pair of vectors will yield a counter-example to the requirement that the subset is closed under vector addition (this can be seen from the fact that, in the first method, we showed that the sum of *any* two vectors will give a vector where the membership criterion gives $2r \neq r$ and this is never satisfied because $r \neq 0$). Although, if we take *any* vector and multiply it by the scalar $\alpha = 1$ we will get a vector that *does* satisfy the membership criterion (this special case was noted when we were showing, again in Method I, that the subset was not closed under scalar multiplication) and so consequently, this will not serve as a counter-example. However, it should be clear that choosing any other value for the scalar, we will get a counter-example $\forall \mathbf{x} \in S_{a,b,c,r}$ (as clearly, $\alpha \cdot r \neq r$ if $\alpha \neq 1$; recall that $r \neq 0$ too!). ♣

Also, in order to firm up your 'intuitions' as to what sorts of subsets are not vector spaces, we can see that geometrically, for any choice of our indices (i.e. a, b, c and r) we will get a plane that does not go through the origin (as $r \neq 0$); for example, see Figure 1.4. Further, we can see that,

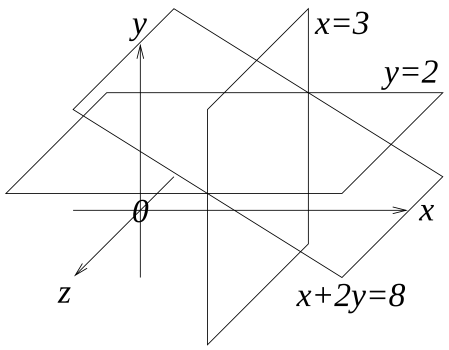

Figure 1.4: *Three planes that are not subspaces of* $\mathbb{R}^3$*: The planes* $x = 3$, $y = 2$ *and* $x + 2y = 8$ *(corresponding to the subsets* $S_{1,0,0,3}$, $S_{0,1,0,2}$ *and* $S_{1,2,0,8}$ *respectively).*

geometrically, the violation of the closure conditions corresponds to the fact that the position vectors of the points in the plane do not lie in the plane itself — see Figure 1.5 — and so, when we perform the operations of vector addition or scalar multiplication on them, we get the position vectors of points that are not in the plane. Consequently, the subset cannot be closed.[18]

[17] In this example, we shall also assume that $a \neq 0$ and $b \neq 0$ (the reason for this assumption is obvious when you look at the solution!). To see what happens if this isn't the case, see Problem Sheet 1.

[18] If you are interested in what happens when we consider the 'special' subset $S_{0,0,0,r}$ (recall, something 'special' happened when we looked at the subspace $S_{0,0,0}$), then notice that when $a = b = c = 0$, the membership criterion reduces to $0 = r$, and so, as $r \neq 0$, we get a contradiction. This means that *no* vector can belong to the subset $S_{0,0,0,r}$, i.e. $S_{0,0,0,r}$ is the empty set, and as we have seen, this is not a vector space.

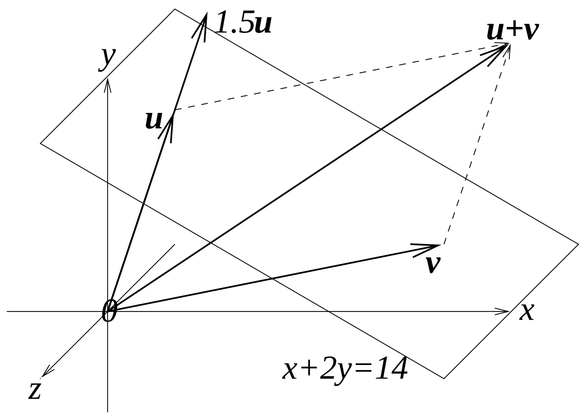

Figure 1.5: *The plane* $x + 2y = 14$ *(corresponding to the subset* $S_{1,2,0,14}$*). Notice that the [position] vectors* $\mathbf{u}, \mathbf{v} \in S_{1,2,0,14}$ *[corresponding to the points* (u_1, u_2, u_3) *and* (v_1, v_2, v_3) *in the plane* $x+2y = 14$*] do not lie in the plane. Further, the vector sum,* $\mathbf{u} + \mathbf{v}$ *and the scalar multiple,* $1.5\mathbf{u}$ *do not give us [the position vectors of] points in the plane.*

1.5 Learning Objectives

At the end of this hand-out you should:

- Be familiar with the logical and set theoretical notation introduced in Section 1.1.
- Understand that a vector space must contain certain vectors and, in particular, it must be closed under the operations of vector addition and scalar multiplication as specified in Section 1.2.
- Be aware of the three examples of a vector space given in Section 1.3 and, in particular, you should know how to represent the vectors that they contain and understand how the operations of vector addition and scalar multiplication are defined within them.
- Understand that some sub*sets* of a vector space will also be a vector space and that we call these *subspaces*. As outlined in Section 1.4.1, you should know that a sub*set* is a sub*space* if and only if it is closed under the operations of vector addition and scalar multiplication.
- Be able to prove, as shown in Section 1.4.2, that:
  - A subset is a subspace (by showing that it satisfies the closure conditions given in Theorem 1.4 using *general* vectors).
  - A subset is not a subspace (by finding a counterexample to, or showing that general vectors in the set fail to satisfy, the closure conditions given in Theorem 1.4).

  as appropriate for the set in question. You should also start to develop a geometric intuition about which sub*sets* of $\mathbb{R}^3$ are, and which are not, sub*spaces*.

This material will be developed in Problem Sheet 1.