Department of Mathematics

PURE MATHS 211-23

Linear Algebra

LINEAR ALGEBRA

Printed at the University of New England, January, 1979.

IN REPLY PLEASE QUOTE
REF

TELEPHONE ARMIDALE [illegible]
AREA CODE [illegible]
TELEX NUMBER [illegible]
POST CODE [illegible]

THE UNIVERSITY OF NEW ENGLAND

ARMIDALE, N.S.W.

DEPARTMENT OF MATHEMATICS,
2nd February, 1979

Dear PM211 student,

Enclosed are notes for the first half of the Linear Algebra course. These should be covered by about the end of April. Notes and assignments for the later sections of the course will be sent to you later in the semester.

These notes cover all the definitions, results and examples which you will be expected to know, and should serve as a course "summary" as well as a suggested "learning order".

However, except where they are not included or insufficiently explained, proofs and worked examples have been referred to the text:

Bernard Kolman "Elementary Linear Algebra", 2nd Edition, 1977, MacMillan.

Henceforth this will be referred to as K. .

The notations and conventions adopted in these notes and those in K. differ considerably. Each significant difference is noted on the first occurrence of it in the notes. You should become familiar with these, so that you can readily translate from one to the other and be able to read or work problems in either. (The preparation of a parallel table of notations might be a useful exercise.) Provided it is clear what you mean I don't mind what notations or style you choose to use in assignments and exams so find one with which you feel comfortable.

Occasionally references have also been made to the linear algebra sections of your calculus text:

Kaplan and Lewis "Calculus and Linear Algebra", Vol. 2 (or combined edition).

This is referred to as K. and L.

Other references are:

Mary Tropper "Linear Algebra", Nelson.

(This is an excellent book, though the style is somewhat terse.)

Warren Brisley "A Basis for Linear Algebra" Wiley, Australia.

(This is at a slightly lower level than our course, but could provide a useful and interesting introduction to most of the material.)

"The Schaum Outline Series" books on Linear Algebra and Matrices, may be of some value.

You may find the early part of the course difficult. It is a different kind of mathematics than you are probably used to. However with perseverance most students find that it eventually fits together and really isn't all that hard.

.../ii

From the outset it is essential that you learn the definition of a linear space (§I of the notes) - write the definition on the ceiling above your bed and recite it through each night. Don't waste too much time worrying about what a 'linear space' is; a feel for it should develop as you progress through the course. 'Linear space' is an abstract concept not a concrete object, though particular examples of it are. <u>Anything</u> which can be made to satisfy the axioms is a linear space.

Knowing definitions and results (not only like a parrot but also with understanding as to what they are saying) is essential and to this end the working of problems is an integral part of the course. Exercises should be attempted as they are encountered not let accumulate until the end. Don't be discouraged if you are unable to do some of them. All mathematicians, at all levels, suffer from this.

Remember, Theorems, lemmas etc. once established are the tools with which we work. You should be constantly on the look out for ways of using previous results to simplify and assist in the proofs of current problems. When you do make an appeal to an earlier result in the solution of a problem acknowledge it with a statement like "... follows by the theorem which states ...".

Always test that you have really learned what you have been studying. After having solved a problem, possibly with the aid of notes, books etc., think back through it without these aids, then check a couple of days later to make sure you can still do it without the aids. This often takes only a few moments concentration while waiting for a bus. Your bookwork should be treated similarly. Many students find 'giving a lesson' on the material to an "imaginary" student a useful device.

Comments, criticisms and requests for additional explanations or assistance are welcomed and will be dealt with as promptly as possible.

Wishing you success and enjoyment in your studies.

Yours sincerely,

Brailey Sims
(Lecturer)

<u>Assignments</u>

There will be 4 assignments. All problems should be attempted but don't delay sending in an assignment for too long because you can't do some of it. Always include an indication of your unsuccessful attempts. These may give us some idea of where you are going wrong and so greatly improve our chances of helping you. It may only be possible to mark a selection of problems from each of the assignments however an effort will be made to supply solutions for all the problems.

CONTENTS

Linear Algebra Assignments

Many of the assignment questions are more "theoretical" than the Exercises given in K. However as preparation for doing the assignment you should work through as many of the exercises in relevant sections of K. as you can. I have chosen K. as the textbook mainly because of the large number of examples, and exercises of an elementary nature. Make use of these to improve your understanding of material in the course.

ASSIGNMENT I (due Friday 23rd March)

1) Show that the following cancellation laws hold in any linear space.
If $\lambda \neq 0$ and $\lambda \underset{\sim}{a} = \lambda \underset{\sim}{b}$ then $\underset{\sim}{a} = \underset{\sim}{b}$.
If $\underset{\sim}{a} \neq \underset{\sim}{0}$ and $\lambda \underset{\sim}{a} = \mu \underset{\sim}{a}$ then $\lambda = \mu$.
If $\underset{\sim}{a} + \underset{\sim}{b} = \underset{\sim}{a} + \underset{\sim}{c}$ then $\underset{\sim}{b} = \underset{\sim}{c}$.
Also prove; $\lambda \underset{\sim}{0} = \underset{\sim}{0}$ for all scalars λ,
and $-(-\underset{\sim}{a}) = \underset{\sim}{a}$ for all $\underset{\sim}{a} \in V$.

2) On $\mathbb{R}^+$ define an operation of addition $\oplus$ by $x \oplus y = xy$ (ordinary multiplication) and define scalar multiplication by $\lambda x = x^{\lambda}$. Equipped with these two operations show that $\mathbb{R}$ is a linear space.

3) Complete the details for Example (5) §II, p.7 of the notes.

4) Which of the following subsets of the linear space of 2×2 matrices are subspaces?
   a) the set of all Singular Matrices
   b) the set of all Nonsingular Matrices
   c) the set of all Symmetric Matrices

(See K. Chapter I for definitions.)

5) Exercise 9 of §III p. 12 of the notes.

6) Let V be the linear space of all Real valued functions on $\mathbb{R}$, i.e. $V = F(\mathbb{R},\mathbb{R})$. Which of the following subsets of V are subspaces?
   a) The non-negative functions, i.e. $f(x) \geq 0$ for all $x \in \mathbb{R}$.
   b) The constant functions, i.e. $f(x) = c$ for all $x \in \mathbb{R}$.
   c) Functions such that $f(0) = 0$.
   d) Functions such that $f(0) = 5$.
   e) The even Functions, i.e. $f(-x) = f(x)$.
   f) The odd Functions, i.e. $f(-x) = -f(x)$.

7) a) Let $\underset{\sim}{v} = (v_1, v_2)$ be any vector in V_2, show that $\langle\{\underset{\sim}{v}\}\rangle$ is a line through the origin (zero vector). [See Your PM111 notes for the vector definition of a line.]

b) Show that any line through the origin in V_2 is a subspace of V_2.

c) Show that a line in V_2 not passing through the origin is not a subspace of V_2.

*d) Show that the only subspaces of V_2 are $\{\underset{\sim}{0}\}$, V_2 itself, and any line passing through the origin.

8) In V_3 show that the span of (2,1,3) and (1,2,0) is a plane passing through the origin.

9) In $F(\mathbb{R},\mathbb{R})$ let V_e and V_o be the subspaces of all even and all odd functions respectively [see assignment problem 6 e) and f)]. Show that $F(\mathbb{R},\mathbb{R}) = V_e \oplus V_o$.

10) For the linear space V let U and W be subspaces such that $V = U \oplus W$, show that for every vector $\underset{\sim}{v} \in V$ there exists a UNIQUE pair of vectors $\underset{\sim}{u} \in U$ and $\underset{\sim}{w} \in W$ such that $\underset{\sim}{v} = \underset{\sim}{u} + \underset{\sim}{w}$.

ASSIGNMENT II (due on Friday, 20th April)

1) Determine which of the following mappings are linear.

(a) $T: V_3 \to V_3$ defined by $T(x_1,x_2,x_3) = (x_1+2x_3, x_1, 3x_2+x_1)$

(b) $T: P_n(\mathbb{R}) \to P_{n+1}(\mathbb{R})$ defined by $P(p)(x) = xp(x)$

(c) $T: V_2 \to V_3$ defined by $T(x_1,x_2) = (x_1-x_2, x_1+1, x_2)$

(d) $T: C(\mathbb{R}) \to C(\mathbb{R})$ defined by $T(f)(x) = f(x+1)$.

For each of the mappings in (a) to (d) which you have proved to be linear find

(i) its range

(ii) its kernel,

hence decide whether or not the mapping is invertible. If it is, find its inverse.

2) Exercises: 2 (p.17), 17 No.1 (p.22), 14 (p.21) and 3 (p.23) of the the Notes.

3) a) Define T: $\mathcal{F}(\mathbb{R},\mathbb{R}) \to \mathcal{F}(\mathbb{R},\mathbb{R})$ by $T(f)(x) = \frac{f(x) + f(-x)}{2}$

   i) Show that: T is linear

   ii) Find Ker T and $T(\mathcal{F}(\mathbb{R},\mathbb{R}))$.

   iii) Show that $\mathcal{F}(\mathbb{R},\mathbb{R}) = \text{Ker } T \oplus T(\mathcal{F}(\mathbb{R},\mathbb{R}))$.

   iv) Show that $T = T^2 (= T \circ T)$.

*b) Any linear mapping P from a linear space U into itself with the property that $P = P^2$ is termed a <u>projection</u>. Thus T of part a) is a projection by iv). Generalize iii) of part a) by showing that for any projection $P \in \mathcal{L}(U,U)$ we have

$$U = \text{Ker } P \oplus P(U).$$

4) Exercises (14) p.33 of §VI.

5) Does the set of vectors $\{x^2+1, x-2\}$ span $P_2(\mathbb{R})$?

6) Which of the following sets of vectors are linearly dependent? For those which are, express one of the vectors as a linear combination of the others.

   i) (1,2,1), (2,3,4), (4,5,10).

   ii) $\{2x^2+x, x+3, x\}$ in $\mathcal{P}(\mathbb{R})$.

   iii) $\{\cos^2 t, \sin^2 t, \cos 2t\}$ in $C(\mathbb{R})$.

7) Find a basis of V_3 which includes the vectors $\begin{bmatrix}2\\1\\3\end{bmatrix}$, $\begin{bmatrix}1\\2\\0\end{bmatrix}$.

8) Find a basis for the solution space of the linear homogeneous system

$$\begin{bmatrix}1 & 0 & 2\\2 & 1 & 3\\3 & 1 & 2\end{bmatrix} \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

9) Classify all subspaces of $\mathbb{R}^3$.

10) (i) If $U = \{\text{all vectors of form } [\alpha, 2\alpha, \beta]\} \leq V_3$, and

$$V = \{\text{all of form } [p,q,2p]\} \leq V_3$$

find a similar description for $U \cap V$. Find also $U + V$.

(ii) Find a pair of subspaces of V_2 such that $U \cap V = \{0\}$ and $U + V = V_2$.

(iii) Show that, in V_3 $\langle\{[1,0,0], [0,0,0]\}\rangle \cap \langle\{[0,0,1]\}\rangle = \{0\}$, but that their sum is V_3.

(iv) Find two subspaces U and V of V_3 such that $\{0\} \lneqq U \lneqq V \lneqq V_3$.

(v) Show that the following is impossible:

$$U \leq V_3 \text{ and } V \leq V_3 \text{ and } U \cap V = \{0\} \text{ and}$$
$$\dim(U) = \dim(V) \text{ and } U + V = V_3.$$

(vi) Find two subspaces U, V of V_4 such that $U \cap V = \{0\}$, $\dim(U) = \dim(V)$ and $U + V = V_4$.

PREREQUISITES

While it is peripheral to our main task, the early first year work on Vector Geometry in 2 and 3 dimensions (Lectures 1 to 7, the work on p.17, and lectures 10 to 12 of the 1978 Pure Mathematics 111 External Notes) will help us to motivate and interpret much of our material. You should therefore reacquaint yourself with this work and keep it in mind throughout the course. (See also K. pp.64-67 and Chapter 3, section 3.1 excluding the Cross Product on pp.120-122.)

Your first year work on Matrices and Simultaneous Linear Equations (Lectures 13 to 17 of the 1978 External Notes) will be assumed and used extensively later in the course. At the appropriate places you should therefore thoroughly revise this material. This material is also covered in Chapter 1 (pp.5-63) of K., which could serve as an alternative to your first year notes and with which you should familiarise yourself.

In the first few sections of this course we will recover and enlarge upon your first year work on Vector Spaces (Lectures 18 to 20, of the 1978 External Notes); if you have these notes available, you would be well advised to revise them (and the exercises) before reading further.
[Note: The material referred to above is also present in the 1976 and 1977 External Notes for Pure Mathematics 111, however the placing within the notes has varied from year to year. Some copies of the 1978 notes are available on request for students who did not complete Pure Mathematics 111 at U.N.E. within the last three years.]

NOTES ON LINEAR ALGEBRA

§1 Definition and some elementary consequences

DEFINITION 1: *By a* LINEAR SPACE *(vector space*[1]*) OVER THE REAL NUMBERS*[2] *we mean:*

A set, V, (the elements of which, sometimes referred to as vectors, will in these notes be denoted by $\underset{\sim}{a}, \underset{\sim}{b}, \ldots, \underset{\sim}{z}$. Note: K. uses greek letters $\alpha, \beta, \ldots,$ while other books frequently use heavy type) *for which two operations are defined -*

Vector addition, *associating with each pair of elements* $\underset{\sim}{a}, \underset{\sim}{b}$ *of* V *another element of* V *which we denote by* $\underset{\sim}{a} + \underset{\sim}{b}$ (Note: K. uses $\oplus$ in place of +. While this serves to emphasize that + represents a formal operation, which which may have little or no connection with ordinary addition, it is a clumsy and unusual notation.) *and*

Scalar multiplication, *associating with each real number (scalar)* λ *and element* $\underset{\sim}{a}$ *of* V *another element of* V *which we denote by* $\lambda\underset{\sim}{a}$. (Note: K. chooses to denote scalars by lower case letters: a, b, c, etc., and uses $\odot$ to represent scalar multiplication. However we will adopt the more usual convention of using greek letters for scalars and rely on juxtaposition to indicate scalar multiplication.)

(definition continued over page)

---

(1) I prefer, and will use, the more descriptive term "linear space" rather than the emotive name "vector space" which often causes confusion by suggesting too strong a connection between our object of study and the 'vectors' of physics or elementary geometry.

(2) In this context the important property of the real numbers $\mathcal{R}$ is that of being a FIELD.
That is, a set on which two binary operations + and × are defined and satisfy:

$\lambda, \mu \in \mathcal{R} \Rightarrow \lambda + \mu \in \mathcal{R}$.	$\lambda, \mu \in \mathcal{R} \Rightarrow \lambda \times \mu \in \mathcal{R}$.
$\lambda + \mu = \mu + \lambda$.	$\lambda \times \mu = \mu \times \lambda$.
$(\lambda+\mu) + \eta = \lambda + (\mu+\eta)$.	$\lambda \times (\mu\times\eta) = (\lambda\times\mu) \times \eta$.
$\exists\ 0 \in \mathcal{R}$ such that $0 + \lambda = \lambda$.	$\exists\ 1 \in \mathcal{R}$ such that $1\times\lambda = \lambda$ all $\lambda \in \mathcal{R}$.
For each $\lambda \in \mathcal{R}\ \exists\ -\lambda \in \mathcal{R}$ such that $\lambda + (-\lambda) = 0$.	For each $\lambda \in \mathcal{R}\backslash\{0\}\ \exists\ \lambda^{-1} \in \mathcal{R}$ such that $\lambda \times \lambda^{-1} = 1$.

$$\lambda \times 0 = 0.$$

$$\lambda \times (\mu+\eta) = (\lambda\times\mu) + (\lambda\times\eta).$$

In some deeper results the additional structure of $\mathcal{R}$ (e.g. the presence of an order relationship "≤") is important. However the elementary theory of linear spaces could be developed with any field substituted for $\mathcal{R}$. Indeed, later in the course we will find it useful to replace $\mathcal{R}$ by the field of complex numbers $\mathcal{C}$, which enjoys the property of being "algebraically closed" (any polynomial of degree n has n roots in $\mathcal{C}$).

These two operations are required to satisfy the following axioms.

(i) $\underset{\sim}{a} + \underset{\sim}{b} = \underset{\sim}{b} + \underset{\sim}{a}$ *for all* $\underset{\sim}{a}, \underset{\sim}{b} \in V$ (commutivity).

(ii) $\underset{\sim}{a} + (\underset{\sim}{b} + \underset{\sim}{c}) = (\underset{\sim}{a} + \underset{\sim}{b}) + \underset{\sim}{c}$ *for all* $\underset{\sim}{a}$, $\underset{\sim}{b}$ *and* $\underset{\sim}{c}$ (associativity)[3].

(iii) *There exists an element* 0 *in* V *such that* $\underset{\sim}{0} + \underset{\sim}{a} = \underset{\sim}{a}$ *for all* $\underset{\sim}{a} \in V$ (existence of a zero element in V).

(iv) *For each* $a \in V$ *there exists an element* $\underset{\sim}{a}'$ *in* V *such that* $\underset{\sim}{a} + \underset{\sim}{a}' = \underset{\sim}{0}$ (existence of additive inverses).

[Axioms (i) to (iv) are summarized by the statement: V is an abelian, or commutative, group under vector addition].

(v) $\lambda(\underset{\sim}{a} + \underset{\sim}{b}) = \lambda\underset{\sim}{a} + \lambda\underset{\sim}{b}$ *for all* $\lambda \in \mathcal{R}$ *and* $\underset{\sim}{a}, \underset{\sim}{b} \in V$ (scalar multiplication is distributive over vector addition).

(vi) $(\lambda + \mu)\underset{\sim}{a} = \lambda\underset{\sim}{a} + \lambda\underset{\sim}{a}$ *for all* $\lambda, \mu \in \mathcal{R}$ *and* $\underset{\sim}{a} \in V$ (scalar multiplication is distributive over addition of scalars).

(vii) $\lambda(\mu\underset{\sim}{a}) = (\lambda\mu)\underset{\sim}{a}$ *for all* $\lambda, \mu \in \mathcal{R}$ *and* $\underset{\sim}{a} \in V$.

(viii) $1\underset{\sim}{a} = \underset{\sim}{a}$ *for all* $\underset{\sim}{a} \in V$ (N.B. 1 denotes the scalar one).

The above is a convenient working definition of a linear space which you should now compare with that of K., section 2.1, definition 2.1, p. 68 (or K. & L. section 9-1 (definition 9-10) p.641). All the axioms are the same except that our (iii) and (iv) have been replaced by the apparently stronger (and hence more difficult to verify) (3) and (4) which involve uniqueness assumptions. The equivalence of our definition with that of K. is established in the following elementary proposition.

PROPOSITION 2: *In a linear space* V,

(i) *there is a UNIQUE zero vector* $\underset{\sim}{0}$,

(ii) *for each element* $\underset{\sim}{a} \in V$ *there is a UNIQUE element* $\underset{\sim}{a}'$ *satisfying* $\underset{\sim}{a} + \underset{\sim}{a}' = \underset{\sim}{0}$. (This unique $\underset{\sim}{a}'$ is referred to as the additive inverse of $\underset{\sim}{a}$ and will be denoted by $-\underset{\sim}{a}$.)

Proof of (i)

The existence of such a zero vector is guaranteed by 1(iii). Now assume there were two such elements $\underset{\sim}{0}$ and $\underset{\sim}{0}'$, i.e.

$$\underset{\sim}{0} + \underset{\sim}{a} = \underset{\sim}{a} \quad \text{for all } \underset{\sim}{a} \in V \tag{1}$$

and

$$\underset{\sim}{0}' + \underset{\sim}{a} = \underset{\sim}{a} \quad \text{for all } \underset{\sim}{a} \in V. \tag{2}$$

---

(3) As a result of this axiom we can omit the parenthesis from expressions like $((\underset{\sim}{a}+\underset{\sim}{a})+\underset{\sim}{a})+\underset{\sim}{b}$ and write $\underset{\sim}{a} + \underset{\sim}{a} + \underset{\sim}{a} + \underset{\sim}{b}$ without introducing any ambiguity.

Then, putting $\underset{\sim}{a} = \underset{\sim}{0}'$ into (1) we have

$$\underset{\sim}{0}' = \underset{\sim}{0} + \underset{\sim}{0}'$$
$$= \underset{\sim}{0}' + \underset{\sim}{0} \quad \text{(by 1(i))}$$
$$= \underset{\sim}{0} \text{ by (2) with } \underset{\sim}{0} \text{ replacing } \underset{\sim}{a}.$$

That is $\underset{\sim}{0} = \underset{\sim}{0}'$, proving uniqueness.

The proof of (ii) is accomplished similarly and is left as an exercise. ■

EXERCISE 3: Prove that each element of a linear space has a unique additive inverse.

The next proposition establishes a certain consistency among the notations adopted above.

PROPOSITION 4: *Let* V *be a linear space over* $\mathcal{R}$, *then:*

(i) $0\underset{\sim}{a} = \underset{\sim}{0}$ *for any* $\underset{\sim}{a} \in V$;

(ii) *For any* $\underset{\sim}{a} \in V$ *the additive inverse* $-\underset{\sim}{a} = (-1)\underset{\sim}{a}$.

[In view of this we can, without confusion, write $\underset{\sim}{a} - \underset{\sim}{b}$ instead of $\underset{\sim}{a} + (-\underset{\sim}{b})$ and $-\lambda\underset{\sim}{a}$ instead of $-(\lambda\underset{\sim}{a})$ or $(-\lambda)\underset{\sim}{a}$.]

(iii) *For any* $n \in \{1,2,3,\ldots\} \subset \mathcal{R}$ *and* $\underset{\sim}{a} \in V$

$$n\underset{\sim}{a} = \underbrace{\underset{\sim}{a} + \underset{\sim}{a} + \ldots + \underset{\sim}{a}}_{n \textit{ times}}$$

Proof (i) Since $0 = 0 + 0$ we have

$$0\underset{\sim}{a} = (0 + 0)\underset{\sim}{a} = 0\underset{\sim}{a} + 0\underset{\sim}{a} \quad \text{(by (1(vi))}$$

adding $-(0\underset{\sim}{a})$ to both sides of this identity yields

$$\underset{\sim}{0} = 0\underset{\sim}{a} + (-(0\underset{\sim}{a}))$$
$$= (0\underset{\sim}{a} + 0\underset{\sim}{a}) + (-(0\underset{\sim}{a}))$$
$$= 0\underset{\sim}{a} + (0\underset{\sim}{a} + (-(0\underset{\sim}{a}))) \quad \text{(by (1(ii))}$$
$$= 0\underset{\sim}{a} + \underset{\sim}{0}$$
$$= 0\underset{\sim}{a}.$$

(ii) By (2) it suffices to show $\underset{\sim}{a} + (-1\underset{\sim}{a}) = \underset{\sim}{0}$,

now $\underset{\sim}{a} + (-1\underset{\sim}{a}) = (1\underset{\sim}{a}) + (-1\underset{\sim}{a})$ (by(1(viii))

$= (1 + (-1))\underset{\sim}{a}$ (by (1(vi))

$= 0\underset{\sim}{a}$

$= \underset{\sim}{0}$ (by (i) above).

(iii) For $n = 1$ the result is ensured by 1(viii) .

The general result is now readily established by induction on n using 1(vi) , and is left as an exercise. ■

<u>EXERCISE 5:</u> (i) Prove 4 (iii).

(ii) Show that any finite sum of scalar multiples of elements in a linear space V is an element of V. That is, if $\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_n \in V$ and $\lambda_1, \ldots, \lambda_n \in \mathcal{R}$, then $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n \in V$.

§II Examples

The importance of our concept lies in the extremely large number of "mathematical objects" which exhibit the structure of a linear space. Some of these objects are listed below. In each case you should convince yourself of the truth of Axioms I(i)-(viii) for the operations defined.

Once a general result for linear spaces has been proved we then immediately know it is true in all these examples. Thus, once we have accumulated a body of results for linear spaces, showing something is a linear space amounts to immediately knowing a lot about it.

Further, as we see below, ordinary two dimensional space is an example of a linear space. Thus when attempting to establish a result for less familiar linear spaces we can see what the problem amounts to in two dimensional space, use our "geometric intuition" to find a proof in that case, and then, provided the proof only relied on the linear space structure of Euclidean space, translate it back to obtain our result.

EXAMPLE (1):

The prototype of linear spaces is the set of directed line segments from an origin in "ordinary" 2 (or 3) dimensional space. Representing these "vectors" by their components with respect to a set of Cartesian axes, we can identify the space with the set of ordered pairs (or triplets) of real numbers $V = \{\underset{\sim}{a} = (a_1, a_2) : a_1, a_2 \in \mathbb{R}\}$ on which vector addition and scalar multiplication is defined component-wise, i.e.,

$$\underset{\sim}{a} + \underset{\sim}{b} = (a_1, a_2) + (b_1, b_2) = (a_1+b_1, a_2+b_2),$$

$$\lambda \underset{\sim}{a} = \lambda(a_1, a_2) = (\lambda a_1, \lambda a_2).$$

That these operations satisfy axioms II(i) to (viii) is easily verified. Refer to your PM111 Notes.

EXAMPLE (2):

Example (1) generalizes to the set of ordered n-tuples of real numbers $\mathbb{R}^n$, with component-wise definitions of vector addition and scalar multiplication.

We will denote this linear space by V_n. See PM111 1978 Notes, p. 76 and K. Example 1, p.68 for details.

A trivial but important special case is that of n = 1, which shows that the real numbers can be regarded as a linear space over themselves. (See K. Example 3, p.69.)

EXAMPLE (3):

Of course we need not stop at finite-tuples, the space of all real (infinite) sequences is a linear space with respect to component-wise definitions of vector addition and scalar multiplication (the proof is essentially the same as for V_n). This space is sometimes denoted by $\mathcal{R}^\infty$.

EXAMPLE (4):

The set of all m×n - matrices with real entries (m and n fixed positive integers) is a linear space over $\mathcal{R}$ when we take the usual addition of matrices as vector addition and the usual multiplication of a matrix by a real number as scalar multiplication. (K. Example 2, p.69).

EXAMPLE (5):

Let D be any set and denote by $F(D,\mathcal{R})$ the set of all real functions with domain D and co-domain the real numbers $\mathcal{R}$, i.e., the set of functions $f: D \to \mathcal{R}$.

Define addition and scalar multiplication of functions point-wise:

For $f, g \in F(D,\mathcal{R})$, $(f+g)$ maps each $x \in D$ to $f(x) + g(x)$ and for $\lambda \in \mathcal{R}$, λf maps each $x \in D$ to $\lambda f(x)$ (1).

Then with respect to these operations $F(D,\mathcal{R})$ is a linear space (2).

[Note: The role of $\mathcal{R}$ in $F(D,\mathcal{R})$ could be taken by any linear space V, and the conclusion would remain valid.]

A CHALLENGE: We need only establish Example (5). All the other examples are really special cases of (5); can you show how this is so?

---

(1) Note, this agrees with the standard practice in calculus and elsewhere when we write $(f + g)' = f' + g'$, $(\lambda f)' = \lambda f'$ etc.

(2) To prove this proceed as follows.
Verification of Axiom I(i): For $f, g \in F(D,\mathcal{R})$ and any $x \in D$

$(f + g)(x) = f(x) + g(x)$ by definition of addition.

$= g(x) + f(x)$ as f(x) and g(x) are real Numbers, the addition of which is commutative (Field axioms).

$= (g + f)(x)$ by definition of addition.

Thus the image of each $x \in D$ under either of the two functions f+g and g+f is the same, and so f+g and g+f are two names for the one function, or $f+g = g+f$.

The other axioms are established similarly (do so).

Note: to establish I(iii) we seek a function $O \in F(D,\mathcal{R})$ such that for every $f \in F(D,\mathcal{R})$ $O + f = f$, that is, for each $x \in D$ $O(x)+f(x) = f(x)$ and so $O(x) = 0$. Thus the only possible candidate for O is the function which maps every $x \in D$ to the scalar D. It only remains to check that this O does the job (obvious!).

Many other examples of linear spaces could be given now, however, using the above examples (in particular Example (5)) and the work of the next section, the verification that axioms I(i) to (viii) are satisfied will be considerably simplified, and so further examples are deferred till §IV.

§III Subspaces

DEFINITION 1: A subspace of the linear spave V is a subset of V which itself forms a linear space with the operations of vector addition and scalar multiplication which are inherited from V.

We will write $W \leqslant V$ to mean W is a subspace of the linear space V.

LEMMA 2: *Let* V *be a linear space and* W *a non-empty subset of* V. *Then* W *is a subspace of* V *if and only if*[1]

(i) *for any pair of vectors* $\underset{\sim}{a}, \underset{\sim}{b} \in W$ *we have that* $\underset{\sim}{a} + \underset{\sim}{b} \in W$ [i.e., W is closed under vector addition].

and (ii) *for any scalar* $\lambda \in \mathbb{R}$ *and vector* $\underset{\sim}{a} \in W$ *we have* $\lambda\underset{\sim}{a} \in W$ [i.e., W is closed under scalar multiplication].

Proof. See the proof of Theorem 2.2, p.72 of K.

EXERCISE 3. Let V be a linear space. Show that the single point set $\{\underset{\sim}{0}\}$ (where $\underset{\sim}{0}$ is the zero vector in V) and V itself are both subspaces. These two subspaces are known as the trivial subspaces of any linear space. A subspace of V other than $\{\underset{\sim}{0}\}$ or V is referred to as a proper subspace.

Lemma 2 is the work-saver promised at the end of §II. If we can recognise a set as a subset of a known linear space V and the operations defined on it as those inherited from V, then we need only verify 2(i) and (ii) to show it is a subspace, hence a linear space in its own right and so automatically establish the validity of axioms 1(i) - (viii) for it.

Subspaces Spanned (generated) by Sets of Vectors

Theorem 4. *Let* V *be a linear space and* S *any subset of* V, *then the set of all vectors which are finite sums of scalar multiples of elements of S is a subspace of* V.

[We will call this the subspace spanned by S and denote it by <S>. Its elements will later be referred to as (finite) linear combinations of elements of S, being of the form

$$\underset{\sim}{a} = \sum_{i=1}^{m} \lambda_i \underset{\sim}{s}_i = \lambda_1 \underset{\sim}{s}_1 + \lambda_2 \underset{\sim}{s}_2 + \ldots + \lambda_m \underset{\sim}{s}_m,$$

where $m \in N$; $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R}$ (the scalars) and $\underset{\sim}{s}_1, \underset{\sim}{s}_2, \ldots, \underset{\sim}{s}_m \in S$.

If W = <S> we will also say the set of vectors S spans W.]

---

(1) It is not difficult to see that the two conditions (i) and (ii) can be condensed into the single equivalent condition: for $\underset{\sim}{a}, \underset{\sim}{b} \in W$ and $\lambda \in \mathbb{R}$, $\underset{\sim}{a} + \lambda\underset{\sim}{b} \in W$.

Proof. Let $\underset{\sim}{a}$, $\underset{\sim}{b} \in \langle S \rangle$, then

$$\underset{\sim}{a} = \lambda_1 \underset{\sim}{s}_1 + \lambda_2 \underset{\sim}{s}_2 + \ldots + \lambda_m \underset{\sim}{s}_m$$

for some $m \in N$; $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathcal{R}$ and $\underset{\sim}{s}_1, \underset{\sim}{s}_2, \ldots, \underset{\sim}{s}_m \in S$

and $\underset{\sim}{b} = \mu_1 \underset{\sim}{t}_1 + \mu_2 \underset{\sim}{t}_2 + \ldots + \mu_n \underset{\sim}{t}_n$

for some $n \in N$; $\mu_1, \ldots, \mu_n \in \mathcal{R}$ and $\underset{\sim}{t}_1, \ldots, \underset{\sim}{t}_n \in S$.

Hence $\underset{\sim}{a} + \underset{\sim}{b} = \lambda_1 \underset{\sim}{s}_1 + \ldots + \lambda_m \underset{\sim}{s}_m + \mu_1 \underset{\sim}{t}_1 + \ldots + \mu_n \underset{\sim}{t}_n$

is a sum of scalar multiples of elements of S and so belongs to $\langle S \rangle$.

Further, for any $\lambda \in \mathcal{R}$

$$\lambda \underset{\sim}{a} = \lambda(\lambda_1 \underset{\sim}{s}_1 + \ldots + \lambda_m \underset{\sim}{s}_m)$$

$$(\lambda\lambda_1)\underset{\sim}{s}_1 + \ldots + (\lambda\lambda_m)s_m \qquad \text{(by I(iv) and (vii))}$$

and so $\lambda \underset{\sim}{a} \in \langle S \rangle$

Thus $\langle S \rangle$ is a subspace by lemma 2. ∎

Another useful way of viewing $\langle S \rangle$ is provided by

Theorem 5: *Let* V *be a linear space and* $S \subseteq V$, *then* $\langle S \rangle$ *is the smallest subspace of* V *containing S.*

i.e., If $W \leqslant V$ and $S \subseteq W$ then $\langle S \rangle \leqslant W$,

Proof. If $\underset{\sim}{a} \in \langle S \rangle$, then $\underset{\sim}{a} = \lambda_1 \underset{\sim}{s}_1 + \ldots + \lambda_m \underset{\sim}{s}_m$ where $\lambda_1, \ldots, \lambda_m \in \mathcal{R}$ and $\underset{\sim}{s}_1, \ldots, \underset{\sim}{s}_m \in S$. Since $S \subseteq W$, hence each $\underset{\sim}{s}_i \in W$, and W is a subspace so closed under scalar multiplication and vector addition it follows that $\lambda_1 \underset{\sim}{s}_1 + \ldots + \lambda_m \underset{\sim}{s}_m \in W$

i.e. $\underset{\sim}{a} \in W$ and so $\langle S \rangle \subseteq W$. ∎

Corollary: $\langle S \rangle$ *is the intersection of all the subspaces of V which contain* S.

Proof. If W is a subspace of V containing S, then by theorem 5 $\langle S \rangle \subseteq W$. Thus certainly $\langle S \rangle$ is contained in the intersection of all such subspaces. But, by Theorem 4 $\langle S \rangle$ is one such subspace and so the intersection cannot be larger than $\langle S \rangle$. ∎

Combining Subspaces

Theorem 6: *Let* V *be a linear space, and* U, W $\leqslant$ V, *then*

(i) $U \cap W \leqslant V$ (and hence any finite intersection of subspaces is a subspace)

and (ii) $U + W \leqslant V$

where U + W *is the set of all vectors of the form* $\underset{\sim}{u} + \underset{\sim}{w}$ *with* $\underset{\sim}{u} \in U$ *and* $\underset{\sim}{w} \in W$. *In fact* $U + W = \langle U \cup W \rangle$.

Proof. (i) This is left as an exercise but follows readily from lemma 2 and the definition of linear space.

(ii) Since $\langle U \cup W \rangle$ is a subspace (Theorem 4) we need only show $U + W = \langle U \cup W \rangle$.

Clearly by the definition of $\langle U \cup W \rangle$ any vector of the form $\underset{\sim}{U} + \underset{\sim}{W}$ belongs to $\langle U \cup W \rangle$, so $U + W \subseteq \langle U \cup W \rangle$.

Now if $\underset{\sim}{a} \in \langle U \cup W \rangle$ then

$$\underset{\sim}{a} = \lambda_1 \underset{\sim}{s}_1 + \ldots + \lambda_m \underset{\sim}{s}_m \text{ with } \underset{\sim}{s}_1, \ldots, \underset{\sim}{s}_m \in U \cup W$$

i.e., each $\underset{\sim}{s}_i \in U$ or W. Using commutivity of vector addition we can rearrange the terms in this sum so that all the $\underset{\sim}{s}_i$'s in U come first (followed by those in V) since U is a linear space (subspace of V) the sum of these first terms is therefore an element of U. Similarly the sum of the remaining terms will be an element of W.

Thus $\underset{\sim}{a}$ is of the form $\underset{\sim}{u} + \underset{\sim}{w}$ and so in U + W.

i.e., $\langle U \cup W \rangle \subseteq U + W$. ■

EXERCISE 7. (i) Prove (i) of Theorem 6.

(ii) The union of two subspaces U and W of the linear space V, $U \cup W$ is not necessarily a subspace.

Otherwise, from Theorems 5 and 6 we would have

$$U \cup W = \langle U \cup W \rangle = U + W,$$

give an example showing that this need not be so (Try in V_2).

REMARK 8. If U and W are subspaces of the linear space V such that $U \cap W = \{\underset{\sim}{0}\}$, the trivial subspace, then we call U + W a direct sum and write $U \oplus W$ for it.

(Note: This is not to be confused with K.'s use of $\oplus$.)

Thus to write $V = U \oplus W$ means

$$V = U + W = \{\underset{\sim}{u} + \underset{\sim}{w}: \underset{\sim}{u} \in U \text{ and } \underset{\sim}{w} \in W\}$$
$$= \langle U \cup W \rangle$$
$$\text{and} \quad U \cap W = \{\underset{\sim}{0}\}.$$

EXERCISE 9: The three problems listed here are trivial, but the results are useful in subsequent calculations.

(i) Prove, if $U \leq W$ and $W \leq V$, then $U \leq V$

(ii) If W is a subspace of V show that $\langle W \rangle = W$, hence conclude that for any subset S of the linear space V, $\langle\langle S \rangle\rangle = \langle S \rangle$ (here $\langle\langle S \rangle\rangle$ means the span of the span of S).

(iii) Let W be a subspace of V and $w_0 \in W$, show that the translate $\underset{\sim}{w}_0 + W = \{\underset{\sim}{w}_0 + \underset{\sim}{w}: \underset{\sim}{w} \in W\}$ equals W.

§IV More Examples

Further examples are obtained by identifying various subspaces (hence, linear spaces in their own right) of the spaces introduced in §II.

EXAMPLE 1. Let W be any subset of V_n (Example 2 of §I) such that the components of every vector in W satisfy a given system of linear homogeneous relationships.

Then W is a subspace of V_n.

For a proof of this see K. Example 14, p.73. However, first note that the above definition of W means there exist real numbers $a_{11}, a_{12}, \ldots, a_{21}, a_{22}, \ldots, a_{mn}$ such that $\underset{\sim}{v} = (v_1, v_2, \ldots, v_n)$ belongs to W if and only if the m linear homogeneous relationships:

$$
\begin{aligned}
a_{11}v_1 + a_{12}v_2 + \ldots + a_{1n}v_n &= 0 \\
a_{21}v_1 + a_{22}v_2 + \ldots + a_{2n}v_n &= 0 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
am_1v_1 + am_2v_2 + \ldots + a_{mn}v_n &= 0,
\end{aligned}
$$

are satisfied.

Or, equivalently, writing $\underset{\sim}{v}$ as the column $\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$

$\underset{\sim}{v} \in W$ if and only if $\underset{\sim}{v}$ is a solution of the homogeneous matrix equation

$$A\underset{\sim}{v} = \underset{\sim}{0}$$

where A is the m × n matrix $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$

A particular case of this is Example 12, p.72 of K., where n = 3 and

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

Note: It is really only necessary to establish the result for a subset of vectors whose components satisfy a single linear homogeneous relationship. The subset satisfying a system of such relationships is then the intersection of the subspaces determined by each of the relationships in the system and so by Theorem III 6 is a subspace.

EXAMPLE 2: From your first year work on sequences; if

$$x_1, x_2, \ldots, x_n, \ldots \quad \text{is a convergent sequence (limit } x)$$

and

$$y_1, y_2, \ldots, y_n, \ldots \quad \text{is a convergent sequence (limit } y),$$

then their "sum"

$$x_1+y_1, x_2+y_2, \ldots, x_n+y_n, \ldots \quad \text{has limit } x+y \text{ and so is a}$$

convergent sequence also,
and for any $\lambda \in \mathbb{R}$

$$\lambda x_1, \lambda x_2, \ldots, \lambda x_n, \ldots \quad \text{has limit } \lambda x \text{ and hence is convergent.}$$

We can now interpret these results as saying: the set $\underline{c}$ of convergent sequences is closed under addition and scalar multiplication and is therefore (Theorem 2 of III) a subspace of the space of all sequences (Example 3 of II). $\underline{m}$, the set of all bounded sequences, is another such subspace, as is $\underline{c_0}$ the set of sequences convergent to zero. Indeed $c_0 \le c \le m \le \mathbb{R}^\infty$.

EXAMPLE 3: For any interval I of real numbers (e.g. $I = [0,1]$ or $I = (-1,1)$), let $C(I)$ denote the set of all continuous functions (refer PM111 notes) from I into $\mathbb{R}$. Then $C(I)$ is a subset of $F(I,\mathbb{R})$ - §II, Example 5 - further your first year theorems:

"The sum of two continuous functions if continuous" and "any multiple of a continuous function is continuous", show that $C(I) \le F(I,\mathbb{R})$.

EXAMPLE 4: From the school results

$$\frac{d}{dx}(f(x) + g(x)) = \frac{d}{dx} f(x) + \frac{d}{dx} g(x)$$

and

$$\frac{d}{dx} \lambda f(x) = \lambda \frac{d}{dx} f(x) \qquad (\lambda \in \mathbb{R})$$

we conclude that the subset of $F(I,\mathbb{R})$ consisting of all differentiable functions is a subspace.

Similarly, the set of all n-times differentiable functions ($n \in \mathbb{N}$) is a subspace of $F(I,\mathbb{R})$, which is sometimes denoted by $\mathcal{D}_n(I)$.

EXAMPLE 5: For any interval I it is readily verified that the set $\mathcal{P}(I)$ of all polynomials on I is a subspace of $\mathcal{C}(I)$ and hence $\mathcal{F}(I,\mathcal{R})$ - see K. Example 5, p.70, but bear in mind Theorem 2 of II.

Also, the set of all polynomials on I of degree less than or equal to n ($n \in N$), $\mathcal{P}_n(I)$ is a subspace of $\mathcal{P}(I)$ (K. Example 10, p.71).

Indeed for any $n \in N$ we have:

$$\mathcal{P}_n(I) \leq \mathcal{P}_{n+1}(I) \leq \ldots \leq \mathcal{P}(I) \leq \ldots \leq \mathcal{D}_{n+1}(I) \leq \mathcal{D}_n(I) \leq \ldots \leq \mathcal{D}_1(I) \leq \mathcal{C}(I) \leq \mathcal{F}(I,\mathcal{R}).$$

EXERCISE 6: Show that the set of functions (solutions) satisfying any given (second order) linear homogeneous differential equation

$$a(x)y'' + b(x)y' + c(x)y = 0 \qquad \text{(refer PM111 Notes)}$$

form a subspace of the set of twice differentiable functions on $\mathcal{R}$.

EXERCISE 7: Let D be any set and Z be a non-empty subset of D. Show that the set of all functions in $\mathcal{F}(D,\mathcal{R})$ which are zero on Z, i.e., $\{f \in \mathcal{F}(D,\mathcal{R}) : f(z) = 0 \text{ all } z \in \mathcal{R}\}$, is a subspace of $\mathcal{F}(D,\mathcal{R})$.
This example is particularly important in the case when $D \subseteq \mathcal{R}$ and Z consists of one or two points, where it shows for example that the set of all functions satisfying the initial condition $f(t_0) = 0$ form a linear space.

Exercises 6 and 7 combined with theorem II 6 i) show that the solutions of any given linear homogeneous initial value, or boundary value, problem,

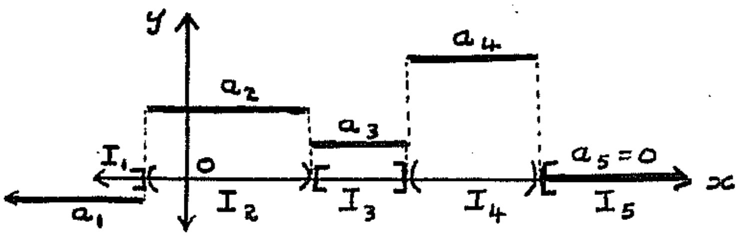
$$\text{E.g.} \quad \begin{cases} a(x)y'' + b(x)y' + c(x)y = 0 \\ y(t_0) = y(t_1) = 0, \end{cases}$$

form a linear space.

EXERCISE 8: By a step function we mean a function f from $\mathcal{R}$ to $\mathcal{R}$ of the form

$$f(x) = a_i \quad \text{for} \quad x \in I_i \qquad i = 1,2,\ldots,n$$

where $a_1, a_2, \ldots, a_n$ are a finite set of real numbers and $I_1, I_2, \ldots, I_n$ are a finite family of disjoint intervals whose union equals $\mathcal{R}$.
Show that the set of step functions $st(\mathcal{R})$ form a subspace of $\mathcal{F}(\mathcal{R},\mathcal{R})$. This subspace is of basic importance in most theories of integration.



Many other examples of linear spaces could be described. However we now have an adequate number on which to illustrate our subsequent theory of linear spaces, and by now you should be convinced that the occurrence of linear spaces is both frequent and varied.

You should now work through all the examples in K. section 2.1 including the "negative" ones, e.g. Example 11.

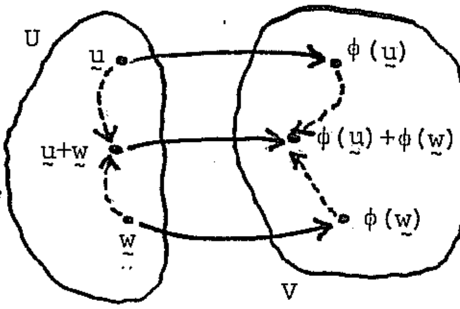
§V Homomorphisms (Part of K. Chapter 4)

Generally, a homomorphism between two similar algebraic structures is a mapping from one to the other which respects the structure.
Thus:-

DEFINITION (1) Let U, V be linear spaces over $\mathcal{R}$. By a *homomorphism* from U to V we mean a mapping $\phi: U \to V$ such that

(i) $\phi(u+w) = \phi(u) + \phi(w)$, all $u, w \in U$.
(additivity)

and (ii) $\phi(\lambda u) = \lambda\phi(u)$, all $u \in U$ and $\lambda \in \mathcal{R}$.
(scalar homogenity)



[Note in (i) the first "+" refers to vector addition in U, whereas the second "+" refers to vector addition in V. A similar remark applies to scalar multiplication in (ii).]

A mapping from one linear space into another which is both additive and scalar homogeneous is termed *linear*.[1]
Thus for linear spaces homomorphism and *linear mapping* are synonyms (henceforth we will use the latter); other synonyms are: linear function; (linear) transformation; (linear) operator.

For two linear spaces U, V we will denote by $\mathcal{L}(U,V)$ the set of all linear mappings from U into V.

Note: $\mathcal{L}(U,V)$ is a subset of $\mathcal{F}(U,V)$ and so in $\mathcal{L}(U,V)$ we have addition and scalar multiplication of linear maps defined point wise.

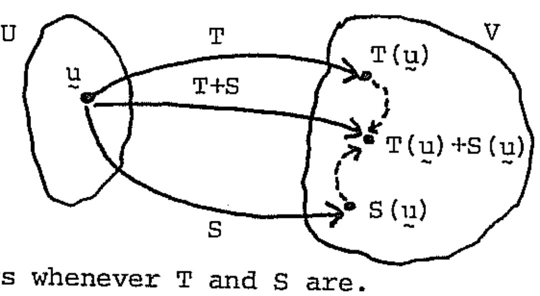
i.e. For $T, S \in \mathcal{L}(U,V)$ we have

$$(T+S)(u) = T(u) + S(u)$$

and for $\lambda \in \mathcal{R}$

$$(\lambda T)(u) = \lambda(T(u))$$

for all $u \in U$



EXERCISE 2. Show that T+S and λT are linear mappings whenever T and S are.

As a result of exercise 2), Example 5 of §II and lemma 2 of §3 we make the important observation that $\mathcal{L}(U,V)$ *is itself a linear space over* $\mathcal{R}$.

---

(1) As you may have gathered already, the term 'linear' refers to the two operations; vector addition and scalar multiplication, present in a "linear" space. Thus a linear mapping respects both of these operations, a linear combination of vectors in formed by the use of the two operations, etc.

EXAMPLES

(3) Let M be an $n \times m$ matrix and define a mapping T from V^m into V^n by

$$T(x_1, x_2, \ldots, x_m) = (y_1, y_2, \ldots, y_n) \quad \text{where}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = M \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

then it follows from the definition of matrix multiplication that T is a linear mapping.

e.g. See examples 5 and 7 of K. p.150.

Also observe that the linear mappings of K., Examples 1(p.148) and 8(p.151) can be regarded in this way for suitable matrices (find them).

(4) Let U be the subspace of $F(I,\mathbb{R})$ described in Example 4 of §IV (i.e., the set of all differentiable functions from I to $\mathbb{R}$) and let D denote the operation of differentiation, i.e., $D \equiv \frac{d}{dx}$ or $D(f) = f'$ the derivative of f, for all $f \in U$. Then D is a linear mapping from U into $F(I,\mathbb{R})$. (K., Example 2, p.148 is a special case.)

(5) Let $C(I)$ be as in Example 3 of §IV where $I = [a,b]$, then from first year we know that

$$I(f) = \int_a^b f(x)dx \quad \text{exists for each } f \in C[a,b].$$

Further, that

$$I(f+g) = I(f) + I(g)$$

$$\text{and} \qquad I(\lambda f) = \lambda I(f)$$

are standard High School results.

Hence I is a linear mapping from $C[a,b]$ into the linear space V^1 $(= \mathbb{R})$. (K., Example 4, p.149 is a special case.)

[A linear mapping from a space U into the linear space $\mathbb{R}$ is sometimes referred to as a *(linear) functional.*]

Similarly, the "primitive" mapping $I_a(f) = \int_a^x f(t)dt$ is a linear mapping from $C[a,b]$ into $F([a,b],\mathbb{R})$.

(6) Let U be the space of infinite sequences described in Example 3 of §II (or any of the subspaces of it discussed in Example 2 of §IV) then the right and left shift operators defined below are linear mappings of U into itself.

$$S_R(x_1,x_2,\ldots,x_n,\ldots) = (0,x_1,x_2,\ldots,\underset{\substack{\uparrow \\ \text{in } (n+1)\text{th place}}}{x_n},\ldots)$$

$$S_L(x_1,x_2,\ldots,x_n,\ldots) = (x_2,x_3,\ldots,\underset{\substack{\uparrow \\ \text{in } (n-1)\text{th place}}}{x_n},\ldots)$$

<u>THEOREM 7</u> (Elementary Properties of Linear Mappings)

Let U, V *be linear spaces and* $T \in \mathcal{L}(U,V)$ *then*

(i) $T(\underset{\sim}{0}) = \underset{\sim}{0}$

$\uparrow$ *zero vector in* U; $\uparrow$ *zero vector in* V

(ii) $T(\lambda\underset{\sim}{u} + \mu\underset{\sim}{\omega}) = \lambda T(\underset{\sim}{u}) + \mu T(\underset{\sim}{\omega})$ *all* $\lambda,\mu \in \mathbb{R}$ *and* $\underset{\sim}{u},\underset{\sim}{\omega} \in U$.

(iii) $T(-\underset{\sim}{u}) = -T(\underset{\sim}{u})$ *all* $\underset{\sim}{u} \in U$.

(iv) $T(\lambda\mu\underset{\sim}{u}) = \lambda(\mu(T(\underset{\sim}{u})))$ *all* $\lambda,\mu \in \mathbb{R}$ *and* $\underset{\sim}{u} \in U$.

etc., etc.; you may add to this list at your discretion.

<u>Proof:</u> <u>Exercise</u>, see also K., Theorem 4.2, p.153.

<u>KERNEL AND RANGE (IMAGE) OF A LINEAR MAPPING</u>

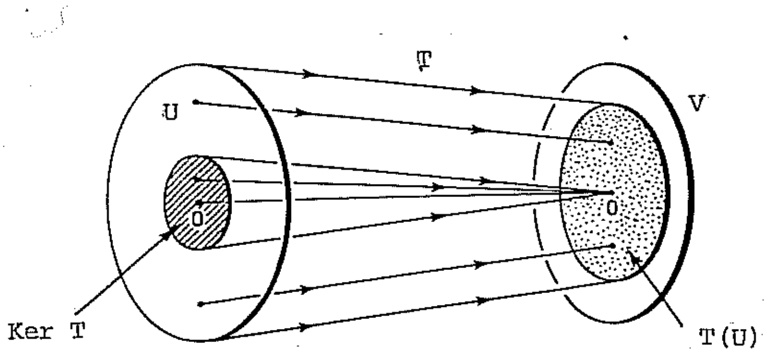
<u>DEFINITION (8):</u> For linear spaces U and V and linear mapping $T \in \mathcal{L}(U,V)$; the *Kernel* of T denoted by <u>Ker T</u>, is the set of vectors in U which T maps to the zero vector of V,

i.e., $\text{Ker } T = \{\underset{\sim}{u} \in U: T(\underset{\sim}{u}) = \underset{\sim}{0}\}$, (K., definition 4.3),

and the *range* of T is the subset of V,

$\underline{T(U)} = \{\underset{\sim}{v} \in V:$ there exists $\underset{\sim}{u} \in U$ with $T(\underset{\sim}{u}) = \underset{\sim}{v}\}$,

(K., definition 4.4).



The concept of range should be familiar to you from your first year work on functions; that of Kernel may be more novel, see K., Examples 3 and 4, pp.156-7 for illustrations of it. (For the moment, ignore the work on dimensions included towards the end of each of these examples.)

<u>THEOREM (9)</u>: *Let* U, V *be linear spaces and* $T \in \mathcal{L}(U,V)$ *then*

Ker T *and* T(U) *are subspaces of* U *and* V *respectively.*

<u>Proof</u>: See K., p.156, part (a) of Theorem 4.3 and p.157, Theorem 4.4.

(10) <u>Recall</u>: A function f from domain X to co-domain Y is

(i) <u>one to one</u> (1-1) if $f(x) = f(y)$ implies $x = y$ (or equivalent $x \neq y$ implies $f(x) \neq f(y)$)

and (ii) <u>onto</u> if its range $f(X) = \{y \in Y: \text{ there exists } x \in X \text{ with } y = f(x)\}$ equals Y.

<u>THEOREM (11)</u>: *Let* U, V *be linear spaces and* $T \in \mathcal{L}(U,V)$, *then:*

(i) T *is onto if and only if* $T(U) = V$;

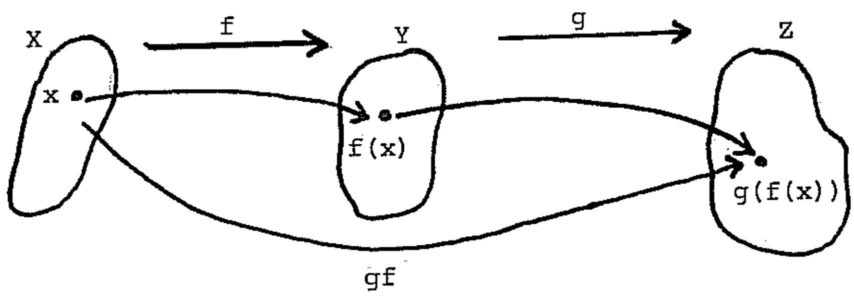
(ii) T *is* 1-1 *if and only if* $\text{Ker } T = \{\underset{\sim}{0}\}$ (<u>i.e.</u>, $T(\underset{\sim}{u}) = 0$ implies $\underset{\sim}{u} = \underset{\sim}{0}$).

<u>Proof</u>: (i) is immediate from the definition of onto;

(ii) see K., p.156, part (b) of Theorem 4.3.

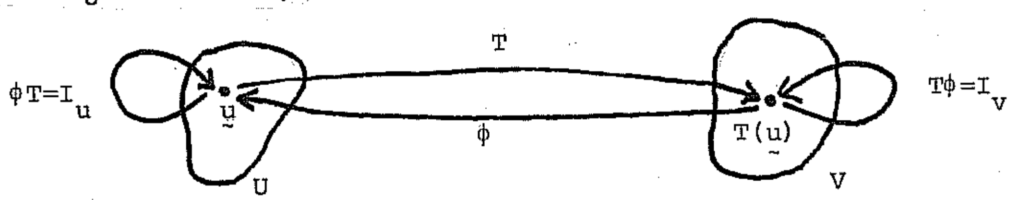
The importance of these concepts lies in the following:

(12) Recall: For sets X, Y, Z and functions $f: X \to Y$, $g: Y \to Z$ the composite gf (or $g \circ f$), defined by $gf(x) = g(f(x))$, is a function from X to Z.



If U, V, W are linear spaces and $T \in \mathcal{L}(U,V)$, $S \in \mathcal{L}(V,W)$, it is readily seen that ST is linear, i.e., $ST \in \mathcal{L}(U,W)$. (Check this.)
Also, for any linear space U let I_U denote the identity mapping defined by $I_U(\underset{\sim}{u}) = \underset{\sim}{u}$ all $\underset{\sim}{u} \in U$. Clearly $I_U \in \mathcal{L}(U,U)$.

DEFINITION (13) (See K., p.161) For linear spaces U and V, $T \in \mathcal{L}(U,V)$ is *invertible* if there exists a mapping $\phi: V \to U$ (not assumed to be linear - see below) such that $\phi T = I_U$ and $T\phi = I_V$.



EXERCISE (14) Show that ϕ in the above definition, if it exists, is unique.
[Hint. Let ψ be another such mapping and consider $\phi T \psi$.]
If $T \in \mathcal{L}(U,V)$ is invertible we will refer to the unique ϕ of the above definition as the inverse of T and denote it by T^{-1}.

THEOREM (15): *Let* U, V *be linear spaces, then* $T \in \mathcal{L}(U,V)$ *is invertible if and only if* T *is* 1-1 *and onto* (i.e., by above, Ker $T = \{0\}$ and $T(U) = V$). *Moreover, its inverse* T^{-1} *is a linear mapping,* i.e., $T^{-1} \in \mathcal{L}(V,U)$, *and* $(T^{-1})^{-1} = T$.

Proof: See K., p.161, Theorem 4.6.

An invertible linear mapping between two linear spaces U, V is sometimes referred to as an isomorphism. If there exists an isomorphism from the linear space U to the linear space V we say U and V are isomorphic. Informally two vector spaces are isomorphic if each is essentially an identical "copy" of the other.

EXAMPLE (16): The space $P_n(I)$ of polynomials on I of degree less than or equal to n (see Example 5, §IV) is isomorphic to V^{n+1}. The mapping

$$T: \quad V^{n+1} \to P_n(I),$$

given by

$$T(a_1, a_2, \ldots, a_{n+1}) = a_1 + a_2x + a_3x^2 + \ldots + a_{n+1}x^n \quad ,$$

is clearly an isomorphism (Prove this).

NOTATION: Henceforth, for a linear mapping T and a vector $\underset{\sim}{u}$ in the domain of T we will write $T\underset{\sim}{u}$ in place of the unnecessarily cumbersome $T(\underset{\sim}{u})$.

EXERCISES (17):

1. Let U, V be linear spaces and $T \in L(U,V)$
   (a) (i) show that $I_vT = TI_u = T$
   (ii) prove Exercise 14 of §V, p.21

   (b) Let U, V, W be linear spaces and let $T \in L(U,V)$ and $S \in L(V,W)$ be invertible mappings. Show that the composite $ST \in L(U,W)$ is invertible with $(ST)^{-1} = T^{-1}S^{-1}$.

2. Let V be the linear space of sequences, show that the right shift operator S_R (§V Example 6) is 1-1 but not onto, while the left shift S_L (§V Example 6) is onto but not 1-1.

   Hence conclude that neither S_R nor S_L are invertible. However $S_LS_R = I_V$, so S_L is a "left inverse" of S_R i.e. S_R is left invertible (similarly S_R is a "right inverse" of S_L i.e. S_L is right invertible) - this shows that the double condition of Definition 13, p.16: $T\phi = I_V$ and $\phi T = I_U$ is necessary for an operator to be invertible

3. Let V be the linear space of all continuous functions from [-1,1] into $\mathbb{R}$ (i.e. $V = C([-1,1])$).

For any fixed point $x_0 \in [-1,1]$ define the evaluation functional E_{x_0} by

$$E_{x_0}(f) = f(x_0)$$

i.e. E_{x_0} maps each function $f \in V$ to its value at x_0.

Show that E_{x_0} is a linear mapping from V into $\mathbb{R}$. Deduce that for scalars $\lambda_1, \lambda_2, \lambda_3, \dots \lambda_n \in \mathbb{R}$ and points $x_1, x_2, \dots, x_n \in [-1,1]$ the mapping defined by

$$T(f) = \lambda_1 E_{x_1}(f) + \lambda_2 E_{x_2}(f) + \dots + \lambda_n E_{x_n}(f) \quad \text{for all } f \in V$$

is also a linear mapping from V into $\mathbb{R}$.

In particular then we have

$$\begin{aligned} S(f) &= \lambda_1 E_{-1}(f) + \lambda_2 E_0(f) + \lambda_3 E_1(f) \\ &= \lambda_1 f(-1) + \lambda_2 f(0) + \lambda_3 f(1) \end{aligned}$$

is a linear mapping from V into $\mathbb{R}$.

The definite integral $I(f) = \int_{-1}^{1} f(x)\, dx$ is also a linear mapping from V to $\mathbb{R}$.

Find values of the constants $\lambda_1, \lambda_2, \lambda_3$ such that S approximates I in the sense that S and I agree for the three functions

$$f(x) = 1, \quad f(x) = x \quad \text{and } f(x) = x^2 .$$

(Can you recognize what you have just established?)

4. Show that each of the following "integral operators" are linear

(a) The Laplace Transform L from the linear space of all real values bounded functions on $[0,\infty)$, defined by

$$L(f)(x) = \int_0^\infty e^{-sx} f(s)\, dx$$

(b) The convolution

$$T(f)(x) = (\sin * f)(x) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin(x-t)\, f(t)\, dt$$

§VI Basic Structure in linear spaces

In this section we develop further the concept of spanning sets introduced in §III Theorem 4. The ideas of this section are among the most important in elementary linear algebra and it is essential that you make every effort to fully understand them.

As foreshadowed in §III, S is a spanning set for the linear space V (ie V = <S>) if every element of V is a linear combination of elements from S, where

DEFINITION 1) by a (finite [1]) *linear combination* of the vectors $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_n$ we mean a vector of the form

$$\underset{\sim}{v} = \sum_{j=1}^{n} \lambda_j \underset{\sim}{v}_j = \lambda_1 \underset{\sim}{v}_1 + \lambda_2 \underset{\sim}{v}_2 + \ldots + \lambda_n \underset{\sim}{v}_n$$

for some set of scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$.

A linear combination is non-trivial if at least one of the scalar coefficients, λ_i, is non-zero. Otherwise, if all the λ_i are zero, it is trivial and necessarily equal to the zero vector.
(Refer K. p.78 definition 2.3, also see Example 1 on the same page.)

EXAMPLE: In the linear space $C(\mathbb{R})$ - Example 3 of §IV - the vector $\sin^3 x$ is a linear combination of sin x and sin 3x. Indeed from elementary trigonometry, $\sin^3 x = \frac{3}{4} \sin x - \frac{1}{4} \sin 3x$.

DEFINITION 2) A linear space V is finite dimensional if it has a finite spanning set, ie.if V = <S> for some finite set (set with only a finite number of elements) S. Otherwise, if every spanning set for V contains an infinity of elements, we say V is infinite dimensional.

EXAMPLES: $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \ldots, \underset{\sim}{e}_n\}$ where $\underset{\sim}{e}_k = (0,0,\ldots,1,\ldots0)$
↑ k'th place

is a finite spanning set for V_n.

$\{f_0, f_1, f_2, \ldots, f_n\}$ where $f_k(x) = x^k$ is a finite spanning set for $P_n(I)$ - see Example 5 §IV p 11. So both V_n and $P_n(I)$ are finite dimensional. Before we can easily give examples of infinite dimensional linear spaces we must develop some further theory.

Finite dimensional linear spaces will dominate our subsequent work, not because infinite dimensional spaces are less important but because of

---

1) Throughout this course all sums will be finite unless otherwise specified.

technical difficulties inherent in any satisfactory theory of them.

DEFINITIONS (3): S is a minimal spanning set for the linear space V if S spans V and no proper subset of S spans V, i.e., $V = \langle S \rangle$ and if $S_0 \subsetneq S$ then $\langle S_0 \rangle \neq V$.

PROPOSITION 4: *Let* V *be a finite dimensional linear space, then any finite spanning set* S *for* V *contains a minimal spanning set for* V.

Proof: Let m be the number of elements in S.
If S is not itself a minimal spanning set, then there is some proper subset S_1 of S (with at most m-1 elements) with $\langle S_1 \rangle = V$.
If S_1 is not a minimal spanning set, then there is some proper subset S_2 of S (with at most m-2 elements) with $\langle S_2 \rangle = V$.

Clearly this process must terminate before m steps and so produce a minimal spanning set. Otherwise we would have $V = \langle \emptyset \rangle = \emptyset$ which is impossible as V contains a zero-vector. □

Clearly, no element of a minimal spanning set S can be a linear combination of other elements of S. [If $\underset{\sim}{s} \in S$ were such a linear combination of elements of $S\backslash\{\underset{\sim}{s}\}$ then we would have $\langle S\backslash\{\underset{\sim}{s}\}\rangle \supseteq S$ and so $V \supseteq \langle S\backslash\{\underset{\sim}{s}\}\rangle = \langle\langle S\backslash\{\underset{\sim}{s}\}\rangle\rangle \supseteq \langle S \rangle = V$, showing $S\backslash\{\underset{\sim}{s}\}$ is a spanning set for V and so contradicting the minimality of S.]

From this simple observation we can derive an extremely useful necessary condition for a subset of S to be a minimal spanning set.

LEMMA 5: *Let* $S = \{v_1, v_2, \ldots, v_n\}$ *be an ordered subset*(1) *of the linear space* V, *then the following are equivalent:*

(i) *There exists some* k, $1 \le k \le n$, *such that* $\underset{\sim}{v}_k$ *is a linear combination of its predecessors* i.e., $\underset{\sim}{v}_k = \sum_{i=1}^{k-1} \lambda_i v_i$ some $\lambda_1, \lambda_2, \ldots, \lambda_{k-1} \in \mathbb{R}$

(ii) *Some element of* S *is a linear combination of other elements of* S

(Continued on the next page)

---

(1) By an *ordered set* $\{v_1, v_2, \ldots, v_n\}$ we really mean an ordered n-tuple of vectors in which no element appears more than once. Thus the order in which the elements are listed is important, $\{\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_n\}$ and $\{\underset{\sim}{v}_2, \underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_n\}$ are not the same ordered set, although they have the same elements. For an ordered set statements such as: the first element, the last element, the next element, the preceding element, make sense.

LEMMA 5(Cont.)

(iii) *There exists a non-trivial linear combination of the elements of* S *equalling the zero vector,* that is

$$\underset{\sim}{0} = \sum_{i=1}^{n} \lambda_i \underset{\sim}{v}_i \quad \text{for some set of scalars } \lambda_1, \lambda_2, \dots \lambda_n$$

not all of which are zero.

<u>DEFINITION 6</u>: A set of vectors satisfying any one (and hence all three) of the conditions listed in Lemma 5 is said to be <u>linearly dependent</u> (one of its elements depends (linearly) on the others). A set of vectors which is not linearly dependent is <u>linearly independent</u>, thus $\{\underset{\sim}{v}_1, \underset{\sim}{v}_2, \dots, \underset{\sim}{v}_n\}$ is linearly independent if $\sum_{i=1}^{n} \lambda_i \underset{\sim}{v}_i = \underset{\sim}{0}$ only when $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$.

<u>Proof</u> (of Lemma 5): That (i) ⇒ (ii) is obvious.

Now, if ii) holds, for some k, $1 \le k \le n$ we have $\underset{\sim}{v}_k = \sum_{\substack{i=1 \\ i \neq k}}^{n} \lambda_i \underset{\sim}{v}_i$ for some set of scalars λ_i some (or all) of which may of course be 0.

But, then

$\underset{\sim}{0} = \lambda_1 \underset{\sim}{v}_1 + \lambda_2 \underset{\sim}{v}_2 + \dots + (-1)\, \underset{\sim}{v}_k + \dots + \lambda_n \underset{\sim}{v}_n$ and the R.H.S. is a non-trivial linear combination of the elements of S as one of the scalar coefficients, namely -1, is non-zero. Thus ii) ⇒ iii). That iii) ⇒ i) is proved in the last paragraph ("Conversely, let S be ...") of the proof to Theorem 2.4 in K. on p 83. □

From the remarks preceeding lemma 5 we have

<u>COROLLARY 7</u>: *A finite minimal spanning set for a linear space is a linearly independent set.*

<u>COROLLARY 8</u>: *Let* $\{\underset{\sim}{v}_1, \underset{\sim}{v}_2, \dots, \underset{\sim}{v}_n\}$ *be an ordered minimal spanning set for the linear space* V, *then for any* $\underset{\sim}{v} \in V$ *there is a unique set of scalars* $\lambda_1, \lambda_2, \dots, \lambda_n$ *such that*

$$\underset{\sim}{v} = \sum_{i=1}^{n} \lambda_i \underset{\sim}{v}_i$$

These unique scalars are known as <u>the coordinates of $\underset{\sim}{v}$ with respect to the minimal spanning set $\{\underset{\sim}{v}_1, \underset{\sim}{v}_2, \dots, \underset{\sim}{v}_n\}$</u>.

(⇐) We first show S_0 is a spanning set for V.
Assume not, ie. there exists $\underset{\sim}{v} \in V \backslash \langle S_0 \rangle$.

Let $S = S_0 \cup \{\underset{\sim}{v}\}$

ie. $S = \{\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_n, \underset{\sim}{v}\}$

then S is a super set of S_0 and so must be linearly dependent, thus, by Lemma 5 i), one of the elements of S must be a linear combination of its predecessors, it cannot be one of the v_k, $k=1,2,\ldots,n$ or again by Lemma 5(i) $S_0 = \{\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_n\}$ would be linearly dependent, so it must be $\underset{\sim}{v}$ ie.

$$\underset{\sim}{v} = \sum_{i=1}^{n} \lambda_i \underset{\sim}{v}_i \quad \text{for some } \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R} .$$

But this means $\underset{\sim}{v} \in \langle S_0 \rangle$ contradicting the choice of $\underset{\sim}{v}$, so no such $\underset{\sim}{v}$ can exist and $V = \langle S_0 \rangle$.
Now, assume S_0 is not minimal as a spanning set ie. $V = \langle S_1 \rangle$ for some $S_1 \subsetneq S_0$, let $\underset{\sim}{v}_k$ be any element of $S_0 \backslash S_1$ then $\underset{\sim}{v}_k$ can be expressed as a linear combination of the elements of S_1. Thus one element of S_0, namely $\underset{\sim}{v}_k$, is a linear combination of other elements of S_0, so by Lemma 5 ii) S_0 is linearly dependent. A contradiction showing that S_0 is a minimal spanning set. □

The next result is one of the most fundamental to the study of finite dimensional spaces and possibly the most tedious to prove.

THEOREM 10: *Let* $S = \{\underset{\sim}{s}_1, \ldots, \underset{\sim}{s}_n\}$ *be a finite basis for the linear space* V, *then any other basis for* V *also has* n *elements.*

Proof: Let T be any minimal spanning set for V.
It suffices to show that T has m elements where $m \leq n$, for then the same argument applied with the roles of S and T interchanged would show $n \geq m$. (See K. Corollary 2.1, p.89. The preceding Theorem 2.7, p.88, also provides an alternative proof to the one we give here.)
Assume the contrary, i.e., $m > n$, then certainly T contains n+1 linearly independent vectors $\underset{\sim}{t}_1, \underset{\sim}{t}_2, \ldots, \underset{\sim}{t}_{n+1}$.
Since S is a spanning set for V each of the vectors $\underset{\sim}{t}_k$ $k = 1,2,\ldots,n+1$

be written as linear combinations of the elements of S, <u>i.e.</u>

$$\underset{\sim}{t}_1 = \mu_{11}\underset{\sim}{s}_1 + \mu_{12}\underset{\sim}{s}_2 + \dots + \mu_{1n}\underset{\sim}{s}_n = \sum_{i=1}^{n} \mu_{1i}\underset{\sim}{s}_i$$

$$\underset{\sim}{t}_2 = \mu_{21}\underset{\sim}{s}_1 + \mu_{22}\underset{\sim}{s}_2 + \dots + \mu_{2n}\underset{\sim}{s}_n = \sum_{i=1}^{n} \mu_{2i}\underset{\sim}{s}_i$$

....

$$\underset{\sim}{t}_k = \mu_{k1}\underset{\sim}{s}_1 + \mu_{k2}\underset{\sim}{s}_2 + \dots + \mu_{kn}\underset{\sim}{s}_n = \sum_{i=1}^{n} \mu_{ki}\underset{\sim}{s}_i$$

....

$$\underset{\sim}{t}_{n+1} = \mu_{n+1\,1}\underset{\sim}{s}_1 + \mu_{n+1\,2}\underset{\sim}{s}_2 + \dots + \mu_{n+1\,n}\underset{\sim}{s}_n = \sum_{i=1}^{n} \mu_{n+1\,i}\underset{\sim}{s}_i$$

By substituting these identities, any linear combination of the $\underset{\sim}{t}_k$'s, $\sum_{k=1}^{n+1} \lambda_k\underset{\sim}{t}_k$ can be written as a linear combination of the $\underset{\sim}{s}_i$'s:

$$\sum_{k=1}^{n+1} \lambda_k\underset{\sim}{t}_k = \sum_{k=1}^{n+1} \lambda_k\left(\sum_{i=1}^{n} \mu_{ki}\underset{\sim}{s}_i\right)$$

$$= \lambda_1(\mu_{11}\underset{\sim}{s}_1 + \mu_{12}\underset{\sim}{s}_2 + \dots) + \lambda_2(\mu_{21}\underset{\sim}{s}_1 + \mu_{22}\underset{\sim}{s}_2 + \dots) + \dots$$

$$= (\mu_{11}\lambda_1 + \mu_{21}\lambda_2 + \dots)\,\underset{\sim}{s}_1 + (\mu_{12}\lambda_1 + \mu_{22}\lambda_2 + \dots)\,\underset{\sim}{s}_2 + \dots$$

$$= \sum_{i=1}^{n}\left(\sum_{k=1}^{n+1} \mu_{ki}\lambda_k\right)\underset{\sim}{s}_i$$

Such a linear combination will equal the zero vector provided each of the coefficients of the $\underset{\sim}{s}_i$ is zero,

Thus $\sum_{k=1}^{n+1} \lambda_k\underset{\sim}{t}_k = \underset{\sim}{0}$ for any set of scalars $\lambda_1, \lambda_2, \dots, \lambda_{n+1}$ which satisfy the equations

$$\sum_{k=1}^{n+1} \mu_{ki}\lambda_k = 0 \qquad i = 1,2,\dots,n.$$

<u>i.e.</u>

$$\mu_{11}\lambda_1 + \mu_{21}\lambda_2 + \dots + \mu_{n+1\,1}\,\lambda_{n+1} = 0$$

$$\mu_{12}\lambda_1 + \mu_{22}\lambda_2 + \dots + \mu_{n+1\,2}\,\lambda_{n+1} = 0$$

....

$$\mu_{1n}\lambda_1 + \mu_{2n}\lambda_2 + \dots + \mu_{n+1\,n}\,\lambda_{n+1} = 0$$

Now, this homogeneous system involves more unknowns (n+1 of them; $\lambda_1, \lambda_2, \ldots, \lambda_{n+1}$) than equations (n of them) and so always has a non-zero solution. [This is a well known result which follows readily from your first year work on equations - a proof is also given in the appendix attached to this section.]

Any such solution provides a set of λ_k's, not all zero, such that $\sum_{k=1}^{n+1} \lambda_k \underset{\sim}{t}_k = 0$ and so shows that the n+1 vectors $\underset{\sim}{t}_1, \ldots, \underset{\sim}{t}_{n+1}$ are linearly dependent, a contradiction which establishes $m \leqslant n$ as required. □

As a result of this Theorem we can now offer

DEFINITION (11) Let V be a finite dimensional linear space, then V has dimension n (written dim V = n) where n is the number of vectors in any one (and hence every) basis for V.

EXAMPLES: $\dim V_n = n$, $\dim P_n(I) = n + 1$.

(See K., pp.89-90, Definition 2.7 and the subsequent examples.)

We have seen that in finite dimensional spaces a basis can be viewed as either a minimal spanning set or as a maximal linearly independent set. We also saw that any spanning set contains a basis.

We now give the complementary result for linearly independent sets.

PROPOSITION (12) Let S be a linearly independent set in a finite dimensional linear space V, then S can be extended to a basis for V, i.e. there exists a basis B of V with $S \subseteq B$.

Proof. See K., Theorem 2.8, p.90.

We are now in a position to establish examples of infinite dimensional linear spaces. To do this we observe that from the above results we need only demonstrate the existence of "infinite" linearly independent sets (why?), i.e. sets containing an infinite number of vectors, no one of which can be written as a finite linear combination of the others.

EXAMPLE (12): P(I) *and hence any linear space containing* P(I) *as a subspace, is infinite dimensional* as $\{1, x, x^2, x^3, \ldots, x^n, \ldots\}$ is an infinite linearly independent set (Prove this).

The linear space of sequences convergent to zero c_0, (and hence any linear space containing it as a subspace) *is infinite dimensional.*

For any n the sequence

$$\underset{\sim}{e}_n = 0,0,\ldots,0,\underset{\uparrow \text{ nth place}}{1},0,0,\ldots.$$

clearly converges to zero, and

$$\{\underset{\sim}{e}_1,\underset{\sim}{e}_2,\ldots,\underset{\sim}{e}_n,\ldots\}$$

is an infinite linearly independent set in C_0.

<u>Theorem (13)</u>. *Let* U, W *be finite dimensional subspaces of a linear space* V, *then*

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U\cap W).$$

In particular then, for a direct sum (see page) we have

$$\dim(U \oplus W) = \dim(U) + \dim(W).$$

<u>Proof</u>. Let $\dim U = n_u$

$\dim W = n_w$

$\dim U\cap W = n_\cap$

and let $S = \{\underset{\sim}{s}_1,\underset{\sim}{s}_2,\ldots,\underset{\sim}{s}_{n_\cap}\}$ be a basis for $U \cap W$.

By Proposition (12) we can extend S to a basis for U, say

$$S_U = \{\underset{\sim}{s}_1,\underset{\sim}{s}_2,\ldots,\underset{\sim}{s}_{n_\cap},\ \underset{\sim}{u}_1,\underset{\sim}{u}_2,\ldots,\underset{\sim}{u}_{n_u-n_\cap}\},$$

by the addition of $n_u - n_\cap$ vectors.
Similarly S may be extended to a basis for W;

$$S_W = \{\underset{\sim}{s}_1,\underset{\sim}{s}_2,\ldots,\underset{\sim}{s}_{n_\cap},\ \underset{\sim}{w}_1,\underset{\sim}{w}_2,\ldots,\underset{\sim}{w}_{n_w-n_\cap}\}$$

Let

$$S_1 = S_U \cup S_W = \{\underset{\sim}{s}_1,\ldots,\underset{\sim}{s}_{n_\cap},\ u_1,\ldots,\underset{\sim}{u}_{n_u-n_\cap},\ \underset{\sim}{w}_1,\ldots,\underset{\sim}{w}_{n_w-n_u}\}.$$

It suffices to show S_1 is linearly independent, for clearly any vector in $U \cup W$ can be written as a linear combination of element of S_1 hence S_1 spans $\langle U \cup W\rangle = U + W$ so S_1 is a basis for $U + W$ and

$$\dim(U + W) = \text{no. of elements in } S_1$$
$$= n_\cap + (n_u - n_\cap) + (n_w - n_\cap)$$
$$= n_u + n_w - n_\cap$$
$$= \dim U + \dim W - \dim(U \cap W)$$

the desired conclusion.

Let $n_u - n_\cap = m$, $n_w - n_\cap = k$ and assume

$$\lambda_1 \underset{\sim}{s}_1 + \ldots + \lambda_{n_\cap} \underset{\sim}{s}_{n_\cap} + \mu_1 \underset{\sim}{u}_1 + \ldots + \mu_m \underset{\sim}{u}_m + \omega_1 \underset{\sim}{w}_1 + \ldots + \omega_k \underset{\sim}{w}_k = \underset{\sim}{0}$$

then

$$\lambda_1 \underset{\sim}{s}_1 + \ldots + \lambda_{n_\cap} \underset{\sim}{s}_{n_\cap} + \mu_1 \underset{\sim}{u}_1 + \ldots + \mu_m \underset{\sim}{u}_m = -\omega_1 \underset{\sim}{w}_1 - \ldots - \omega_k \underset{\sim}{w}_k .$$

R.H.S. is an element in W while L.H.S. is an element in U so both are in U ∩ W.

Thus both are equal to a linear combination of elements in S, i.e.

$$-\omega_1 \underset{\sim}{w}_1 - \ldots - \omega_k \underset{\sim}{w}_k = \theta_1 \underset{\sim}{s}_1 + \ldots + \theta_{n_\cap} \underset{\sim}{s}_{n_\cap}$$

and

$$\lambda_1 \underset{\sim}{s}_1 + \ldots + \lambda_{n_\cap} \underset{\sim}{s}_{n_\cap} + \mu_1 \underset{\sim}{u}_1 + \ldots + \mu_m \underset{\sim}{u}_m = \theta_1 \underset{\sim}{s}_1 + \ldots + \theta_{n_\cap} \underset{\sim}{s}_{n_\cap}$$

or

$$\theta_1 \underset{\sim}{s}_1 + \ldots + \theta_{n_\cap} \underset{\sim}{s}_{n_\cap} + \omega_1 w_1 + \ldots + \omega_k \underset{\sim}{w}_k = \underset{\sim}{0}$$

and so, since S_W is linearly independent

$$\theta_1 = \ldots = \theta_{n_\cap} = \omega_1 = \ldots = \omega_k = 0.$$

But then

$$\lambda_1 \underset{\sim}{e}_1 + \ldots + \lambda_{n_\cap} \underset{\sim}{s}_{n_\cap} + \mu_1 \underset{\sim}{u}_1 + \ldots + \mu_m \underset{\sim}{u}_m = \underset{\sim}{0} \qquad (\text{all } \theta_i = 0)$$

so, since S_U is linearly independent

$$\lambda_1 = \ldots = \lambda_{n_\cap} = \mu_1 = \ldots = \mu_m = 0,$$

showing S_1 is linearly independent as required. □

EXERCISES (14): (The results of these exercises will be assumed in subsequent work).

1) Prove that any set of vectors containing the zero vector is linearly independent.

2) Show that if W is a nonzero subspace of a finite-dimensional vector space V, then dim W $\leq$ dim V.

3) Show that if W is a subspace of a finite-dimensional vector space V and dim W = dim V, then W = V.

4) Let V be an n-dimensional vector space.

   (a) If S = $\{\underset{\sim}{a}_1, \underset{\sim}{a}_2, \ldots, \underset{\sim}{a}_n\}$ is a linearly independent set of vectors in V, then S is a basis for V.

   (b) If S = $\{\underset{\sim}{a}_1, \underset{\sim}{a}_2, \ldots, \underset{\sim}{a}_n\}$ spans V, then S is a basis for V.

5) Let V be an n-dimensional vector space. Show that any n+1 vectors in V form a linearly dependent set. Hence any linearly independent set of vectors in V has at most n elements.

6) Show that a set of vectors is linearly dependent if any subset of it is a linearly dependent set.

7) A subspace U of the linear space V is complemented if there exists a subspace W of V such that V = U $\oplus$ W.
   Show that any subspace of a finite-dimensional space is complemented.
   [Hint: Let $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_k\}$ be a basis for U and extend it to a basis for V, $\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_k, \underset{\sim}{u}_{k+1}, \ldots, \underset{\sim}{u}_n\}$.]

APPENDIX TO §VI

Linear Homogeneous Systems involving more unknowns than equations

Here we give a proof of the well known result used in the proof of Theorem 9. For the system of simultaneous equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m &= 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m &= 0 \\ \dots\dots\dots\dots\dots\dots\dots & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m &= 0 \end{aligned}$$

where $m > n$ there exists at least one solution set $x_1 = \alpha_1$, $x_2 = \alpha_2$, ..., $x_m = \alpha_m$ such that not all the α_i are zero.

The proof is taken from Warren Brisley "A Basis for Linear Algebra" and is reproduced for the private use of students in PM211 only.

Proof. By induction on the positive integer m. $\mathscr{P}(m)$ is the whole statement of the Theorem. $\mathscr{P}(1)$ is void, $\mathscr{P}(2)$ is the assertion that "given an equation $a_{11}x_1 + a_{12}x_2 = 0$, there is a non-trivial (i.e. not-all-zero) solution".

If a_{11} and a_{12} are both zero, put $x_1 = 1$, $x_2 = 1$ as solution.
If $a_{11} \neq 0$, put $x_2 = 1$ and $x_1 = -a_{12}/a_{11}$.
If $a_{11} = 0$, but $a_{12} \neq 0$, put $x_1 = 1$ and $x_2 = 0$.
So $\mathscr{P}(2)$ *is* true.

For the second leg of induction, operate under the assumption that $\mathscr{P}(k)$ is true for some $k \geqslant 2$. Examine the system involved in $\mathscr{P}(k+1)$

$$\begin{bmatrix} a_{11} & \dots & a_{1,k+1} \\ \vdots & & \\ a_{21} & \dots & a_{2,k+1} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{r,k+1} \end{bmatrix} \begin{bmatrix} x_1 \\ \\ x_2 \\ \vdots \\ \vdots \\ x_{k+1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix} \text{ with } r < k+1.$$

If *all* the a_{ij} here are zero, put $x_1 = x_2 = \dots = x_{k+1} = 1$ as a solution.

If *not all* the a_{ij} are zero, we can rearrange the equations and rename the a_{ij}, so that, without loss of generality, we have $a_{r,k+1} \neq 0$. Say, $a_{r,k+1} = \beta \neq 0$. Then the last equation gives

(i) $$\frac{a_{r1}}{\beta}x_1 + \frac{a_{r2}}{\beta}x_2 + \dots + \frac{a_{rk}}{\beta}x_k + x_{k+1} = 0.$$

If we now subtract "$a_{i,k+1}$ times this equation" from each (other) ith equation ($i = 1, 2, \dots, r-1$), we have

(ii) $$\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ & & & \\ b_{21} & b_{22} & \dots & b_{2k} \\ \vdots & & & \\ b_{r-1,1} & b_{r-1,2} & \dots & b_{r-1,k} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ \vdots \\ x_k \end{bmatrix} \quad \left(b_{ij} = a_{ij} - \frac{a_{i,k+1} \cdot a_{rj}}{\beta} \right)$$

and (ii) together with (i) is equivalent to the original set of equations.

Now $\mathscr{P}(k)$ is true by assumption, so there is a solution-set

$$\{x_1 = \alpha_1, x_2 = \alpha_2, \dots, x_k = \alpha_k\}$$

(with *not all* the α_i zero) for (ii). Use (i) to calculate

$$\alpha_{k+1} = -\frac{a_{r1}\alpha_1 + \dots + a_{rk}\alpha_k}{\beta}.$$

Now $\{x_1 = \alpha_1, x_2 = \alpha_2, \dots, x_{k+1} = \alpha_{k+1}\}$ is a non-trivial solution for the system. Thus we have

(i) $\mathscr{P}(2)$ true,
(ii) whenever $\mathscr{P}(k)$ is true, ($k \geqslant 2$), then $\mathscr{P}(k+1)$ is true. □