Solutions

BLOCK 3:
SELECTED TOPICS IN LINEAR ALGEBRA

Pretest

1. 2

2. $x_4 = 19x_1 + 9x_2 - 8x_3$

3. All scalar multiples of $3u_1 - 2u_2 + u_3$

4. -40

5. (a) c = 2 and c = 3

   (b) $V_2 = [\alpha_1] \oplus [\alpha_2]$ where $\alpha_1 = u_2 + u_3$ and $\alpha_2 = 3u_1 + 2u_3$

6. $B = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix}$.

7. (a) $\frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\sin(2n + 1)x}{2n + 1}$

   (b) $\frac{\pi}{4}$

Unit 1: The Case Against n-Tuples

3.1.1(L)

In our discussion of this exercise, we shall restrict our study to 2-dimensional spaces, but the discussion is easily generalized to higher dimensional spaces, as we shall show in the later exercises.

Let us imagine that we are viewing E^2 as the space of vectors in the xy-plane, and that, as usual, we think of all vectors as being linear combinations of $\vec{i}$ and $\vec{j}$. In this context, it is conventional to view the 2-tuple (x,y) as an abbreviation for $x\vec{i} + y\vec{j}$.

In particular, then, if

$$\vec{\gamma} = 5\vec{i} + 4\vec{j} \qquad (1)$$

we would abbreviate (1) by writing

$$\vec{\gamma} = (5,4). \qquad (2)$$

Now suppose we consider a second pair of vectors in the xy-plane, say,

$$\vec{\alpha} = 3\vec{i} + 4\vec{j} \qquad (3)$$

and

$$\vec{\beta} = 2\vec{i} + 3\vec{j}. \qquad (4)$$

With respect to our given convention, (3) and (4) may be rewritten as

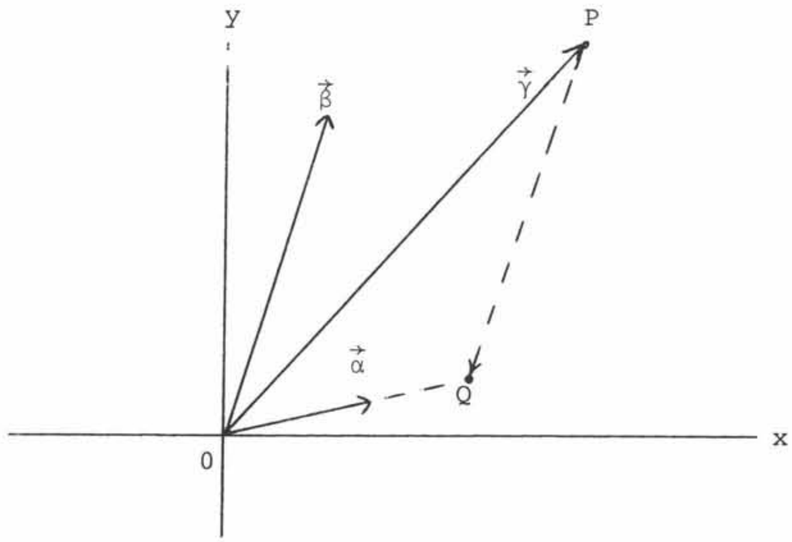
$$\vec{\alpha} = (3,4) \qquad (3')$$

and

$$\vec{\beta} = (2,3). \qquad (4')$$

3.1.1(L) continued

We notice, at least pictorially, that each vector in the xy-plane has a unique representation as a linear combination of $\vec{\alpha}$ and $\vec{\beta}$. Namely, quite in general, if $\vec{\alpha}$ is not a scalar multiple of $\vec{\beta}$ we have



$$\begin{aligned}\vec{\gamma} &= \vec{OP}\\ &= \vec{OQ} + \vec{QP}\\ &= a\vec{\alpha} + b\vec{\beta}\end{aligned}$$

Since vector arithmetic and numerical arithmetic have the same structure, we may "invert" (3) and (4) to express $\vec{i}$ and $\vec{j}$ in terms of $\vec{\alpha}$ and $\vec{\beta}$. Namely:

$$\left.\begin{aligned}-2\vec{\alpha} &= -6\vec{i} - 8\vec{j}\\ 3\vec{\beta} &= 6\vec{i} + 9\vec{j}\end{aligned}\right\} .$$

Hence

$$\vec{j} = -2\vec{\alpha} + 3\vec{\beta} \qquad (5)$$

and similarly

$$\left.\begin{aligned}3\vec{\alpha} &= 9\vec{i} + 12\vec{j}\\ -4\vec{\beta} &= -8\vec{i} - 12\vec{j}\end{aligned}\right\} .$$

3.1.1(L) continued

Hence,

$$\vec{i} = 3\vec{\alpha} - 4\vec{\beta} . \tag{6}$$

Using (5) and (6), we see that (1) may be rewritten as:

$$\begin{aligned}\vec{\gamma} &= 5(3\vec{\alpha} - 4\vec{\beta}) + 4(-2\vec{\alpha} + 3\vec{\beta}) \\ &= 7\vec{\alpha} - 8\vec{\beta} .\end{aligned} \tag{7}$$

If we now elect to view $\vec{\gamma}$ with respect to $\vec{\alpha}$ and $\vec{\beta}$ coordinates by letting (a,b) now denote $a\vec{\alpha} + b\vec{\beta}$, equation (7) becomes

$$\vec{\gamma} = (7,-8). \tag{8}$$

Now $\vec{\gamma}$ is the same vector whether we view it in terms of $\vec{i}$ and $\vec{j}$ components or in terms of $\vec{\alpha}$ and $\vec{\beta}$ components.

Comparing (2) and (8), it appears that (5,4) = (7,-8), but this is no contradiction since (5,4) is the representation of $\vec{\gamma}$ with respect to i and j, while (7,-8) is the representation of $\vec{\gamma}$ with respect to $\vec{\alpha}$ and $\vec{\beta}$.

The definition which says that $(a_1,a_2) = (b_1,b_2) \leftrightarrow a_1 = b_1$ and $a_2 = b_2$ presupposes that (a_1,a_2) and (b_1,b_2) are representations with respect to the same pair of vectors.

Our main point is that the notation (a,b) is ambiguous since it means $a\alpha_1 + b\alpha_2$, and this in turn depends on α_1 and α_2. For example, with respect to $\vec{i}$ and $\vec{j}$, (5,4) denotes $5\vec{i} + 4\vec{j}$; but with respect to $\vec{\alpha}$ and $\vec{\beta}$, it denotes

$$\begin{aligned}5\vec{\alpha} + 4\vec{\beta} &= 5(3\vec{i} + 4\vec{j}) + 4(2\vec{i} + 3\vec{j}) \\ &= 23\vec{i} + 32\,\vec{j}.\end{aligned}$$

Note #1:

Our main aim in this exercise was to back up our assertion that there is some ambiguity involved when we write vectors of E^n in n-tuple notation. Nevertheless, it would be a shame to throw away the nice structure of n-tuple arithmetic just for this one reason.

3.1.1(L) continued

So what we do is make some sort of compromise, or convention. Namely, given E^n we assume that we have a specific set of n vectors, $\{\vec{u}_1,\ldots,\vec{u}_n\}$, such that each vector of E^n can be expressed uniquely (i.e., in one and only one way) as a linear combination of $\vec{u}_1,\ldots,$ and $\vec{u}_n$. With this explicit assumption, we then abbreviate each vector in E^n by $(a_1,\ldots,a_n)$, where $(a_1,\ldots,a_n)$ means $a_1\vec{u}_1 + \ldots + a_n\vec{u}_n$. Since no two different linear combinations of $\vec{u}_1,\ldots,$ and $\vec{u}_n$ can yield the same vector of E^n, we see that our n-tuple abbreviation obeys the usual rules for n-tuple arithmetic.

The question of how we find a set of vectors such as the above-described $\{\vec{u}_1,\ldots,\vec{u}_n\}$ will be discussed within the next few Unit of this Block.

The question of how we translate $(a_1,\ldots,a_n)$ into $(b_1,\ldots,b_n)$, where $(b_1,\ldots,b_n)$ means $b_1\vec{v}_1 + \ldots b_n\vec{v}_n$ and $\{\vec{v}_1,\ldots,\vec{v}_n\}$ is another set of vectors such that any element of E^n can be expressed uniquely as a linear combination of $v_1,\ldots,$ and v_n will be discussed in Exercise 3.1.3.

Note #2

Because of the similarities between vector and numerical arithmetic, we may invert (3) and (4) by the row-reduced matrix technique. Namely,

$$\begin{array}{cccc} \vec{i} & \vec{j} & \vec{\alpha} & \vec{\beta} \end{array}$$

$$\begin{bmatrix} 3 & 4 & 1 & 0 \\ 2 & 3 & 0 & 1 \end{bmatrix} \sim$$

$$\begin{bmatrix} 6 & 8 & 2 & 0 \\ 6 & 9 & 0 & 3 \end{bmatrix} \sim$$

$$\begin{bmatrix} 6 & 8 & 2 & 0 \\ 0 & 1 & -2 & 3 \end{bmatrix} \sim$$

$$\begin{bmatrix} 6 & 0 & 18 & -24 \\ 0 & 1 & -2 & 3 \end{bmatrix} \sim$$

3.1.1(L) continued

$$\begin{array}{cccc} \vec{i} & \vec{j} & \vec{\alpha} & \vec{\beta} \end{array}$$
$$\begin{bmatrix} 1 & 0 & 3 & -4 \\ 0 & 1 & -2 & 3 \end{bmatrix} \quad (9)$$

From (9) we see that $\vec{i} = 3\vec{\alpha} - 4\vec{\beta}$ and $\vec{j} = -2\vec{\alpha} + 3\vec{\beta}$ which agrees with (5) and (6).

This connection between representing a vector in different coordinate systems can be extended to show how matrix multiplication allows us to transform a vector from an n-tuple with respect to one coordinate system to an n-tuple in another coordinate system. The technique is very similar to the one we described in the course of our discussion in Block 4 of Part 2 concerning the inversion of a system of n linear algebraic equations in n unknowns. We shall follow up this connection also in Exercise 3.1.3.

3.1.2(L)

a. In the usual mechanical way, we obtain

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ 3 & 5 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \quad (1)$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 2 & 5 & -3 & 0 & 1 \end{bmatrix} \sim$$

$$\begin{bmatrix} 1 & 0 & -1 & 3 & -1 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{bmatrix} \sim$$

$$\begin{bmatrix} 1 & 0 & 0 & 4 & -3 & 1 \\ 0 & 1 & 0 & -4 & 5 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{bmatrix} \sim \quad (2)$$

b. If our coding system in (a) is that the first three columns of our matrix denote $\vec{i}$, $\vec{j}$, and $\vec{k}$ while the last three columns denote $\vec{\alpha}$, $\vec{\beta}$, and $\vec{\gamma}$ then (1) becomes

3.1.2(L) continued

$$\begin{array}{cccccc} \vec{i} & \vec{j} & \vec{k} & \vec{\alpha} & \vec{\beta} & \vec{\gamma} \\ \hline 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ 3 & 5 & 8 & 0 & 0 & 1 \end{array}$$

and this says that

$$\left.\begin{aligned} \vec{\alpha} &= \vec{i} + \vec{j} + \vec{k} \\ \vec{\beta} &= 2\vec{i} + 3\vec{j} + 4\vec{k} \\ \vec{\gamma} &= 3\vec{i} + 5\vec{j} + 8\vec{k} \end{aligned}\right\} \qquad (3)$$

Matrix (2) then codes the inverse of (3). Namely

$$\left.\begin{aligned} \vec{i} &= 4\vec{\alpha} - 3\vec{\beta} + \vec{\gamma} \\ \vec{j} &= -4\vec{\alpha} + 5\vec{\beta} - 2\vec{\gamma} \\ \vec{k} &= \vec{\alpha} - 2\vec{\beta} + \vec{\gamma} \end{aligned}\right\} \qquad (4)$$

c. If

$$\vec{\xi} = 5\vec{i} + 3\vec{j} - 2\vec{k}, \qquad (5)$$

then we obtain from (4) that

$$\begin{aligned} \xi &= 5(4\vec{\alpha} - 3\vec{\beta} + \vec{\gamma}) + 3(-4\vec{\alpha} + 5\vec{\beta} - 2\vec{\gamma}) - 2(\vec{\alpha} - 2\vec{\beta} + \vec{\gamma}) \\ &= 6\vec{\alpha} + 4\vec{\beta} - 3\vec{\gamma}\,. \end{aligned} \qquad (6)$$

Comparing (5) and (6) we see that $\vec{\xi} = (5,3,-2)$ is relative to the coordinate system $\{\vec{i},\vec{j},\vec{k}\}$ while $\vec{\xi} = (6,4,-3)$ is relative to the coordinate system $\{\vec{\alpha}, \vec{\beta}, \vec{\gamma}\}$.

d. Every polynomial of degree ≤ 2 can be written uniquely in the form $a_o + a_1x + a_2x^2$. We may now abbreviate $a_o + a_1x + a_2x^2$ by (a_o, a_1, a_2). In this way $1 = (1,0,0)$, $x = (0,1,0)$, and $x^2 = (0,0,1)$. Our first observation is that we need no longer think of arrows and $\vec{i}$, $\vec{j}$, and $\vec{k}$ to interpret $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$.

Now, suppose we let

3.1.2(L) continued

$$\left.\begin{aligned} p_1(x) &= 1 + x + x^2 \\ p_2(x) &= 2 + 3x + 4x^2 \\ p_3(x) &= 3 + 5x + 8x^2 \end{aligned}\right\} \qquad (7)$$

Then, our matrix (1) may be viewed as the code:

$$\begin{array}{cccccc} 1 & x & x^2 & p_1(x) & p_2(x) & p_3(x) \\ \hline 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ 3 & 5 & 8 & 0 & 0 & 1 \end{array}$$

whereupon (2) means

$$\begin{array}{cccccc} 1 & x & x^2 & p_1(x) & p_2(x) & p_3(x) \\ \hline 1 & 0 & 0 & 4 & -3 & 1 \\ 0 & 1 & 0 & -4 & 5 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}$$

This, in turn, tells us that

$$\left.\begin{aligned} 1 &= 4p_1(x) - 3p_2(x) + p_3(x) \\ x &= -4p_1(x) + 5p_2(x) - 2p_3(x) \\ x^2 &= p_1(x) - 2p_2(x) + p_3(x) \end{aligned}\right\}$$

or, by (7),

$$\left.\begin{aligned} 1 &= 4(1 + x + x^2) - 3(2 + 3x + 4x^2) + (3 + 5x + 8x^2) \\ x &= -4(1 + x + x^2) + 5(2 + 3x + 4x^2) - 2(3 + 5x + 8x^2) \\ x^2 &= (1 + x + x^2) - 2(2 + 3x + 4x^2) + (3 + 5x + 8x^2) \end{aligned}\right\} \qquad (8)$$

e. Using (8)

$$\begin{aligned} 5 + 3x - 2x^2 &= 20(1 + x + x^2) - 15(2 + 3x + 4x^2) + 5(3 + 5x - 8x^2) \\ &\quad - 12(1 + x + x^2) + 15(2 + 3x + 4x^2) - 6(3 + 5x + 8x^2) \\ &\quad \underline{- 2(1 + x + x^2) + 4(2 + 3x + 4x^2) - 2(3 + 5x + 8x^2)} \\ &= 6(1 + x + x^2) + 4(2 + 3x + 4x^2) - 3(3 + 5x + 8x^2) \end{aligned}$$

3.1.2(L) continued

That is,

$5 + 3x - 2x^2 = (5,3,-2)$, if (a,b,c) means $a + bx + cx^2$

but

$5 + 3x - 2x^2 = (6,4,-3)$, if (a,b,c) means $ap_1(x) + bp_2(x) + cp_3(x)$.

Our main observation in parts (d) and (e) is that they exhibit an isomorphism between the "arrows" of xyz-space and the polynomials of degree no greater than 2. That is, with respect to addition and scalar multiplication we cannot distinguish between these two different models. In other words, both are models of a 3-dimensional vector space; where in one model the "coordinate" vectors are $1 = (1,0,0)$, $x = (0,1,0)$, and $x^2 = (0,0,1)$.

Thus, it seems that both models used in this exercise are special cases of the following more general problem.

Suppose that the vectors in E^3 are each linear combinations of the vectors α_1, α_2, and α_3; and that they are also linear combinations of the vectors β_1, β_2, and β_3. Given a vector which is expressed as a linear combination of α_1, α_2, and α_3, how do we express it as an equivalent linear combination of β_1, β_2, and β_3? We pursue this in the next exercise.

3.1.3(L)

We have that every vector in E^3 is a unique linear combination of α_1, α_2, and α_3.* Hence, we may use (a,b,c) as an abbreviation for $a\alpha_1 + b\alpha_2 + c\alpha_3$.

*We no longer write $\vec{\alpha}_1$, $\vec{\alpha}_2$, $\vec{\alpha}_3$ since there is no need to restrict E^3 to the "arrow" model. For example, $\vec{\alpha}$, could be a polynomial of degree ≤ 2.

3.1.3(L) continued

We are now told that

$$\left.\begin{aligned} \beta_1 &= \alpha_1 + \alpha_2 + \alpha_3 \\ \beta_2 &= 2\alpha_1 + 3\alpha_2 + 4\alpha_3 \\ \beta_3 &= 3\alpha_1 + 5\alpha_2 + 8\alpha_3 \end{aligned}\right\} \quad (1)$$

If we now let (x,y,z) denote $x\beta_1 + y\beta_2 + z\beta_3$, we have from (1) that $(x,y,z) = x\beta_1 + y\beta_2 + z\beta_3$

$$\begin{aligned} &= x(\alpha_1 + \alpha_2 + \alpha_3) + y(2\alpha_1 + 3\alpha_2 + 4\alpha_3) \\ &\qquad + z(3\alpha_1 + 5\alpha_2 + 8\alpha_3) \\ &= (x + 2y + 3z)\alpha_1 + (x + 3y + 5z)\alpha_2 \\ &\qquad + (x + 4y + 8z)\alpha_3. \end{aligned} \quad (2)$$

In other words, we see from (2) that if

$\gamma = x\beta_1 + y\beta_2 + z\beta_3$,

then

$\gamma = (x,y,z)$, relative to $\{\beta_1, \beta_2, \beta_3\}$

while

$$\gamma = (x + 2y + 3z,\ x + 3y + 5z,\ x + 4y + 8z) \text{ relative to } \{\alpha_1, \alpha_2, \alpha_3\} \quad (3)$$

Equation (3) may be viewed as the product of two matrices. Namely,

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 4 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + 2y + 3z \\ x + 3y + 5z \\ x + 4y + 8z \end{bmatrix}, \quad (4)$$

3.1.3(L) continued

where the right side of (4) is the 3-tuple (3), written as a column matrix (vector).

The major "hang-up" in (4) is that the matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 4 & 8 \end{bmatrix}$$

is not the matrix of coefficients in (1). Rather it is the transpose of that matrix. (Recall that the transpose A^T of the m by n matrix A is the n by m matrix which is obtained from A by interchanging its rows and columns).

If we want our matrix product to reflect the fact that we are using the matrix of coefficients in (1), then instead of using (4), we express (3) in the form

$$[x + 2y + 3z \quad x + 3y + 5z \quad x + 4y + 8z] = [x \ y \ z] \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 3 & 5 & 8 \end{bmatrix} \quad (5)$$

The relationship between (4) and (5) can be summarized and somewhat generalized as follows. Suppose that every vector in E^3* may be expressed uniquely as a linear combination of β_1, β_2, and β_3; and that β_1, β_2, and β_3 are themselves linear combinations of α_1, α_2, and $\alpha_3 \varepsilon E^3$. Suppose, in particular, that

$$\left.\begin{aligned} \beta_1 &= a_{11}\alpha_1 + a_{12}\alpha_2 + a_{13}\alpha_3 \\ \beta_2 &= a_{21}\alpha_1 + a_{22}\alpha_2 + a_{23}\alpha_3 \\ \beta_3 &= a_{31}\alpha_1 + a_{32}\alpha_2 + a_{33}\alpha_3 \end{aligned}\right\} \quad (6)$$

Then, if

$$\begin{aligned} \gamma &= x_1\beta_1 + x_2\beta_2 + x_3\beta_3 \\ &= (x_1, x_2, x_3) \text{ relative to } \{\beta_1, \beta_2, \beta_3\} \end{aligned}$$

*The remainder of this summary is equally valid if we replace E^3 by E^n and talk about n-tuples rather than 3-tuples.

3.1.3(L) continued

and we want to express γ relative to $\{\alpha_1, \alpha_2, \alpha_3\}$ then γ $(y_1, y_2, y_3) = y_1\alpha_1 + y_2\alpha + y_3\alpha_3$ where

$$[y_1 \quad y_2 \quad y_3] = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{7}$$

More symbolically, if we let $Y = [y_1\ y_2\ y_3]$, $X = [x_1\ x_2\ x_3]$ and A the matrix of coefficients in (6), then (7) becomes

$$Y = XA. \tag{8}$$

Moreover, since $(AB)^T = B^T A^T$, if we want (8) expressed so that the 3 by 3 matrix appears to the left of X, we may deduce from (8) that

$$Y^T = (XA)^T$$

or

$$Y^T = A^T X^T. \tag{9}$$

That is,

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{10}$$

Note:

Hopefully our present discussion of inverting systems of equations resembles our discussion in Part 2. To carry this analogy still further, an associated problem with coordinate system representation centers around inverting system (6). Namely, suppose we are given the vector $v \in E^3$ but in terms of α -coordinates rather than β-coordinates, say, $v = (y_1, y_2, y_3) = y_1\alpha_1 + y_2\alpha_2 + y_3$ and we want to express v as

3.1.3(L) continued

$$v = (x_1, x_2, x_3)$$

$$= x_1\beta_1 + x_2\beta_2 + x_3\beta_3 ,$$

Clearly, this is the inverse of the problem we just tackled.

Our point is that if A^{-1} exists (i.e., A is a non-singular matrix), we may then multiply both sides of (8) on the right by A^{-1} to obtain

$$YA^{-1} = (xA)A^{-1}$$

$$= x(AA^{-1})$$

$$= xI_3$$

$$= x.$$

That is,

$$x = YA^{-1}. \qquad (11)$$

Since A is known and A^{-1} exists, we may compute A^{-1} and since $Y = [y_1\ y_2\ y_3]$ where y_1, y_2, and y_3 are given values; we may solve (11) for $x = [x_1\ x_2\ x_3]$.

3.1.4

a. To invert

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 4 & 9 & 9 \end{bmatrix} \qquad (1)$$

we have

$$\begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 7 & 0 & 1 & 0 \\ 4 & 9 & 9 & 0 & 0 & 1 \end{bmatrix} \sim$$

3.1.4 continued

$$\begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 1 & -3 & -4 & 0 & 1 \end{bmatrix} \sim$$

$$\begin{bmatrix} 1 & 0 & 1 & 5 & -2 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 0 & -4 & -2 & -1 & 1 \end{bmatrix} \sim$$

$$\begin{bmatrix} 4 & 0 & 4 & 20 & -8 & 0 \\ 0 & 4 & 4 & -8 & 4 & 0 \\ 0 & 0 & -4 & -2 & -1 & 1 \end{bmatrix} \sim$$

$$\begin{bmatrix} 4 & 0 & 0 & 18 & -9 & 1 \\ 0 & 4 & 0 & -10 & 3 & 1 \\ 0 & 0 & -4 & -2 & -1 & 1 \end{bmatrix} \sim$$

$$\begin{bmatrix} 1 & 0 & 0 & \frac{18}{4} & -\frac{9}{4} & \frac{1}{4} \\ 0 & 1 & 0 & -\frac{10}{4} & \frac{3}{4} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{2}{4} & \frac{1}{4} & -\frac{1}{4} \end{bmatrix},$$

from which we conclude that

$$A^{-1} = \begin{bmatrix} \frac{18}{4} & -\frac{9}{4} & \frac{1}{4} \\ -\frac{10}{4} & \frac{4}{4} & \frac{1}{4} \\ \frac{2}{4} & \frac{1}{4} & -\frac{1}{4} \end{bmatrix}. \qquad (2)$$

b. Given that

$$\beta_1 = \alpha_1 + 2\alpha_2 + 3\alpha_3$$

$$\beta_2 = 2\alpha_1 + 5\alpha_2 + 7\alpha_3$$

$$\beta_3 = 4\alpha_1 + 9\alpha_2 + 9\alpha_3$$

3.1.4 continued

and that $\gamma_1 = (3, -2, 1)$ relative to $\{\beta_1, \beta_2, \beta_3\}$, That is, $\gamma_1 = 3\beta_1 - 2\beta_2 + \beta_3$. Then using equation (7) of the previous exercise, we have that

$$\gamma_1 = (y_1, y_2, y_3)$$

relative to $\{\alpha_1, \alpha_2, \alpha_3\}$; where

$$[y_1\ y_2\ y_3] = [3\ -2\ \ 1]\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 4 & 9 & 9 \end{bmatrix}$$

$$= [(3 - 4 + 4)\ (6 - 10 + 9)\ (9 - 14 + 9)$$

$$= [3\ 5\ 4];$$

whereupon

$$y_1 = 3$$
$$y_2 = 5$$
$$y_3 = 4.$$

In other words,

$\gamma_1 = (3,-2,1)$ relative to β -coordinates

but

$\gamma_1 = (3,5,4)$ relative to α -coordinates.

c. Make sure that you notice the difference between this part and part (b). In this part $\gamma_2 = (3,-2,1)$ is relative to α-coordinates, while in the previous part (3,-2,1) was relative to β -coordinates.

We now want to express γ_2 in β-coordinates. Using equation (11) of the previous exercise, we have that

$$\gamma_2 = x_1\beta_1 + x_2\beta_2 + x_3\beta_3$$

3.1.4 continued

where $[x_1\ x_2\ x_3] = [3\ -2\ 1]\ A^{-1}$, or, from part (a)

$$[x_1\ x_2\ x_3] = [3\ -2\ 1]\begin{bmatrix} \frac{18}{4} & -\frac{9}{4} & \frac{1}{4} \\ -\frac{10}{4} & \frac{3}{4} & \frac{1}{4} \\ \frac{2}{4} & \frac{1}{4} & -\frac{1}{4} \end{bmatrix} \qquad (3)$$

$$= \left[\frac{54 + 20 + 2}{2} \quad \frac{-27 - 6 + 1}{4} \quad \frac{3 - 2 - 1}{4}\right]$$

$$= [19\ -8\ \ 0].$$

Therefore,

$x_1 = 19$, $x_2 = -8$, $x_3 = 0$;

and

$$\gamma_2 = 19\beta_1 - 8\beta_2 + 0\beta_3$$

$= (19, -8, 0)$ relative to β -coordinates.

<u>Check</u>

$$\begin{aligned} 19\beta_1 - 8\beta_2 &= 19\alpha_1 + 38\alpha_2 + 57\alpha_3 \\ &\quad - 16\alpha_1 - 40\alpha_2 - 56\alpha_3 \\ &= 3\alpha_1 - 2\alpha_2 + \alpha_3 \\ &= \gamma_2. \end{aligned}$$

<u>Note:</u>
What should be the same as part (b) is that if we hadn't done (b) and were told to convert $\gamma = 3\alpha_1 + 5\alpha_2 + 4\alpha_3$ into β-coordinates, we should obtain as our answer $\gamma = 3\beta_1 - 2\beta_2 + \beta_3$. That this is the case, follows just as in our derivation of (3), only with [3 5 4] replacing [3 -2 1]. Namely, $\gamma = x_1\beta_1 + x_2\beta_2 + x_3\beta_3$ where

3.1.4 continued

$$[x_1 \ x_2 \ x_3] = [3 \ 5 \ 4]\begin{bmatrix} \frac{18}{4} & -\frac{9}{4} & \frac{1}{4} \\ -\frac{10}{4} & \frac{3}{4} & \frac{1}{4} \\ \frac{2}{4} & \frac{1}{4} & -\frac{1}{4} \end{bmatrix}$$

$$= \left[\frac{54 - 50 + 8}{4} \quad \frac{-27 + 15 + 4}{4} \quad \frac{3 + 5 - 4}{4}\right]$$

$$= [\ 3 \quad -2 \quad 1\ ].$$

3.1.5(L)

Our aim here is to reinforce the role of the cancellation theorem in determining the structure of a vector space V. We begin with the proof of the cancellation theorem.

a. Since $\beta + \alpha = \gamma + \alpha$, we may add $-\alpha$*

$(\beta + \alpha) + (-\alpha) = (\gamma + \alpha) + (-\alpha)$,

or since addition is associative

$\beta + [\alpha + (-\alpha)] = \gamma + [\alpha + (-\alpha)]$.

This, in turn, since $\alpha + (-\alpha) = 0$, implies that

$\beta + 0 = \gamma + 0$;

and by the property of 0, the conclusion $\beta = \gamma$ follows. That is, for $\alpha, \beta, \gamma \varepsilon V$; $\beta + \alpha = \gamma + \alpha \rightarrow \beta = \gamma$.

b. We could mimic part (a) but the quicker way is to use the commutative property and then invoke the result of part (a). Namely,

---

*Notice here that we are using the fact that V is a vector space by assuming that the element $-\alpha \ \varepsilon \ V$ exists. That is, we are using the property that for $\alpha \varepsilon V$, there exists $-\alpha \varepsilon V$ such that $\alpha + (-\alpha) = 0$, etc.

3.1.5(L) continued

$\alpha + \beta = \alpha + \gamma \rightarrow$

$\beta + \alpha = \gamma + \alpha$, and this, by part (a), implies that $\beta = \gamma$.

Notice that (a) and (b) together tell us that the cancellation theorem allows us to cancel the common term regardless of the position it occupies in the equation.

We now look at a few consequences of the cancellation theorem on the structure of a vector space.

c. $0\alpha = (0 + 0)\alpha \rightarrow$
$0\alpha = 0\alpha + 0 \rightarrow$
$0 + 0 = + 0\alpha \rightarrow$
$0 = 0\alpha$, or $0\alpha = 0$.

d. $c\vec{0} = c(\vec{0} + \vec{0}) \rightarrow$
$c\vec{0} = c\vec{0} + c\vec{0} \rightarrow$
$c\vec{0} + \vec{0} = c\vec{0} + c\vec{0}$
$\vec{0} = c\vec{0}$, or $c\vec{0} = \vec{0}$.

e. $\alpha + (-1)\alpha = 1\alpha + (-1)\alpha \rightarrow$
$\alpha + (-1)\alpha = [1 + (-1)]\alpha \rightarrow$
$\alpha + (-1)\alpha = 0\alpha$; hence from (c),
$\alpha + (-1)\alpha = 0$.

f. $\alpha + (-\alpha) = 0$, by definition of $(-\alpha)$. Moreover, by (e), $\alpha + (-1)\alpha = 0$. Since $0 = 0$, it follows that $\alpha + (-1)\alpha = \alpha + (-\alpha)$; so that by cancellation, $(-1)\alpha = -\alpha$.

Part (f) can be generalized as follows. If $\alpha \varepsilon V$ and $\beta \varepsilon V$ and if $\alpha + \beta = 0$, then $\beta = -\alpha$(that is, $-\alpha$ is usually uniquely determined by α). Namely, since $\alpha + (-\alpha)$ is also 0, we have that $\alpha + \beta = \alpha + (-\alpha)$. Hence, by cancellation, $\beta = -\alpha$. Part (f) was a special case of this more general result with $\beta = (-1)\alpha$. This special case plays an important role in many applications, one of which will occur in the next exercise.

g. Since $\alpha + \beta = \alpha$ and since $\alpha + 0 = \alpha$, we have that $\alpha + \beta = \alpha + 0$; whence, again by cancellation, $\beta = 0$.

Notice that part (g) tells us that the element 0 of V is uniquely determined. In other words, it is not possible to have two elements of V, say 0 and 0' such that $v + 0 = v + 0' = v$ for $v \varepsilon V$, unless $0 = 0'$.

3.1.5(L) continued

From another point of view, what this says is that suppose we have a non-empty subset S of V with the property that there exists an element 0_s in S usch that $s + 0_s = s$ for each $s \epsilon S$. Then $0_s = 0$, itself. Namely, since S is a subset of V, the fact that $s \epsilon S$ implies that $s \epsilon V$; and since $s \epsilon V$, the definition of 0 implies that $s + 0 = s$. Hence, $s + 0_s = s + 0$, so by calcellation, $0_s = 0$. Some of the implications of this discussion will appear in the next exercise.

3.1.6(L)

a. Our main purpose here is to emphasize the difference between a subset and a subspace. For a start, then, let us oassume that V is a vector space in the sense that it satisfies the nine properties mentioned in the lecture; and that S is any <u>subset</u> of V.

Suppose now that α, β, γ now refer to members of S while c, c_1, c_2, etc. still refer to real numbers. Then, each of the following axioms (or, properties) for V are automatically obeyed by S.

(2)* $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$

(5) $\alpha + \beta = \beta + \alpha$

(6) $c(\alpha + \beta) = c\alpha + c\beta$

(7) $(c_1 + c_2) \quad = c_1 \quad + c_2$

(8) $c_1(c_2 \;) = (c_1 c_2)$

(9) $1\alpha = \alpha$.

The reason for this is that since $S \subseteq V$, the fact that $\alpha, \beta, \gamma \epsilon S$ also implies that $\alpha, \beta, \gamma \epsilon V$; and axioms (2), (5), (6), (7), (8), and (9) hold for all elements of V since V is a vector space.

b. What we can <u>not</u> be sure of, however, is

(1) $\alpha, \beta \epsilon S \rightarrow \alpha + \beta \epsilon S$.

In fact, all we know for sure is that $\alpha + \beta \epsilon V$, since α and β both belong to V. In essence, if a structural set is closed with respect to a given operation, it need not be true that an arbitrary subset of the structured set is also closed with respect to this given operation. For example, the integers are closed with respect to addition since the sum of two integers is again an

*These numbers refer to those used in numbering the vector space axioms as presented in the lecture.

3.1.6(L) continued

integer. $S = \{1,2\}$ is a subset of the integers but S is not closed with respect to addition since, for instance, $1 \varepsilon S$ and $2 \varepsilon S$, but $1 + 2 = 3 \notin S$.

(3) There exists $0_s \varepsilon S$ such that $\alpha + 0_s =$ for each $\alpha \varepsilon S$. We have written 0_s to emphasize that the identity element must belong to the set under consideration. For example by property (3) $\alpha + 0 =$ for each $\alpha \varepsilon S$ but there is no guarantee that $0 \varepsilon S$.*

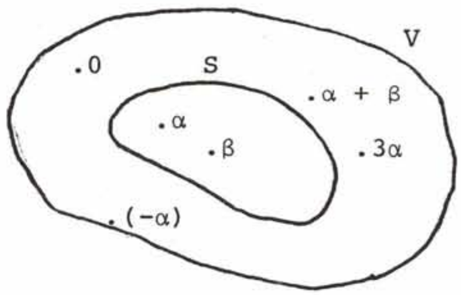
(4) $\alpha + (-\alpha) = 0$.

Again, what we do know by virtue of V being a vector space is that there exists $(-\alpha) \varepsilon V$, but not necessarily in S, such that $\alpha + (-\alpha) = 0$, for each $\alpha \varepsilon S$.*

(6') $\alpha \varepsilon S \rightarrow c\alpha \varepsilon S$.

We use (6') here to take into account the fact that our blackboard writing was "sketchy". When we talk about scalar multiplication, it is assumed that $c\alpha$ belongs to the same set as α. Notice that since $S \subset V$, $\alpha \varepsilon S \rightarrow \alpha \varepsilon V$; so that since V is a vector space, we can be sure that $c\alpha \varepsilon V$ but we cannot be sure that **c** $\alpha \varepsilon S$.

Pictorially, these are things that could happen.



*What we showed in the previous exercise is that if 0_s exists, then it is the identity element, 0, of V itself.

**Again consider our earlier example $S = \{1,2\}$. Certainly there exists an integer -2 such that $2 + (-2) = 0$ but $-2 \notin S$.

3.1.6(L) continued

c. Using part (a) for background, suppose that we now have the following additional knowledge about S.

1. $\alpha, \beta \varepsilon S \rightarrow \alpha + \beta \varepsilon S$
2. $c \varepsilon R, \quad \alpha \varepsilon S \rightarrow c\alpha \varepsilon S$

Clearly, (1) and (2) make sure that S now obeys (1) and (6'). So all we have to do now is show that (1) and (2) are also sufficient to guarantee that (3) and (4) are also obeyed by S, since then all nine axioms will be obeyed by S.

Since $-\alpha = (-1)\alpha$, and $(-1) \varepsilon R$, we see by (2) that $\alpha \varepsilon S \rightarrow (-1)\alpha \varepsilon S \rightarrow -\alpha \varepsilon S$. Now by (1), since α and $-\alpha$ both belong to S so also does $\alpha + (-\alpha)$; but $\alpha + (-\alpha) = 0$. Hence, $0 \varepsilon S$. The facts that 0 and $-\alpha$ belong to S is all that we need to verify that S obeys (3) and (4).

This justifies the criteria described in the lecture. Namely, if S is a subset of the vector space V, then S is a subspace of V $\leftrightarrow$.

1. $\alpha \varepsilon S, \ \beta \varepsilon S \rightarrow \alpha + \beta \varepsilon S$
2. $c \varepsilon R, \ \alpha \varepsilon S \rightarrow c\alpha \varepsilon S$

3.1.7(L)

Here $V = E^2 \ \{(x_1, x_2): x_1, x_2 \varepsilon R\}$.

a. $S = \{(0,0)\}$

Clearly, S is a subset of V since $(0,0) \varepsilon E^2$. Since $\alpha \varepsilon S \rightarrow \alpha = 0$, it is easily verified that each of our axioms (1) through (9) is obeyed.

In terms of the "short-cut" we have

1. $\alpha, \ \beta \varepsilon S \rightarrow \alpha + \beta \varepsilon S$
2. $c \varepsilon R, \ \alpha \varepsilon S \rightarrow c\alpha \varepsilon S$

Namely,

3.1.7(L) continued

$\alpha,\beta\varepsilon S \to \alpha = \beta = 0 \to \alpha + \beta = 0\varepsilon S$

and

$\alpha\varepsilon S \to c\alpha = c0 \to c\alpha = 0\varepsilon S.$

Note:
Part (a) generalizes as follows. If V is any vector space, then $S = \{\vec{0}\}$ is a subspace of V. Namely $0 + 0 = 0$ and $c0 = 0$ for each $c\varepsilon R$.

Since the space which consists of only the 0-vector is rather trivial, one frequently stipulates that the subspace have more than just the 0 vector.

Notice that once a subspace S contains a non zero vector α, it must contain infinitely many vectors, namely all scalar multiples of α. That is $\alpha\varepsilon S$ implies $c\alpha\varepsilon S$ for all $c\varepsilon R$ if S is a subspace of V. Notice that if $c_1\alpha = c_2\alpha$ then $c_1 - c_2 = 0$ or $(c_1 - c_2)\alpha = 0$. Hence, either $c_1 - c_2 = 0$ or $\alpha = 0$. Therefore, if $\alpha \neq 0$, then $c_1 - c_2 = 0$ or $c_1 = c_2$. In other words, if $\alpha \neq 0$ and c_1 and c_2 are different (unequal) scalars, then $c_1\alpha \neq c_2\alpha$. In other words, if $\alpha \neq 0$, there are infinitely many different scalar multiples of α.

b. $S = \{(x_1,x_2): x_1 = 0 \text{ or } x_2 = 0\}$.

Now, S contains not only (0,0) but, among others, (0,1) and (1,0). But every vector in V (not just in S) is a linear combination of (0,1) and (1,0). Hence S is not closed with respect to addition. For example, $(2,3)\notin S$ (since neither 2 nor 3 equals 0); yet $(2,3) = (2,0) + (0,3)$ where both (2,0) and (0,3) belong to S.

In this example, all that S lacks for being a subspace of $V = E^2$ is closure with respect to addition. All the other axioms for a vector space are possessed by S.

c. $S = \{(x_1,x_2): x_2 = x_1 + 1\}$

One quick way of concluding that S is not a subspace of V is that $0\notin S$. Namely, $0 = (0,0)$ in which case $x_2 \neq x_1 + 1$ (i.e., $0 \neq 0 + 1$). Another way is to observe that if $(x_1,x_2)\varepsilon S$ and $(y_1,y_2)\varepsilon S$, then

3.1.7(L) continued

$x_2 = x_1 + 1$
$y_2 = y_1 + 1.$

Hence, $x_2 + y_2 = x_1 + y_1 + 2 \neq x_1 + y_1 + 1$. Therefore, (x_1,x_2) and $(y_1,y_2) \in S$, but $(x_1,x_2) + (y_1,y_2) = (x_1 + y_1, x_2 + y_2) \notin .$

d. $S = \{(x_1,x_2): x_2 = 3x_1\}$

In this case, suppose (x_1,x_2) and (y_1,y_2) both belong to S.

Then,

$x_2 = 3x_1$
$y_2 = 3y_1.$

Hence, $x_2 + y_2 = 3(x_1 + y_1)$. Therefore, $(x_1 + y_1, x_2 + y_2) \in S$. In other words, $(x_1, x_2) \in S, (y_1,y_2) \in S \rightarrow (x_1,x_2) + (y_1,y_2) \in S$. Similarly, $c(x_1,x_2) = (cx_1,cx_2)$ and since $x_2 = 3x_1 \rightarrow cx_2 = 3cx_1$; $(cx_1,cx_2) \in S$.

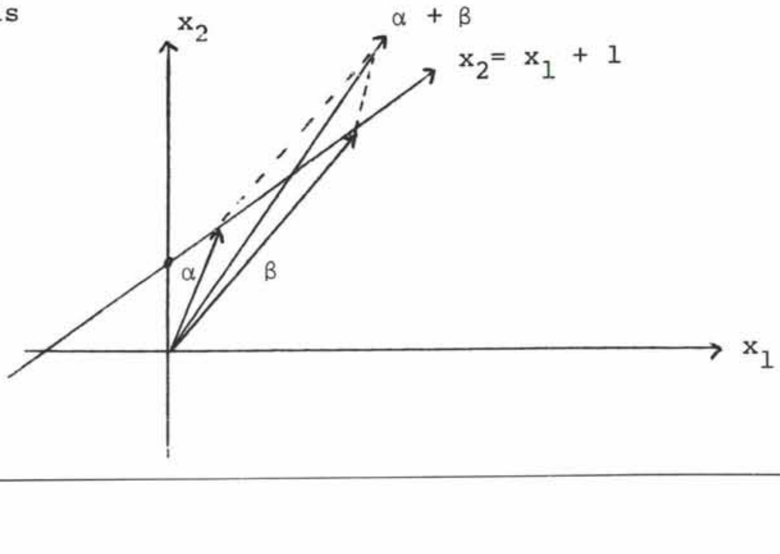
Hence, S is closed with respect to addition and scalar multiplication, so S is a subspace of V.

Geometric Interpretation of Parts (c) and (d).

If we assume that each vector $(x_1,x_2) = x_1\vec{i} + x_2\vec{j}$ in the x_1x_2-plane originates at (0,0), then it terminates at the point (x_1,x_2).

Thus, $\{(x_1,x_2): x_2 = x_1 + 1\}$ represents the set of all vectors which originate at (0,0) and terminate on the line $x_2 = x_1 + 1$.

That is



3.1.7(L) continued

1. α and β belong to S; i.e., α and β terminate on $x_2 = x_1 + 1$.
2. $\alpha + \beta$ doesn't terminate on $x_2 = x_1 + 1$.

The main idea is that unless the line passes through the origin, the set of vectors which originate at (0,0) and terminate on the given line will not be closed with respect to either addition or scalar multiplication. If the line passes through the origin, then all our vectors have the same direction, namely, that of the line itself. This is why in linear algebra we require that all "lines" pass through 0 (the "origin").

3.1.8

$V = \{f: \text{dom } f = [0,1]\}$

a. $S = \{f \varepsilon V: f(x) = f(1 - x)\}$

To test whether S is a subspace of V, it is necessary and sufficient to show that $f\varepsilon S$ and $g\varepsilon S \rightarrow f + g\varepsilon S$ and $c\varepsilon R$, $f\varepsilon S$ $cf\varepsilon S$. To this end, suppose $h = f + g$ where $f, g\varepsilon S$. Then, for x [0,1]

$$\begin{aligned} h(x) &= f(x) + g(x) \\ &= f(1 - x) + g(1 - x) \\ &= h(1 - x). \end{aligned}$$

Since $h(x) = h(1 - x)$, $h\varepsilon S$.

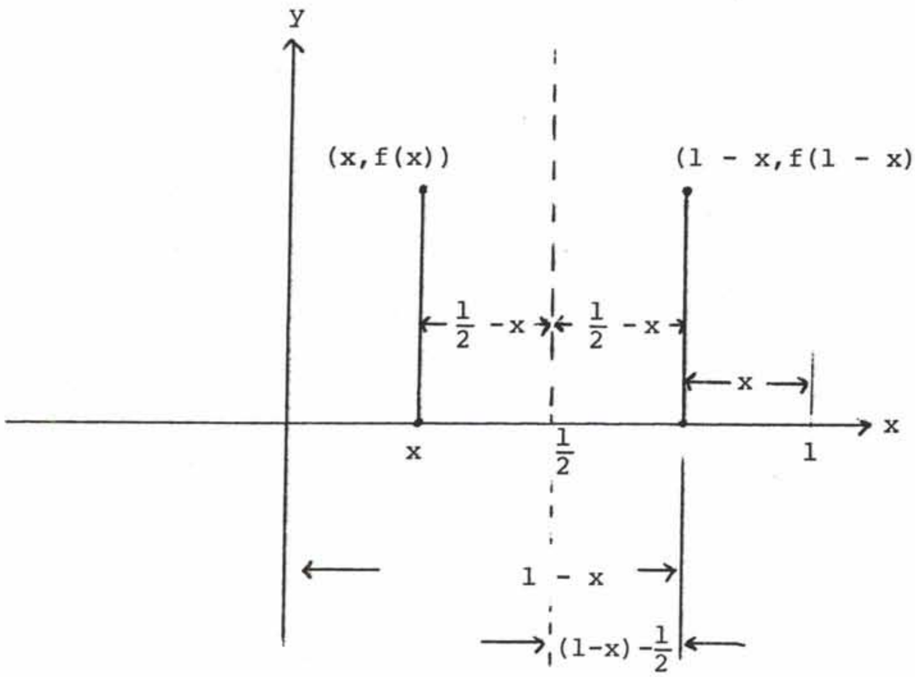
Next, if $k(x) = cf(x)$ where $f\varepsilon S$, we have

$$\begin{aligned} k(x) &= cf(x) \\ &= cf(1 - x) \\ &= k(1 - x) . \end{aligned}$$

Hence, $k(= cf)\varepsilon S$. Therefore, S is a subspace of V.

<u>Geometric Note</u>:
The set of all curves $y = f(x)$ defined for $0 \leq x \leq 1$ such that $f(x) = f(1 - x)$ is precisely that set of curves which are symmetric with respect to the line $x = 1/2$. Pictorially,

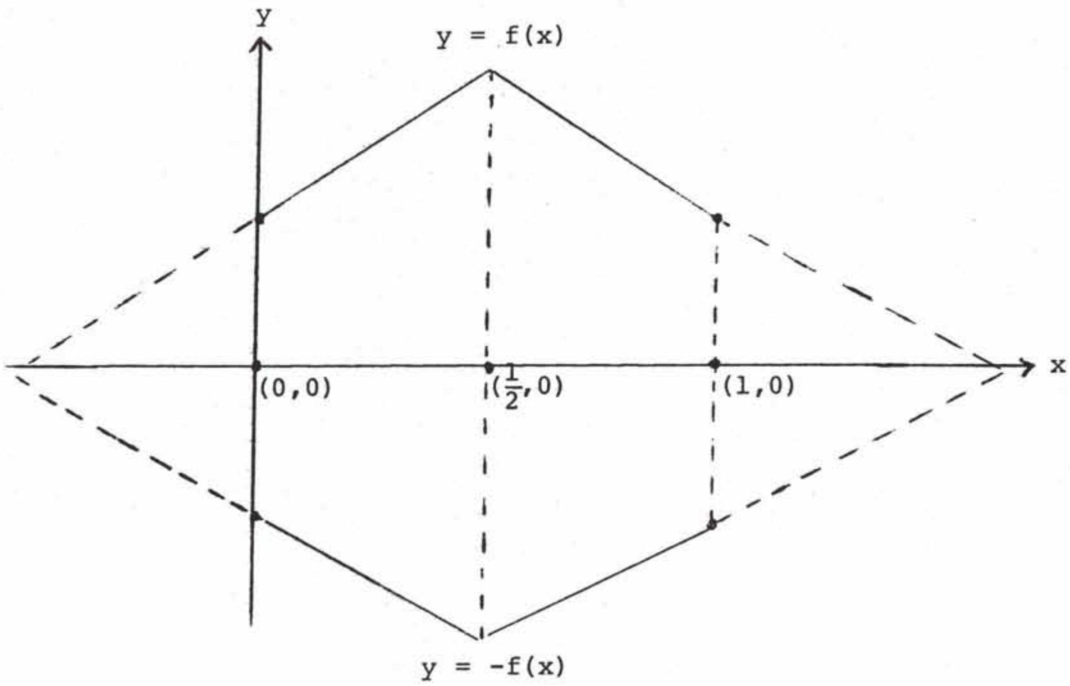
3.1.8 continued



Thus, what we are saying in this exercise is that the sum of two curves symmetric with respect to the line $x = 1/2$ is also symmetric with respect to the line $x = 1/2$; and any scalar multiple of such a curve is symmetric to the line $x = 1/2$.

As a check that the truth of $f \varepsilon S$ and $g \varepsilon S \rightarrow f + g \varepsilon S$ and $c \varepsilon R$, $f \varepsilon S \rightarrow cf \varepsilon S$ really do imply Axioms (1) through (9), we essentially need only check that $0 \varepsilon S$ and that $-f \varepsilon S$ if $f \varepsilon S$. All the other axioms are obeyed by all functions defined on [0,1]. Well, clearly if $f(x) \equiv 0$ for all x [0,1], then $f(x) = f(1 - x)$ for all $x \varepsilon [0,1]$ since both expressions equal 0. Moreover, since $y = -f(x)$ is simply the reflection of $y = f(x)$ about the x-axis, then the fact that $y = f(x)$ is symmetric to the line $x = 1/2$ guarantees that $y = -f(x)$ is also symmetric with respect to $x = 1/2$. For example

3.1.8 continued



b. $S = \{f\varepsilon V: f(0) = 2\}$

In this case, if $f\varepsilon S$ and $g\varepsilon S$, then $f(0) = g(0) = 2$. Therefore, $f(0) + g(0) = 2 + 2 = 4$. That is, letting $h = f + g$, $h(0) = f(0) + g(0) = 4$. Hence $h\notin S$.

Therefore, S is not a subspace of V.

Geometrical Interpretation

S is the set of all curves defined on [0,1] which pass through (0,2). Not only is the sum of two such curves not a member of S, but S fails to admit the 0-function or inverses. For example, if $f(x) \equiv 0$, then $y = f(x)$ passes through (0,0) not (0,2); and if $f\varepsilon S$ and $f(x) + g(x) \equiv 0$, then $g(0) = -2$, not 2, since $f(0) + g(0) = 0 \rightarrow 2 + g(0) = 0$. Hence $g\notin S$.

As far as scalar multiplication is concerned, if $f(0) = 2$, then $cf(0) = 2 \leftrightarrow c = 1$. Hence, if $f\varepsilon S$ and $c \neq 1$, then $cf\notin S$.

3.1.9(L)

a. By <u>definition</u>

$$W = S(\alpha_1,\alpha_2) = \{x_1\alpha_1 + x_2\alpha_2 : x_1,x_2 \varepsilon R\} . \quad (1)$$

b. Hence,

$$\begin{aligned} w\varepsilon W &\leftrightarrow w = x_1\alpha_1 + x_2\alpha_2 \\ &\leftrightarrow w = x_1(1,2,3) + x_2(3,5,5) \\ &\leftrightarrow w = (x_1,2x_1,3x_1) + (3x_2 + 5x_2 + 5x_2) \\ &\leftrightarrow w = (x_1 + 3x_2,\ 2x_1 + 5x_2,\ 3x_1 + 5x_2). \end{aligned} \quad (2)$$

c. According to (2) we see that for (2,3,2) to belong to W, there must exist real numbers x_1 and x_2 such that

$$\left.\begin{aligned} x_1 + 3x_2 &= 2 \\ 2x_1 + 5x_2 &= 3 \\ 3x_1 + 5x_2 &= 2 \end{aligned}\right\} \quad (3)$$

Clearly (3) is a special case of the more general result that $(a,b,c)\varepsilon w \leftrightarrow$

$$\left.\begin{aligned} x_1 + 3x_2 &= a \\ 2x_1 + 5x_2 &= b \\ 3x_1 + 5x_2 &= c \end{aligned}\right\} \quad (3')$$

The point is that (3') is a system which has more equations (three) than unknowns (two). Consequently, either (3') will be an inconsistent system or else at least one of the equations must be contained in the others. In fact, using our row-reduced matrix technique, we see that (3') becomes

$$\begin{matrix} x_1 & x_2 & a & b & c \end{matrix}$$

$$\begin{bmatrix} 1 & 3 & 1 & 0 & 0 \\ 2 & 5 & 0 & 1 & 0 \\ 3 & 5 & 0 & 0 & 1 \end{bmatrix} \sim$$

$$\begin{bmatrix} 1 & 3 & 1 & 0 & 0 \\ 0 & -1 & -2 & 1 & 0 \\ 0 & -4 & -3 & 0 & 1 \end{bmatrix} \sim$$

3.1.9(L) continued

$$\begin{bmatrix} 1 & 0 & -5 & 3 & 0 \\ 0 & -1 & -2 & 1 & 0 \\ 0 & 0 & 5 & -4 & 1 \end{bmatrix} \sim$$

$$\begin{bmatrix} x_1 & x_2 & a & b & c \\ \hline 1 & 0 & -5 & 3 & 0 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 5 & -4 & 1 \end{bmatrix} \qquad (4)$$

From (4) we now have the following interesting information:

$$\begin{aligned} x_1 &= -5a + 3b \\ x_2 &= 2a - b \end{aligned} \qquad (5)$$

<u>and</u>

$$0 = 5a - 4b + c; \text{ or } c = 4b - 5a. \qquad (6)$$

Therefore, we see from (6) that unless $c = 4b - 5a$, $(a,b,c) \notin W$. On the other hand, if $c = 4b - 5a$, then not only is $(a,b,c) \in W$ but [from (5)] $(a,b,c) = x_1(1,2,3) + x_2(3,5,5)$, where $x_1 = -5a + 3b$ and $x_2 = 2a - b$. In still other words,

$$(a,b,c) \in \{x_1(1,2,3) + x_2(3,5,5)\} \leftrightarrow c = 4b - 5a$$

and in this case

$$(a,b,c) = (-5a + 3b)(1,2,3) + (2a - b)(3,5,5). \qquad (7)$$

Applying this discussion to (2,3,2), we have that $a = 2$, $b = 3$, and $c = 2$. Hence, $4b - 5a = 12 - 10 = c$ so that by (6), $(a,b,c) \in w$. Moreover, from (5) $x_1 = -5(2) + 3(3) = -1$ and $x_2 = 2(2) - 3 = 1$ so that

$$\begin{aligned} (2,3,2) &= x_1(1,2,3) + x_2(3,5,5) \\ &= -(1,2,3) + (3,5,5) \\ &= (2,3,2). \end{aligned}$$

3.1.9(L) continued

d. Since $4(4) - 5(3) = 1 \neq 3$, we have from (6) that $(3,4,3) \notin W$.

e. Since $4(4) - 5(3) = 1$, $(3,4,c) \in W \leftrightarrow c = 1$. Hence, $(3,4,1) \in W$. Moreover, from (5) we have that $x_1 = -5(3) + 3(4) = -3$ and $x_2 = 2(3) - 4 = 2$. Hence, $(3,4,1) = -3(1,2,3) + 2(3,5,5)$.

<u>Check:</u>

$-3(1,2,3) + 2(3,5,5) = (-3,-6,-9) + (6,10,10) = (3,4,1)$.

f. W is the plane spanned by $\alpha_1 = \vec{i} + 2\vec{j} + 3\vec{k}$ and $\alpha_2 = 3\vec{i} + 5\vec{j} + 5\vec{k}$. The vector $v = a\vec{i} + b\vec{j} + c\vec{k}$ lies in the plane $w \leftrightarrow c = 4b - 5a$ and in this case

$$\vec{v} = (-5a + 3b)\alpha_1 + (2a - b)\alpha_2.$$

In the next Unit we shall discuss the idea behind this problem in more detail. In particular, we shall discuss how we find a pair of more "informative" vectors with which we may describe the space W. For now our hope is that it is clear what we mean when we talk about the subspace spanned by a set of vectors of V. While we have chosen an exercise in which we could relate the answer to a simple geometric interpretation, notice that the general idea does not depend on our having to talk about a 2-dimensional subspace (namely a plane) of 3-dimensional space.

The point is that with this exercise as an introductory example, we may use it as reinforcement in our more general treatment of the next Unit.

3.1.10 (optional)

a. If we let $w = S(\alpha_1,\alpha_2)$, then by definition of w we have that if $w_1 \in W$ and $w_2 \in W$, then

$$\left.\begin{array}{l} w_1 = a_1\alpha_1 + a_2\alpha_2 \quad (1) \\ w_2 = b_1\alpha_1 + b_2\alpha_2 \quad (2) \end{array}\right\} \text{ where } a_1,a_2,b_1,b_2 \in R.$$

Hence, by the usual vector arithmetic structure,

3.1.10 continued

$$w_1 + w_2 = \underbrace{(a_1 + b_1)}_{\varepsilon R}\alpha_1 + \underbrace{(a_2 + b_2)}_{\varepsilon R}\alpha_2 \,. \tag{3}$$

Since $w_1 + w_2$ is a linear combination of α_1 and α_2, it follows by the definition of w that:

$$1. \quad w_1 \varepsilon W,\ w_2 \varepsilon W \rightarrow w_1 + w_2 \varepsilon W. \tag{4}$$

Moreover, for an $c\varepsilon R$ and $w_1 \varepsilon W$, we have from (1) that

$$\begin{aligned} cw_1 &= c(a_1\alpha_1 + a_2\alpha_2) \\ &= ca_1\alpha_1 + ca_2\alpha_2 \\ &= \underbrace{(ca_1)}_{\varepsilon R}\alpha_1 + \underbrace{(ca_2)}_{\varepsilon R}\alpha_2, \end{aligned}$$

so that cw_1 is a linear combination of α_1 and α_2.

Hence,

$$2. \quad c\varepsilon R,\ w_1 \varepsilon W \rightarrow cw_1 \varepsilon W. \tag{5}$$

From (4) and (5) we see that $W = S(\alpha_1, \alpha_2)$ is a subspace of V.

<u>Note #1</u>:

The proof given here generalizes very nicely to the case of $\alpha_1, \ldots, \alpha_n$. Namely, the sum of two linear combinations of $\alpha_1, \ldots,$ and α_n is also a linear combination of $\alpha_1, \ldots,$ and α_n. For example, if

$$\beta = b_1\alpha_1 + \ldots + b_n\alpha_n$$

and

$$\gamma = c_1\alpha_1 + \ldots + c_n\alpha_n$$

then

3.1.10 continued

$\beta + \gamma = (b_1 + c_1)\alpha_1 + \ldots + (b_n + c_n)\alpha_n.$

Note #2:

If W is any subspace of V which contains the vectors $\alpha_1,\ldots,$ and α_n, then W must contain $S(\alpha_1,\ldots,\alpha_n)$. Namely, by virtue of W being a subspace, any linear combination of vectors in W must be a member of W. In other words, if $\alpha_1,\ldots,\alpha_n \varepsilon W$ where W is a subspace of V and $\alpha_1,\ldots,\alpha_n$ linearly independent; then $S(\alpha_1,\ldots,\alpha_n)$ is a subspace of W. Thus, $S(\alpha_1,\ldots,\alpha_n)$ is the smallest subspace of V which contains $\alpha_1,\ldots,$ and α_n.

b. If S and T are subspaces of V and $\alpha \varepsilon S \cap T$ and $\beta \varepsilon S \cap T$; then by definition of intersection

$\alpha \varepsilon S$ and $\alpha \varepsilon T$
$\beta \varepsilon S$ and $\beta \varepsilon T$.

Since S is a subspace, $\alpha, \beta \varepsilon S \rightarrow \alpha + \beta \varepsilon S$. Similarly, since T is a subspace $\alpha, \beta \varepsilon T \rightarrow \alpha + \beta \varepsilon T$. Therefore, $\alpha + \beta \varepsilon S$ and $\alpha + \beta \varepsilon T$. That is, $\alpha + \beta \varepsilon S \cap T$. Moreover, $\alpha \varepsilon S \rightarrow c\alpha \varepsilon \cap S$; $\alpha \varepsilon T$ $c\alpha \varepsilon$ T. Hence $c\alpha \varepsilon S \cap T$.

Consequently, $S \cap T$ is also a subspace of V since it is closed with respect to both addition and scalar multiplication.

Note:

Part (b), together with a little math induction, tells us that the intersection of any number of subspaces of V is again a subspace of V. Combining this with part (a), we have that the intersection of all subspaces of V which contain $\alpha_1,\ldots,$ and α_n is $S(\alpha_1,\ldots,\alpha_n)$.

c. We define S + T to be the set of all elements of V which can be written as the sum of an element in S and an element in T. That is,

$S + T = \{ s + t\colon s \varepsilon S, T \varepsilon T \}.$

Now, suppose S and T happen to be subspaces of V (rather than merely subsets of V).

3.1.10 continued

Then, if α_1 and α_2 belong to S + T, we have

$$\alpha_1 = s_1 + t_1$$
$$\alpha_2 = s_2 + t_2$$

where $s_1, s_2 \varepsilon S$ and $t_1, t_2 \varepsilon T$.

Hence,

$$\alpha_1 + \alpha_2 = (s_1 + s_2) + (t_1 + t_2). \tag{6}$$

The key now is that since S and T are subspaces $s_1 + s_2 \varepsilon S$ and $t_1 + t_2 \varepsilon T$. Thus, from (6) we conclude that $\alpha_1 + \alpha_2 \varepsilon S + T$. We also have that for $c \varepsilon R$ and $\alpha = s + t \varepsilon S + T$,

$$c\alpha = \underbrace{cs + ct}_{\varepsilon S} \ \varepsilon \ \underbrace{S + T}_{\varepsilon T}$$

<u>Note</u>:
If S and T are subspaces of V, then the subspace S + T is called the <u>linear sum</u> of S and T. Since s = s + 0 and t = 0 + t; and since 0 is a member of each subspace of V, we have, as might be expected that $S \subset S + T$ and $T \subset S + T$. We shall talk more about S + T in Unit 3.

3.1.11 (optional)

The main aim of this exercise is to explore some of the subtleties that are involved in making up an axiomatic system. At least in the more pragmatic cases, the axiomatic system is usually obtained by abstracting certain properties of a <u>known</u> physical model. For example, in the present Unit we have defined a vector space axiomatically,based on how we know that "arrows" behave.

Thus, the fact that our axiomatic system is derived from a real model means that we are spared one difficult problem that besets the pure mathematician. Namely, we do not have to worry about whether our axioms are consistent. What is, if our axioms were contradictory then there could be no real model which obeyed each of the axioms.

3.1.11 continued

There are, however, other problems that occur. For example, even though our axioms are consistent we often would like to know if they are independent. In other words, are some of the axioms derivable from the others? That is, are some of the axioms deducible as theorems from the other axioms? If this is the case, then these axioms may be deleted from the list of axioms and added to the list of theorems. Of course, even if an axiom is a theorem, it may be simpler to state it as an axiom anyway; and this can cause no harm since anything which follows inescapably from our axioms is as valid as the axioms themselves.

In addition to this, there is the related case in which one axiom seems to be so simple that it appears unnecessary to have to state it separately. There are many examples in elementary geometry in which these situations occur (in fact part of the "new" geometry is to emphasize these logical aspects of the structure of geometry), but it is not our place to pursue this here. Rather we shall be content to show how one may go about the business of showing that one axiom, obeyed by a particular physical model, is independent of the other axioms. The procedure, quite simply (at least in concept), is to construct another model (even though the model may seem far-fetched from a practical point of view - all that's required is that the model be consistent) in which every axiom <u>except</u> the one in question is obeyed. Then, since we now have two models, one of which obeys every axiom and the other which obeys all but the one in question, we may conclude that the axiom in question is independent of the others. Namely, if it weren't, it would have to be obeyed once the others were obeyed.

We illustrate this idea in part (a) of this exercise. In part (b) of this exercise we show how it is possible that an axiom which seems to be independent of the others may in actuality not be; but this need cause no harm.

a. Since vector addition is defined as before, axioms (1) through (5) (as listed in the lecture) for a vector space must still be obeyed.

3.1.11 continued

Now, according to our new definition of scalar multiplication, wherein $c\alpha = \vec{0}$* for each $c \varepsilon R$ and $\alpha \varepsilon V$ we see that axioms (6), (7), and (8) are also obeyed. Namely:

(6) $\left.\begin{aligned} c(\alpha + \beta) &= \vec{0} \\ c\alpha + c\beta &= \vec{0} + \vec{0} = \vec{0} \end{aligned}\right\}$ Therefore, $c(\alpha + \beta) = c\alpha + \beta$.

(7) $\left.\begin{aligned} (c_1 + c_2)\alpha &= \vec{0} \\ c_1\alpha + c_2\alpha &= \vec{0} + \vec{0} = \vec{0} \end{aligned}\right\}$ Therefore, $(c_1 + c_2)\alpha = c_1\alpha + c_2\alpha$.

(8) $\left.\begin{aligned} c_1(c_2\alpha) &= c_1\vec{0} = \vec{0} \\ (c_1c_2)\alpha &= \vec{0} = \vec{0} \end{aligned}\right\}$ Therefore, $c_1(c_2\alpha) = (c_1c_2)\alpha$.

Clearly, however, if V contains more than the 0-vector, then

$1\alpha = \alpha$ for all $\alpha \varepsilon V$

must be false, since by definition

$$c\alpha = \vec{0} \qquad (1)$$

for all $c \varepsilon R$, in particular then when $c = 1$. Therefore, if $\alpha \neq \vec{0}$, we see from (1) that $1\alpha = \vec{0} \neq \alpha$ so that (9) is not obeyed.

This proves that axiom (9) is independent of axioms (1) through (8) since in both models, (1) through (8) are obeyed, but in one case (9) is obeyed and in the other it isn't.

---

*Notice that we're not concerned here with the question of why one would want to invoke such a "sterile" definition. To be sure it might not have much (if any) practical application, but the definition is meaningful, hence, "legal". Thus, if we can show that our new structure (the vectors are the same but scalar multiplication is different) obeys axioms (1) through (8) but not (9), then we have succeeded in proving that (9) cannot be derived as a theorem from (1) through (8).

3.1.11 continued

b. Treating $\alpha + \beta$ as a single element, say γ, we have from Axiom (7) that

$$\begin{aligned}(1 + 1)(\alpha + \beta) &= (1 + 1)\gamma \\ &= 1\gamma + 1\gamma ,\end{aligned}$$

so by Axiom (9):

$$\begin{aligned}(1 + 1)(\alpha + \beta) &= \gamma + \gamma \\ &= (\alpha + \beta) + (\alpha + \beta). \qquad (1)\end{aligned}$$

On the other hand, treating $|+|$ as a single number, we may use Axiom (6) to conclude that

$$\begin{aligned}(1 + 1)(\alpha + \beta) &= (1 + 1)\alpha + (1 + 1)\beta \\ &= (\alpha + \alpha) + (\beta + \beta). \qquad (2)\end{aligned}$$

Equating the expressions for $(1 + 1)(\alpha + \beta)$ in equation (1) and (2), we obtain

$$(\alpha + \alpha) + (\beta + \beta) = (\alpha + \beta) + (\alpha + \beta),$$

or since vector addition is associative, we may omit parenthesis and write:

$$\alpha + \alpha + \beta + \beta = \alpha + \beta + \alpha + \beta$$

Hence

$$-\alpha + (\alpha + \alpha + \beta + \beta) - \beta = -\alpha + (\alpha + \beta + \alpha + \beta) - \beta,$$

or

$$(-\alpha + \alpha) + (\alpha + \beta) + (\beta - \beta) = (-\alpha + \alpha) + (\beta + \alpha) + (\beta - \beta).$$

Therefore,

$$0 + (\alpha + \beta) + 0 = 0 + (\beta + \alpha) + 0;$$

3.1.11 continued

or

$\alpha + \beta = \beta + \alpha$.

Thus, we have shown that Axiom (5) is actually redundant since it may be derived as a theorem from the other axioms. In other words, had Axiom (5) been omitted, it would still be valid. Nevertheless, because Axiom (5) is so easy to accept, coupled with the fact that it is consistent with the other eight axioms, we prefer to include it as one of our axioms.

From a different perspective, what we are saying is that it is impossible to find a real model that will obey all nine of our axioms except for axiom (5). Once axioms (1), (2), (3), (4), (6), (7), (8), and (9) are obeyed, (5) must also be obeyed.

Notice also that at least in our proof, the validity of Axiom (5) as a consequence of the other axioms required that we accept Axiom (9). That is, relative to part (a) of this exercise, if Axiom (9) is omitted it is no longer clear that Axiom (5) can be derived from the remaining seven axioms.

As a closing note, we should mention that much of pure mathematics is concerned with finding the minimum number of axioms that can be used to give an equivalent definition of a structure defined by a greater number of axioms. This search involves either trying to delete some of the given axioms because they are logical derivations of the others or else it involves finding an entire new set of axioms. Again, further discussion of this point is far removed from our present investigation of vector spaces.