APPLICATIONS OF MATRICES AT SCHOOL LEVEL

T J FLETCHER

Introduction

In this talk I am mainly concerned with the applications of matrix algebra which it is possible to study with school pupils aged 15-18. It will be necessary for me to say a little about the English school system in order to compare it with the Austrian system, and I will also say a little about the matrix algebra which is, and which has been, studied in English schools by pupils aged 11-16.

In the Lehrplan for the Oberstufe in Austria we find:-

5. Klasse Lineare Gleichungen mit zwei und drei Variablen
Wiederholung einiger geometrischer Abbildungen
Vektoren der Ebene und des Raumes
[Lineare Ungleichungen in zwei Variablen
Lineare Optimierung
Lineare Abhängigkeit; Rang eines Gleichungsystems.]

The material in square brackets is only in the courses for the Realistischen Gymnasium, Naturwissenschaftlichen Realgymnasium and the Mathematischen Gymnasium. Matrices are not needed for the study of the above items, but all these items provide possible applications for matrices.

6. Klasse Additionstheoreme für die Sinus und Kosinusfunktion
[Rechnerische Behandlung geometrische Abbildungen unter Verwendung von Matrizen
Addition und Multiplikation von Matrizen.]

Again, the first item does not need matrix algebra, but it provides an application of it; indeed the book by Szirucsek et al (19) introduces matrices in connection with the problem of handling rotations and their composition.

7. Klasse Der Körper der komplexen Zahlen.

If one wishes it is possible to demonstrate to pupils that complex numbers can be seen as 2x2 matrices of a special kind.

We can conclude that the study of matrices is included only in some of the mathematics courses of the Oberstufe, and study of the textbooks shows that the main application of matrices is to transformations of the Cartesian plane.

Further, matrices can be applied to some other areas of the course, although they are certainly not essential.

The Lehrplan does not explicitly mention the inverse matrix, although it is included in some textbooks. Neither Lehrplan nor textbooks include eigenvalues and eigenvectors. This observation is relevant when we come to consider the work in some English courses.

The <u>didaktische Grundsätze</u> of the Austrian <u>Lehrplan</u> contain the following:-

Beider Einführung eines neuen Stoffgebietes sind als Ausgangspunkte nach Möglichkeit Probleme aus anderen Wissenschaften oder aus dem täglichen Leben zu wählen. Im Anschluss an motivierende Beispiele ist eine Formalisierung oder Exaktifizierung in Form der zu behandlelnden Definitionen und Lehrsätze durchzuführen . . . .

Schliesslich ist die Behandlung eines bestimmten Lehrstoffgebietes durch möglichst zahlreiche und vielfaltige Anwendungen und Übungsaufgaben abzuschliessen; . . . . Bei der Anwendungen sind die vielfaltigen Querverbindungen zwischen der Mathematik und anderen Wissenschaften, insbesondere den Naturwissenschaften, der Technik, den Wirtschafts- und Sozialwissenschaften sowie der Philosophie (mathematische Logik) aufzuzeigen.

Durch Klarstellung der dabei gemachten Annahmen, insbesondere der Vereinfachungen, sowie allenfalls durch Beispiele, bei denen derselbe Sachverhault mit Hilfe verschiendener Modelle behandelt wird, sollen die Möglichkeiten, die Schwierignkeiten und die Grenzen der Anwendbarkeit der Mathematik aufgezeigt werden.

The aim of my talk today is to show how matrices and linear algebra can be approached in this spirit. The examples which follow are my own attempts to realise these intentions - they do not illustrate general practice in England. But first I must explain the English school system.

<u>The educational system in England</u>

The national educational system is based on a partnership between central government and local authorities. The Department of Education and Science (the Ministry) does not manage schools itself, neither does it employ teachers or prescribe textbooks or syllabuses. The courses of study for the abler pupils are in fact determined closely by the examinations for the General Certificate of Education (GCE). These examinations are conducted by eight independent examination boards. They are taken at two levels - 0 level, which is generally taken at age 16, and A level, which is taken at 18. 20-25% of the population (ie the age group) take 0 level. There are examinations in mathematics for other pupils but I cannot explain this system in detail here. Historically the GCE was an examination for grammar schools. We are planning a unified system of examinations moresuited to comprehensive schools and we are at the moment in a stage of transition.

A level is an examination in separate subjects. In England it is possible to specialise at this level more than in other countries and usually a course at this level contains only three main subjects. It is therefore possible for pupils, even those going on to universities, to choose courses which contain no mathematics at all. In practice each main subject takes about one fifth of the time. Another fifth will be given to general studies (including perhaps English, religion, arts and music, games etc) and the remaining fifth to independent study.

It is possible to take mathematics as a double subject, ie for two fifths of the time. When this is done the student usually takes "Mathematics" and "Further Mathematics", although sometimes the courses are described as "Pure Mathematics" and "Applied Mathematics".

Therefore, A level pupils of the ages 16-18 may be studying,

i) no mathematics at all,
ii) single mathematics
iii) double mathematics. (Most common is Maths and Further Maths)

Note that most of these mathematics courses will contain some applied mathematics, either theoretical mechanics or statistics or both. We have a long tradition of studying mechanics as part of mathematics courses as well as part of physics courses.

The school pupils studying mathematics as a main subject at A level form about 10% of the total age group (ie those at school and those outside school), and of these only about 2 in 10 study double mathematics. You will see therefore that these courses are given only to a quite small, elite group of pupils. As previously said, examination syllabuses are not prescribed centrally. Examinations at O and A level are conducted by eight examination boards. Each school can choose its examination board, and furthermore there is usually a choice of mathematics syllabuses within each board. This means that there is a large number of possible mathematics syllabuses at this level. Universities complain that this is a source of confusion and difficulty, and moves are taking place to produce more uniform patterns.

<u>The matrix algebra contained in English O level syllabuses</u>

Mathematics courses for pupils aged 11-16 in English schools are "differentiated". That is to say pupils of different abilities spend about the same time on mathematics but some get much further than others. In this talk I am concerned principally with the work done on O level courses taken by the abler pupils, from whom come those who subsequently study mathematics to A level.

I can only simplify a complex situation, and say there are three types of course.

i) Traditional courses with no matrix algebra at all.

ii) Modern courses, of which the School Mathematics Project course is the most well known. The content of such courses is described below.

iii) The St Dunstan's course, developed in the early 1960s at an independent school in South London, carried matrix work much further than anywhere else. It is therefore an extreme case of particular interest to an international audience.

There are different editions of the SMP course, but generally speaking we might find:-

Year 1	(11+)	No matrices
Year 2	(12+)	Incidence matrices of networks. Matrices as tables of information
Year 3	(13+)	Matrix multiplication (motivated perhaps by scoring systems in games).
Year 4	(14+)	Matrices of geometric transformations, composition of transformations and matrix multiplication.
Year 5	(15+)	Revision of the above work. Inverses of 2 x 2 matrices.

Those who followed the St Dunstan's course were examined on essentially the same range of work as this, but the textbooks (13) covered a much wider range of ideas, treated in an elementary fashion. These included matrix codes, the determinant of a matrix (2 x 2 case only), group structure and groups of matrices, and work on probability matrices leading to eigenvalues and eigenvectors (although only for the 2x2 case).

Matrix algebra in A level syllabuses

I have explained that there is a variety of courses leading to A level. Because of the eight examination boards, because of single and double mathematics, and because there are syllabuses which are traditional, modern (and in-between) I must again simplify a complex situation. The amount of matrix algebra varies greatly. Three levels are typical.

i) There are traditional courses with none.

ii) In a typical single mathematics course one might find matrices in connection with linear equations, linear dependence and geometrical transformations; inverse matrices (usually no more than the 3 x 3 case); perhaps $(AB)^{-1} = B^{-1} A^{-1}$; perhaps applications to electrical circuits (SMP).

iii) Modern courses in double mathematics might include more work on vector spaces, more applications to geometry, and eigenvalues and eigenvectors (eg SMP and the new London syllabus).

When any techniques are introduced into any course it is important to consider the initial motivation, the applications, and the range of exercises which will be possible with the technique.

In the study of matrices there are two fundamental landmarks.

i) the inverse matrix,

ii) eigenvalues and eigenvectors.

Some elementary courses introduce inverses of 2x2 matrices only - and work these out by an algebraic trick which is quite untypical of later developments. This is bad pedagogy. Likewise (in my view) it is bad pedagogy to waste time learning to invert 3x3 matrices by using 2x2 minors. A more general algorithm should be studied.

With more advanced courses it is clearly a problem whether or not to go as far as eigenvectors and eigenvalues. These are the key to a rich store of applications, but the techniques appear complicated even in the 3x3 cases. I believe that if matrices are applied to practical problems (eg in probability or in mechanics) eigenvectors can be studied more easily. In some cases, such as mechanics, physical intuition indicates the eigenvectors, and the calculations required in the context of the application are often much easier than abstract examples on the general theory might suggest. Time permitting I will give examples of this later.

Applications at school level

The following applications are easily accessible in school texts in the English and German languages and I will say no more about them:-

i) geometrical applications of 2x2 matrices,
ii) applications to scheduling in business and industry,
iii) Stochastic matrices and Markov chains (5).

The first few applications discussed below show the teacher a broader, more philosophical view of matrices than is usual in school texts. We see how matrices describe relationships and how the normal definition of matrix multiplication is only one of many possible definitions - depending on the type of relationship involved. The remaining examples relate to the physical sciences (electricity, optics and mechanics) and to statistics.

Matrices describing relationships

In elementary teaching the most common approach to matrices is by way of their application to linear transformations. A matrix embodies a linear transformation. There are other approaches which have certain advantages. Matrices often describe relationships. It seems to me important for teachers to realise this aspect of matrices even though they may feel that they do not have time within the framework of these syllabuses to develop it. Certainly teachers who teach abstract exercises on matrices and who admit that they do not know any applications of these ideas need to appreciate this very general property of matrices.

The first example assumes some preliminary knowledge of matrix algebra, but I include it as a very elementary example and I do not see anything like it in school courses. It seems to me to offer possibilities for classroom discussion.

Example 1

Four people meet together, their ability to communicate in German, French and English is indicated in the following matrix M.

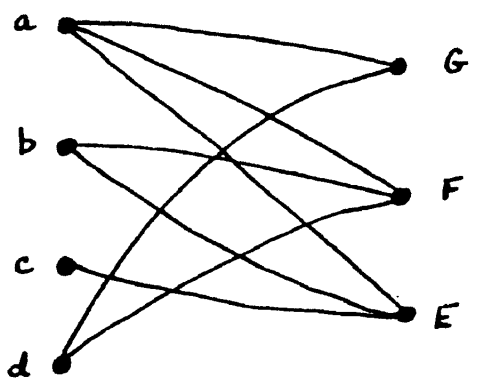
M =

	G	F	E
a	1	1	1
b	.	1	1
c	.	.	1
d	1	1	.

Form the matrix products MM^T and M^TM, and find meanings for the terms in these matrices

$$MM^T = \begin{bmatrix} 3 & 2 & 1 & 2 \\ 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 2 \end{bmatrix} \quad M^TM = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 3 \end{bmatrix}$$

The situation can also be represented (modelled) by a network

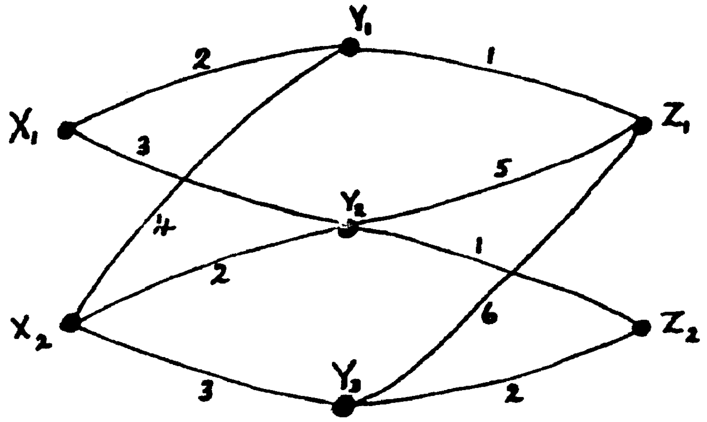


Some school syllabuses in England give a lot of space to the incidence matrices of networks, but they do not usually go on to consider many applications.

Important areas of application open up if we consider networks more generally.

Example 2

It is possible to use network problems to motivate matrix multiplication in the ordinary sense. For example we can consider an airline network.



In this diagram the numbers indicate the number of alternative ways of flying (eg different airlines) between the airports in different countries. Between country X and country Y the possibilities are indicated by the matrix.

$$A = \begin{bmatrix} 2 & 3 & - \\ 4 & 2 & 3 \end{bmatrix},$$

and between Y and Z by

$$B = \begin{bmatrix} 1 & - \\ 5 & 1 \\ 6 & 2 \end{bmatrix}$$

What is the corresponding matrix indicating the possibilities between X and Z? Direct consideration of the possibilities gives the matrix.

$$C = \begin{bmatrix} 17 & 3 \\ 32 & 48 \end{bmatrix}$$

The general law of combination is the familiar $c_{rt} = \sum a_{rs} \; b_{st}$.

But it is more instructive in this example to consider what happens if the numbers indicate other things. For example if the numbers on the links indicate the shortest times or the shortest distances the appropriate combination is

$$C = \begin{bmatrix} 3 & 4 \\ 5 & 3 \end{bmatrix}$$

$$c_{rt} = \min_s (a_{rs} + b_{st}).$$

If the numbers indicate the maximum load that a link will carry we have

$$C = \begin{bmatrix} 3 & 1 \\ 3 & 2 \end{bmatrix}$$

$$c_{rt} = \max_s \left[\min \left\{ a_{rs} \, , \, b_{st} \right\} \right].$$

There are many other laws of combination corresponding to different interpretations which can be given to the network (7,8).

I have not seen much of this done at school level although this is a splendid modelling example. Even if we do not consider these examples suitable for school pupils I think it is instructive for teachers to consider how matrix multiplication embodies very fundamental structural relationships - the series parallel properties of networks.

Example 3

As a boy at school I had to learn the complicated relationships between the traditional Imperial measures of length.

12 inches = 1 foot
3 feet = 1 yard
22 yards = 1 chain
10 chains = 1 furlong
8 furlongs = 1 mile.

As a result of this, until a few years ago the diary which the Department of Education and Science provided for my use contained the table,

	ins.	ft.	yds.	ch.	fur.	mls.
ins.	1					
ft.	12	1				
yds.	36	3	1			
ch.	792	66	22	1		
fur.	7920	660	220	10	1	
mls.	63360	5280	1760	80	8	1

This table seems to have a very complicated structure. How is it constructed? The essential conversions are contained in,

$$C = \begin{bmatrix} . & . & . & . & . & . \\ 12 & . & . & . & . & . \\ . & 3 & . & . & . & . \\ . & . & 22 & . & . & . \\ . & . & . & 10 & . & . \\ . & . & . & . & 8 & . \end{bmatrix}$$

If we call the original table M, then

$$M = I + C + C^2 + C^3 + C^4 + C^5 ,$$

all higher powers of C are zero.

In fact,

$$M = (I - C)^{-1} ,$$

and we have a Leontief inverse (3).

Examples in physics

The next few examples are all taken from physical science. Whilst some knowledge of physics is necessary in every case the essential knowledge is quite small - indeed it is usually only the appreciation of a single important principle. As some teachers may have only a small knowledge of physics we will explain what is involved each time.

It is important to ask whether the methods described below are really helpful to the science teacher or not. Science teaching has developed without the pupils (or the teachers) knowing matrix algebra, and school science always has other ways of discussing the problems which follow. However, input-output systems are important in science, and we are going to show how the same mathematical ideas apply to three quite different areas of physics.

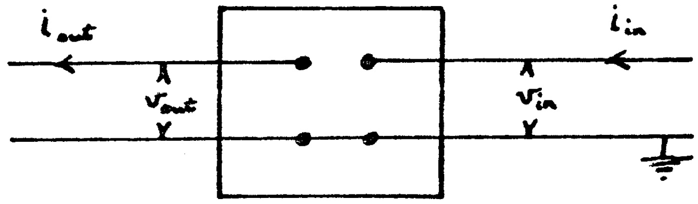
Four-terminal networks

Example 4

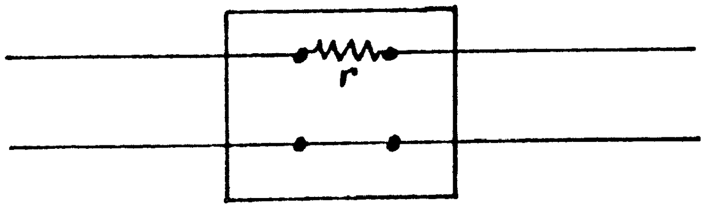
The first application is to a problem in electricity. We need the simplest knowledge of current and voltage; and we need to know Ohm's law,

voltage = current x resistance.

Imagine that we have a "black box" with two input terminals and two output terminals. An electric current flows in to the top right terminal and flows out of the top left hand terminal. The lower terminals are both earthed (ie they remain at zero voltage). We imagine that we have ways of measuring the input and output voltages and currents.



The relation between the input and output voltages and currents depends on the internal connections in the box, but for many electrical components the relation is linear. For example the box may contain a resistor r wired in as shown.



By considering the voltage charge across the resistor,

$$v_{out} = v_{in} - re_{in}.$$

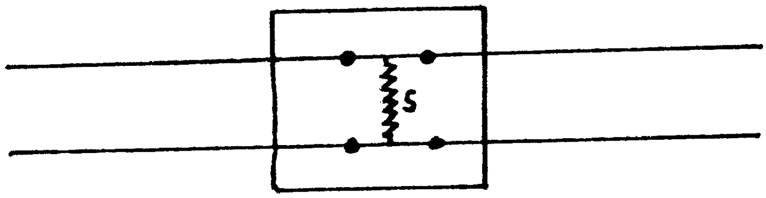
Also, current is conserved so

$$i_{out} = i_{in}.$$

These equations may be written with matrices as follows,

$$\begin{bmatrix} v_{out} \\ i_{out} \end{bmatrix} = \begin{bmatrix} 1 & -r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_{in} \\ i_{in} \end{bmatrix}.$$

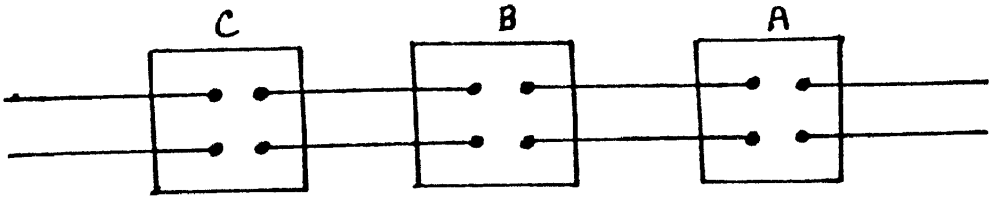
If the box is wired internally in a different manner,



then there is no change in the voltage, but a current v_{in}/s flows down through the resistor and the current flowing along the top wire is reduced by this amount. Hence the equations are,

$$\begin{bmatrix} v_{out} \\ i_{out} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1/s & 1 \end{bmatrix} \begin{bmatrix} v_{in} \\ i_{in} \end{bmatrix}.$$

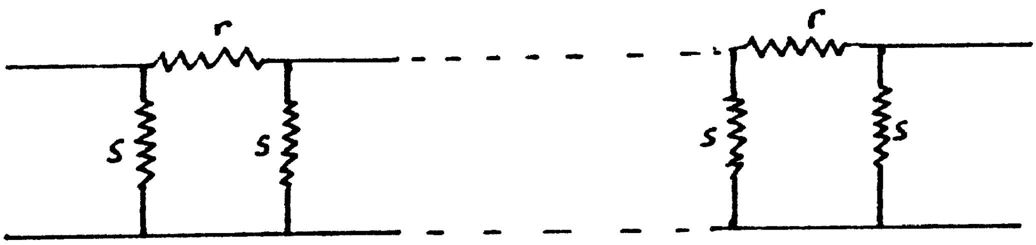
In each case the box is replaced by a matrix, and we can in fact take a series of boxes and replace each by its matrix in order to calculate the total effect. Thus a circuit,



becomes represented by the matrix equation

$$\begin{bmatrix} v_{out} \\ i_{out} \end{bmatrix} = CBA \begin{bmatrix} v_{in} \\ i_{in} \end{bmatrix}$$

The two matrices given above enable us to calculate the properties of more complicated circuits such as



This circuit will have a matrix SRS ... SRS.

By using complex numbers this theory can be extended to circuits with capacitors and inductances, and it is helpful in the study of electrical filters.

More general electrical networks can be studied using special kinds of incidence matrix. There is not space to go into the details here, but an approach at school level is contained in references (6, 16, 17).

Optical theory and matrices.

Methods similar to those used above for four-terminal networks can be applied to the theory of lenses as they are studied in school courses on physics. As a ray of light proceeds along the axis of an optical instrument its direction is changed in various ways. It is sometimes possible to describe these transformations by matrices. It is something of a problem to identify the parameters which change linearly, but it can be shown that if we make the assumptions about paraxial rays which are usually made in school physics then the parameters y (denoting the displacement of the ray from the optical axis) and θ (deonoting the inclination of the ray to the optical axis) have the required properties.

Before considering the details of the problem we must make a few remarks about notation. In the electrical filter diagrams above we drew the current flowing from right to left to match the arrangment of the matrices. In elementary mathematics we most often use matrices to operate on column vectors, and write our operations from right to left. But it is equally possible to operate on row vectors and write our matrices from left to right. The student needs eventually to become accustomed to both methods - although we may not want to introduce him to both at school.

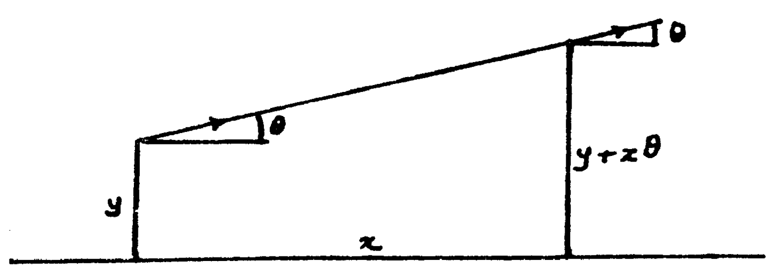
In optical diagrams light rays usually go from left to right, so here we will write our matrix operators the same way, and this entails working with row vectors.

Example 5

A ray of light proceeds through an optical instrument such as a telescope or a microscope. As it goes it is displaced in various ways which we have to study. In order to keep the problem simple we assume that the displacements are "small" - that is to say, first order approximations are sufficiently accurate.

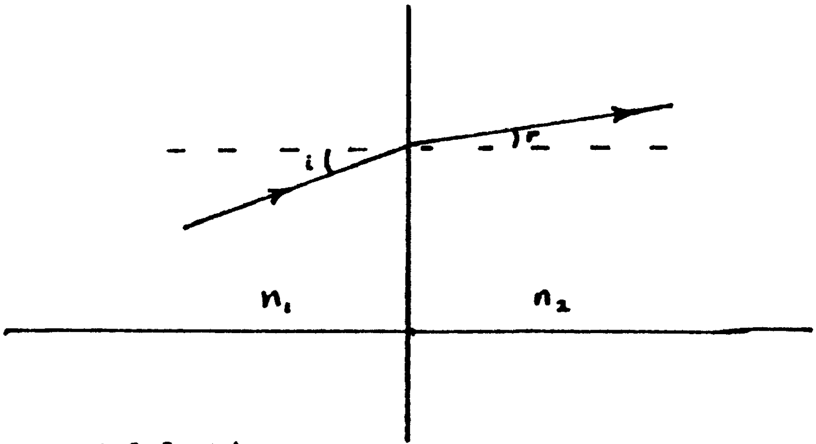
The state of the ray of light can be described by a row vector (y,θ), where y is the lateral displacement of the ray from the axis, and θ is the angle of inclination to the axis, measured in radians.

If the ray travels along the axis for a distance x then (y,θ) becomes $(y + x\theta, \theta)$. In this calculation we are using the usual approximations when θ is small. The change in the state of the ray is thus described by the matrix relation.



$$\begin{bmatrix} y & \theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ x & 1 \end{bmatrix} = \begin{bmatrix} y' & \theta' \end{bmatrix}$$

If a ray passes through the interface between two optical media of refractive indices n_1 and n_2 then a relation studied in school physics applies.



The fundamental law is

$$\frac{\sin i}{\sin r} = \frac{n_2}{n_1} ,$$

where i and r are the angles of incidence and the angle of refraction.

Making the first order approximation once again, we have this time

$$\begin{bmatrix} y & \theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & n_1/n_2 \end{bmatrix} = \begin{bmatrix} y' & \theta' \end{bmatrix} .$$

The value of y is unaltered but the ray undergoes an instantaneous change in direction at the interface.

In a similar way we can get a matrix representation of each optical feature. The output at each stage is always the input to the next stage. Associated with a curved interface between two optical media, with radius of curvature R, there is a matrix

$$\begin{bmatrix} 1 & \left(\frac{n_1}{n_2} - 1\right) \frac{1}{R} \\ 0 & n_1/n_2 \end{bmatrix} .$$

Associated with a converging lens of focal length f there is a matrix.

$$\begin{bmatrix} 1 & -1/f \\ 0 & 1 \end{bmatrix},$$

and so on.

If these examples were used in a school it would be necessary to discuss details of notation with the science teacher. In particular, various sign conventions are used in optics and it would be helpful to use the one with which the pupils were familiar.

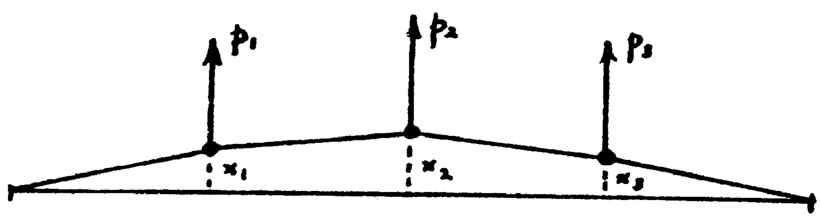
From these matrices it is possible to derive all the usual formulae at this level, and furthermore the matrix method indicates how the ideas could be applied to thick lenses, to lenses with a space between them, and to a range of more complicated problems.

Elastic displacements of mechanical systems

We will consider how matrices can describe distortions in an elastic structure such as a suspension bridge or an aircraft frame, but we will consider only very simple examples. Demonstration apparatus can be made from elastic, from thread and from wooden beads or from similar material.

Example 6

First we consider three particles, equally spaced on an elastic string. They are displaced by lateral forces. We consider lateral displacements only and ignore gravity.



There are tensions in the strings and the system is in equilibrium. The forces may be seen as inputs and the displacements of the particles as outputs. We assume that for small displacements the relations between the forces and displacements are linear. (It will be necessary for the teacher to discuss very carefully with the class what this means.)

We now consider what happens when a unit force is applied to the first particle. The displacements of the three particles are in the ratio 3:2:1, and this information can be recorded in the first column of a matrix.

$$\begin{bmatrix} 3 & \cdot & \cdot \\ 2 & \cdot & \cdot \\ 1 & \cdot & \cdot \end{bmatrix}.$$

In a similar way, if we apply a unit force to the third particle the displacements are in the ratio 1:2:3. This we can also record:

$$\begin{bmatrix} 3 & \cdot & 1 \\ 2 & \cdot & 2 \\ 1 & \cdot & 3 \end{bmatrix}.$$

Now if a unit force is applied to the second particle it is clear that the displacements are in the ratio 1:2:1, but it is not clear how the magnitudes of the displacements in this case relate to the magnitudes in the other cases.

At this point I will introduce a piece of theory which I will not explain. We can prove that matrices which describe systems of this kind are always symmetric, provided that energy is conserved in the system. (The proof of this is short and simple.) If we assume this fact we see that the middle column of our matrix must contain 2,4,2.

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

This matrix is called the flexibility matrix of the system. It is a table which records the displacements of the particles when unit forces are applied. In position (r s) of the flexibility matrix we record the displacement at point r when a unit force is applied at point s (with zero forces elsewhere).

The matrix now enables us to express a more general relationship. If forces (p_1, p_2, p_3) are applied simultaneously then the displacements are (x_1, x_2, x_3), where

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} .$$

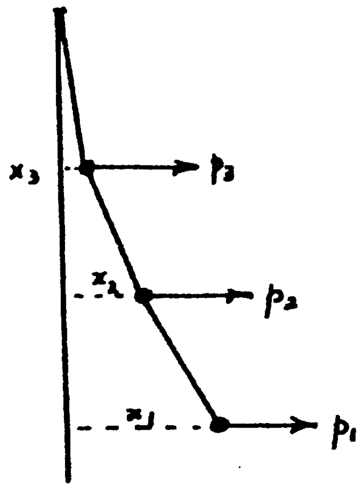
How can we calculate the forces if we observe the displacements? This is an application of the inverse matrix. What is the inverse matrix in this case?

<u>Example 7</u>

We may find the matrices which describe other systems. With three equally spaced particles on a vertical string, with gravity acting, and with horizontal perturbing forces, the flexibility matrix is

$$\begin{bmatrix} 11 & 5 & 2 \\ 5 & 5 & 2 \\ 2 & 2 & 2 \end{bmatrix} .$$

where the particles are numbered from the bottom.



What is the inverse matrix this time?

The inverse of the flexibility matrix is sometimes called the <u>stiffness matrix</u>. If the stiffness matrix is denoted by M, then we have $\underline{p} = M\underline{x}$. Note that in position (r s) of the stiffness matrix we have the force acting at the point r when there is a unit displacement at point s (with zero displacements elsewhere).

The importance of these matrices is not so much to the static problems we have considered as to dynamic problems. If systems of this kind are given a shock then they continue to vibrate under their own internal forces (resonance);

they have natural modes of vibration at certain resonant frequencies. This can be a dangerous condition. If students are familiar with simple harmonic motion in physics then we can explain how structures resonate.

A particle performs a simple harmonic motion when the restoring force is proportional to its displacement. Therefore we must ask - under what conditions is the restoring force on each particle proportional to the particle's displacement? Algebraically this condition is $\underline{p} = \lambda \underline{x}$.

But we have seen that the force $\underline{p}$ is always given by $\underline{p} = M\underline{x}$, so when these natural modes of oscillation occur

$$M\underline{x} = \lambda \underline{x},$$

[The teacher is warned that certain care is needed with this argument in the classroom. With beginners it will be necessary to distinguish carefully between perturbing forces and restoring forces, and to overcome any difficulties which the students have with the algebraic signs.]

The physical problem leads to the equation which defines eigenvectors and eigenvalues. These ideas are motivated by this physical application, and there are good historical and good didactical arguments for approaching them in this way. There are many interesting examples in which physical intuition indicates eigenvectors more easily than algebraic calculation does - and so these make good teaching exercises (9). See also references (10, 11).

<u>Matrices in statistics</u>

The Austrian didaktische Grundsätze refer to the value of examples taken from other areas of the school curriculum. The problem of drawing a straight line to give the best fit to a set of data points occurs in many sciences and of course in statistics. This can be studied at school level.

First we consider the familiar problem of finding the arithmetic mean of a set of numbers; but we use an unexpected method. For this sample problem the method may seem unnecessarily complicated, but the method is very useful with more difficult questions.

Example 8

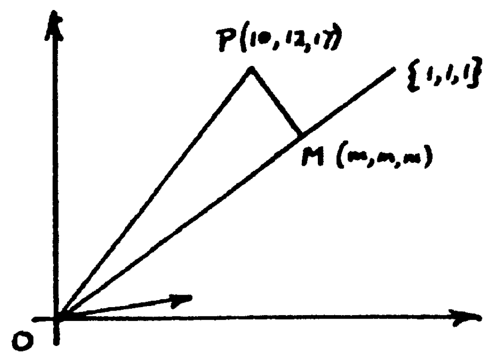
Given the numbers 10, 12, 17, which single number m approximates to them "best"? (We have to define the meaning of "best").

Consider the sum of the squares of the differences of the numbers from m,

$$(10-m)^2 + (12-m)^2 + (17-m)^2 .$$

We define the "best" value of m to be the value which minimises this sum of squares.

In 3-space this is the square of the distance between the point P(10, 12, 17) and the point M (m,m,m), which may be seen as a variable point on the line through 0 in the direction $\{1,1,1\}$.



M must be chosen to minimise PM.

This means choosing M so that PM is perpendicular to OM.
Hence,

$$1x(10-m) + 1x(12-m) + 1x(17-m) = 0,$$

so

$$m = (10 + 12 + 17)/3.$$

This is only the ordinary arithmetic mean, but it is the method which is interesting. With n variables it might be thought that the proof is in n dimensions. In a sense it is, but in fact it is in a two-dimensional subspace only.

It does not concern us here, but this approach is a very helpful one to the study of variance in statistics. We will merely point out that with n variables the quantity PM^2 is equal to $n\sigma^2$, where σ^2 is the variance. This enables easy geometrical proofs to be given of many of the formulae involving variance.

We go on to consider the problem of calculating the line of best fit for a set of data, by a method which could be used at school. The method will be illustrated by a problem with three points only, but the method clearly applies just as well to any number.

Example 9

Given the points (2,3), (4,5), (7,7) find the line of best fit, y = mx+c.

If the three points were collinear, we would be able to satisfy the 3 equations,

$$\begin{aligned} 2m + c &= 3 \\ 4m + c &= 5 \\ 7m + c &= 7, \end{aligned} \qquad (1)$$

or $\underline{A}\,\underline{m} = \underline{y}$,

where

$$\underline{A} = \begin{bmatrix} 2 & 1 \\ 4 & 1 \\ 7 & 1 \end{bmatrix} \qquad \underline{m} = \begin{bmatrix} m \\ c \end{bmatrix} \qquad \underline{y} = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}$$

We need to find $\underline{m}$, but these equations are inconsistent. If mx + c is used to estimate y the errors are

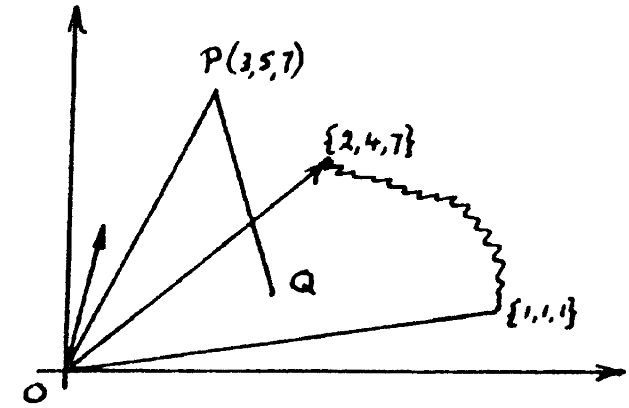
$$2m + c - 3, \quad 4m + c - 5, \quad 7m + c - 7.$$

How can we choose m and c so as to minimise the sum of the squares of the errors; ie to minimise

$$(2m + c - 3)^2 + (4m + c - 5)^2 + (7m + c - 7)^2 \quad ?$$

This is the square of the distance between P(3,5,7) and Q(2m + c, 4m + c, 7m + c) which is a variable point in the plane through 0 defined by

$\{2,4,7\}$ and $\{1,1,1\}$.



This means choosing Q so that PQ is perpendicular to $\{2,4,7\}$ and $\{1,1,1\}$; i.e. so that

$$2\,(2m + c - 3) \;+\; 4(4m + c - 5) \;+\; 7(7m + c - 7), = 0,$$

and $$(2m + c - 3) \;+\; (4m + c - 5) \;+\; (7m + c - 7), = 0.$$

This means that

$$\begin{bmatrix} 2 & 4 & 7 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2m + c - 3 \\ 4m + c - 5 \\ 7m + c - 7 \end{bmatrix} = 0,$$

or $$A^T(A\underline{m} - \underline{y}) = 0$$

or $$A^T A\underline{m} = \underline{A}^T \underline{y}. \qquad (2)$$

It is remarkable that we use A^T to convert the inconsistent equations(1) to the consistent equations (2), which provide the solution needed.

$$\underline{m} = (A^T A)^{-1}\, A^T \underline{y}.$$

This will generalise. Also it is a practical formula in school if you have access to a computer with BASIC.

At the beginning of my lecture I quoted a section of the didaktische Grundsätze in the Austrian Lehrplan for mathematics. I hope that my examples will assist you to carry out the excellent advice which this Lehrplan contains.

BIBLIOGRAPHY

1. T E Brand and A J Sherlock. Matrices: Pure and Applied, Arnold 1970
2. J R Branfield and A W Bell. Matrices and their applications, Macmillan 1970.
3. H Burger, et al. Mathematik Oberstufe 2. Holder-Pichler-Tempsky, 1979.
4. Department of Education & Science, Das Bildungswesen in England and Wales, 1973.
5. A Engel. Wahrscheinlichkeitsrechnung and Statistik. Band 2. Klett, 1976.
6. T J Fletcher. Exemplarische Übugen zur moderner Mathematik. Herder, 1967.
7. T J Fletcher. A heuristic approach to matrices. Educational Studies in Mathematics 1 166-180, 1968.
8. T J Fletcher. Combining Matrices. Mathematical Gazette L11, 379, 23-30, February 1968.
9. T J Fletcher. Matrices and molecular vibrations. Int.J.Math.Educ.Sci. Tech.2 129-142, 1971.
10. T J Fletcher. Linear Algebra through its Applications. Van Nostrand, 1972.
11. G Herzburg. Infrared and Ramam spectra of Polyatomic Molecules. Van Nostrand, 1944-6.
12. J Laub, et al. Lehrbuch der Mathematik. 2 Band. Holder-Pichler-Tempsky, 1979.
13. G Matthews. Matrices, 2 vols. Arnold, 1964.
14. School Mathematics Project. Books A-H, X-Z, Cambridge, 1968-74. (In practice, a course for comprehensive schools, ages 11-16).
15. School Mathematics Project. Book 1-5, Cambridge 1965-69. (An older course, in practice for grammar schools, ages 11-16).
16. School Mathematics Project. Advanced Mathematics, 4 vols. Cambridge, 1967-8. (Single Mathematics for A level)
17. Schools Mathematics Project. Revised Advanced Mathematics. 3 vols. Cambridge, 1973.
18. School Mathematics Project. Linear Algebra and Geometry. Cambridge, 1970. (A supplementary book for double mathematics A level).
19. E Szirucsek, et al. Arbeitslehrbuch Mathematik 6/2. Ueberreuter, 1980.

T J FLETCHER
1 August 1981