APPENDIX I

MATRIX ALGEBRA

For those with inadequate knowledge of the fundamentals of matrix algebra, a brief introduction is given here.

Definition of a Matrix

In the most general sense a matrix is a rectangular array of numbers, or symbols for numbers, which may be combined with other such arrays according to certain rules. When a matrix is written out in full, it has an appearance of which the following is typical:

$$\begin{bmatrix} 4 & -7 & 6 & 0 \\ 2 & 9 & -1 & -8 \\ 2 & 0 & 5 & 4 \\ -8 & 7 & 0 & -3 \\ 6 & 3 & -4 & 7 \end{bmatrix}$$

Note the use of square brackets to enclose the array; this is a conventional way of indicating that the array is to be regarded as a matrix (instead of, perhaps, as a determinant).

In order to discuss matrices in a general way, certain general symbols are commonly used. Thus we may write a symbol for an entire matrix as a script letter, for example, $\mathscr{A}$, which stands for

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ & & & \vdots & \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

We may also represent this matrix by $[a_{ij}]$. The vertical sets are called *columns*, and the horizontal ones *rows*. The symbol a_{ij} represents that element of the matrix $\mathscr{A}$ that stands in the ith *row* and the jth *column*. The m and n tell us the order of the matrix; m gives the number of rows and n the number of columns. A matrix in which $m = n$ is called a *square matrix* and will be of special importance to us. The elements in the set a_{ij} with $i = j$, that is, a_{11}, a_{22}, a_{33}, and so on, in a square matrix are called the *diagonal elements* because they lie entirely on the line running diagonally from upper left to lower right corners. A square matrix in which all of the diagonal elements are equal to 1 and all of the other elements are equal to 0 is called a *unit matrix* and conventionally represented by the symbol $\mathscr{E}$.

A type of matrix that is of considerable importance is the one-column matrix. To have the convenience of writing such a matrix all on one line, it is sometimes written out horizontally but enclosed in braces, { }, so as to distinguish it from a one-row matrix, which is normally written on one line in square brackets. The chief significance of the column matrix, at least for our purposes, is that it affords a way of representing a vector. Indeed it is sometimes actually called a vector.

Let us consider a vector in ordinary three-dimensional space. We can specify the length and direction of this vector in the following way. We arrange to have one end of the vector lie at the origin of a Cartesian coordinate system. The other end is then at a point which may be specified by its three Cartesian coordinates, x, y, z. In fact, these three coordinates completely specify the vector itself provided it is understood that one end of the vector is at the origin of the coordinate system. We can then write these three coordinates as a column matrix, in this case one with three rows, $\{x \quad y \quad z\}$, and say that the matrix represents the vector in question.

Obviously this notation can easily be generalized for vectors in abstract spaces of any dimension. In p-dimensional space a vector can be specified by a column vector of order $(p \times 1)$. The geometrical significance of the elements of this vector matrix is the same as in real space: They give the orthogonal (Cartesian in a general sense) coordinates of one end of the vector if the other end is at the origin of the coordinate system.

It should be noted that each of the coordinates of the outer terminus of the vector is numerically equal to the length of a projection of this vector on the axis concerned. Thus the set of numbers which define the vector in the sense discussed above may also be thought of as defining it in the sense of specifying its projections on a set of p orthogonal axes in the p-dimensional space in which it exists.

Combination of Matrices

There are certain rules for adding, subtracting, and dividing matrices; these are the rules of *matrix algebra*. It should be noted first that two matrices are equal only if they are identical. If $\mathscr{A} = \mathscr{B}$, then $a_{ij} = b_{ij}$ for all i and j.

To add or subtract two matrices, say $\mathscr{A}$ and $\mathscr{B}$, to give a sum or difference $\mathscr{C}$, the three matrices must be of the same dimensions. The elements of $\mathscr{C}$ are given by

$$c_{pq} = a_{pq} \pm b_{pq}$$

A matrix may be multiplied by a scalar number or by another matrix. For multiplication of a matrix $[c_{ij}]$ by a scalar, α, we have

$$\alpha[c_{ij}] = [\alpha c_{ij}] = [c_{ij}\alpha] = [c_{ij}]\alpha$$

Multiplication of a matrix by a matrix is somewhat more complicated. In the first place, it can be done only if the two matrices are *conformable*. This means that, if we wish to take the product $\mathscr{A}\mathscr{B} = \mathscr{C}$, the number of columns in $\mathscr{A}$ must be equal to the number of rows in $\mathscr{B}$. If this requirement is satisfied, so that $\mathscr{A}$ is of order $(n \times h)$ while $\mathscr{B}$ is of order $(h \times m)$, then $\mathscr{C}$ will be of order $(n \times m)$. Each element of the product matrix is given by the following expression:

$$c_{il} = \sum_k a_{ik} b_{kl} \tag{AI-1}$$

This sum may be written out explicitly as follows:

$$c_{il} = a_{i1}b_{1l} + a_{i2}b_{2l} + a_{i3}b_{3l} + a_{i4}b_{4l} + \cdots + a_{ih}b_{hl}$$

where a_{ih} is the last element in the ith row of $\mathscr{A}$, and b_{hl} is the last element in the lth column of $\mathscr{B}$. Perhaps this will be still clearer if we explicitly write out the process of multiplying a 3×2 matrix into a 2×4 matrix.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \end{bmatrix}$$

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} \qquad c_{21} = a_{21}b_{11} + a_{22}b_{21}$$

$$c_{12} = a_{11}b_{12} + a_{12}b_{22} \qquad c_{22} = a_{21}b_{12} + a_{22}b_{22}$$

$$c_{13} = a_{11}b_{13} + a_{12}b_{23} \qquad c_{23} = a_{21}b_{13} + a_{22}b_{23}$$

$$c_{14} = a_{11}b_{14} + a_{12}b_{24} \qquad c_{24} = a_{21}b_{14} + a_{22}b_{24}$$

$$c_{31} = a_{31}b_{11} + a_{32}b_{21}$$

$$c_{32} = a_{31}b_{12} + a_{32}b_{22}$$

$$c_{33} = a_{31}b_{13} + a_{32}b_{23}$$

$$c_{34} = a_{31}b_{14} + a_{32}b_{24}$$

A mnemonically helpful way of summarizing the process is to say that the ijth element of the product is obtained by taking the ith row of the first matrix into the jth column of the second, with emphasis on the "row-into-column" aspect. From this discussion of the process of multiplication, the conformability requirement is readily obvious. If a row of matrix $\mathscr{A}$ is to be multiplied into a column of $\mathscr{B}$, then clearly the number of elements in that row, which is the number of columns in the matrix $\mathscr{A}$, must be equal to the number of elements in a column of $\mathscr{B}$, which is the number of rows in the matrix $\mathscr{B}$.

It should be noted specifically that matrix multiplication is not in general commutative. If the matrices $\mathscr{A}$ and $\mathscr{B}$ are conformable in the sense $\mathscr{AB}$, they need not necessarily be conformable in the sense $\mathscr{BA}$. They will always be conformable in both ways when both are square and of the same order. But even when the conformability requirement is satisfied, commutation is not in general possible. For example, consider the following two products:

$$\begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 6 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 3 & 5 \end{bmatrix}$$

Matrix multiplication does, however, always obey the associative law. This can easily be proved by extension of AI-1, and working through this proof is a recommended exercise.

The quotient $\mathscr{A}/\mathscr{B}$ may be equally well regarded as the product $\mathscr{AB}^{-1}$, that is, as $\mathscr{A}$ multiplied into the inverse of $\mathscr{B}$. We thus reduce the question of how to carry out a division to the question of how to find an inverse. In order to find the inverse of a matrix certain properties of the corresponding determinant must be used. The subject is treated in detail later.

The product of a matrix and its inverse is commutative and equals a unit matrix

$$\mathscr{Q}\mathscr{Q}^{-1} = \mathscr{Q}^{-1}\mathscr{Q} = \mathscr{E}$$

Use of Matrices to Express Geometric Transformations

In Chapter 4 this use of matrices will be explored in great detail, but prior to that (Section 3.10) a few applications to coordinate transformations are given. We give here the rudiments necessary. We may designate any point in Cartesian space by a column matrix that gives its coordinates, x, y, z, namely,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Suppose we wish to reflect this point x, y, z through the origin. Its coordinates will now be $-x, -y, -z$. We may express this by the following matrix equation

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x \\ -y \\ -z \end{bmatrix}$$

If we want to rotate the point around the z axis by $\pi/2$, or reflect it through the xy plane we can write

$$\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} -x \\ -y \\ -z \end{bmatrix} = \begin{bmatrix} y \\ -x \\ -z \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} y \\ -x \\ -z \end{bmatrix} = \begin{bmatrix} y \\ -x \\ z \end{bmatrix}$$

If we want to do all of these things, one after the other, we may obtain the matrix to do it by multiplying the individual matrices to get one matrix that does it all. Since matrix multiplication is associative it does not matter how we choose the pairs. Thus, we may write

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

We can now apply this to the original point and get, as before,

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ -x \\ z \end{bmatrix}$$

Evaluation and Expansion of Determinants

To find the inverse of a matrix we shall have to employ the corresponding determinant. Determinants are, by definition, square. They consist of a square above, that $p > r$, and that x transpositions are required to put the row indices in serial order. Then, if exactly the same x transpositions are carried out in the term

$$a_{l1}a_{m2}a_{n3} \cdots a_{ri}a_{pj} \cdots$$

it will still be necessary to make an additional $2(p - r - 1) + 1$ transpositions to put a_{ri} and a_{pj} in their proper places, making $x + 2(p - r - 1) + 1$ transpositions in all. Thus, if x is even, $x + 2(p - r - 1) + 1$ must be odd and vice versa. It therefore follows that all the terms in the expansion will cancel out in a pairwise fashion. Obviously a similar argument could be made if we assume two columns to be identical.

The Adjoint Matrix

Before defining the adjoint matrix we must define the transpose of a matrix. This is a matrix of which the columns are the rows, and vice versa, of the original matrix. Symbolically, the transpose of the matrix $[a_{ij}]$ is $[a_{ji}]$. Now, the adjoint matrix of a matrix $[a_{ij}]$ is defined as follows:

$$\text{Adjoint of } [a_{ij}] = [A^{ji}]$$

That is, we treat the array of elements constituting $[a_{ij}]$ as a determinant, write the cofactor A^{ij} of each element in place of the element giving the matrix $[A^{ij}]$, and then make the transpose of $[A^{ij}]$. The matrix adjoint to $\mathscr{A}$ will be symbolized $\hat{\mathscr{A}}$.

THE INVERSE OF A MATRIX

The inverse $\mathscr{A}^{-1}$ of a matrix $\mathscr{A}$ is, by definition, such that

$$\mathscr{A}\mathscr{A}^{-1} = \mathscr{A}^{-1}\mathscr{A} = \mathscr{E}$$

Let us now look at the product $\mathscr{A}\hat{\mathscr{A}}$ for a square matrix of order 3:

$$\mathscr{A}\hat{\mathscr{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} A^{11} & A^{21} & A^{31} \\ A^{12} & A^{22} & A^{32} \\ A^{13} & A^{23} & A^{33} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}A^{11} + a_{12}A^{12} + a_{13}A^{13} & a_{11}A^{21} + a_{12}A^{22} + a_{13}A^{23} & a_{11}A^{31} + a_{12}A^{32} + a_{13}A^{33} \\ a_{21}A^{11} + a_{22}A^{12} + a_{23}A^{13} & a_{21}A^{21} + a_{22}A^{22} + a_{23}A^{23} & a_{21}A^{31} + a_{22}A^{32} + a_{23}A^{33} \\ a_{31}A^{11} + a_{32}A^{12} + a_{33}A^{13} & a_{31}A^{21} + a_{32}A^{22} + a_{33}A^{23} & a_{31}A^{31} + a_{32}A^{32} + a_{33}A^{33} \end{bmatrix}$$

We see that each diagonal element is the expansion of the determinant $|A|$ in terms of a row and its cofactors. On the other hand, each off-diagonal

element is the sum of products of the elements of a certain row, say the ith, with the cofactors of the elements of some other row, say the jth. Such a sum is, in fact, the expansion in the elements of the ith row with their cofactors of a determinant in which the ith and jth rows are identical. Since we have already seen that the value of such a determinant must be zero, all off-diagonal elements of the product $\mathscr{A}\hat{\mathscr{A}}$ are zero. It is also easy to see that $\mathscr{A}\hat{\mathscr{A}} = \hat{\mathscr{A}}\mathscr{A}$.

Thus we have the result

$$\mathscr{A}\hat{\mathscr{A}} = \hat{\mathscr{A}}\mathscr{A} = \begin{bmatrix} |A| & 0 & 0 & 0 & \cdots & 0 \\ 0 & |A| & 0 & 0 & \cdots & 0 \\ & & & & \vdots & \\ 0 & & & & \cdots & |A| \end{bmatrix}$$

$$= |A| \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ & & & & \vdots & \\ 0 & & & & \cdots & 1 \end{bmatrix}$$

$$= |A|\,\mathscr{E}$$

Now, referring to the definition of $\mathscr{A}^{-1}$, we see that

$$\mathscr{A}^{-1} = \frac{\hat{\mathscr{A}}}{|A|}$$

That is, each element of $\mathscr{A}^{-1}$ is the element of $\hat{\mathscr{A}}$ divided by $|A|$. Since division by zero is not defined, only matrices for which the corresponding determinants are nonzero can have inverses. A matrix $\mathscr{A}$ such that $|A| = 0$ is said to be *singular* (no inverse), whereas matrices of which the corresponding determinants are nonzero are said to be *nonsingular*. Only nonsingular matrices can occur in representations of a group. It is also clear that since only square matrices can have corresponding determinants, only square matrices can have inverses.