Physics 200-04
Matrices

Matrices will be used time and time again during the course of this course. Both special relativity and quantum mechanics need them, and are far more easily expressed in terms of them. It is ironic that matrices were not taught in school, either high school or university, to physicists in the first quarter of the 20th century. They were considered a highly esoteric mathematical subject of little relevance to physics. Heisenberg, the first developer of quantum mechanics knew nothing about them, and had to be told by Max Born that what he had invented was a form of matrix mechanics. The concept is actually very simple, and we will spend half a lecture on them. They are also very powerful, and thinking about matrices algebraically allows one to prove powerful theorems simply.

However their algebra (manipulation) is also counter-intuitive, and they take some getting used to, so do not be discouraged if at first they seem strange.

Matrices

A matrix is a large n by m array of numbers, variables, symbols, whatever. In A is a matrix, we designate the element in the ith row and jth column by A_{ij}.

If two matrices, A and B have the same number or rows and columns, then the matrix $C = A + B$ is defined by having C have the same number of rows and columns as both A and B have, and any element of C is just the sum of the elements with the same row and column number in A and B. Ie, for any row number I and column number j, the sum is defined by

$$C_{ij} = A_{ij} + B_{ij} \tag{1}$$

We can also multiply Any matrix by a constant. If α is some constant, then the matrix αC (ie α times C) is just the matrix obtained by multiplying each element of C by α.

$$(\alpha C)_{ij} = \alpha C_{ij} \tag{2}$$

Finally we can multiply a matrix by another matrix, but only if the second has the same number of columns as the first.

Thus the matrix AB, the product of the matrices A and B is defined only if A has the same number of columns as B has rows. This product is defined by

$$(AB)_{ij} = \sum_k A_{ik} B_{kj} \tag{3}$$

The sum is over all of the columns of A and thus rows of B. Note that each element in the product is a sum over the products of the elements with the same row number in A with the elements with the same column number in B.

It is extremely important to remember that in general the Product AB of two matrices is **not** the same as the product BA. You have spent about 12 years of your life learning and using the fact that in the multiplication of numbers or of variables, the order in which you multiply them does not matter. You now have to unlearn this. The order is crucial with matrices. If the matrices are not square (have different number of rows and columns) the product in one order may not even be defined. But even for square matrices, the order in general matters. Thus for example, if A and B are matrices, ABA is **not** the same as A^2B, where A^2 is of course AA.

Thus as an example,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -1 & 1 & 3 \\ 0 & 1 & 2 \end{pmatrix} \tag{4}$$

$$= \begin{pmatrix} 1(-1) + 2 \cdot 0 & 1 \cdot 1 + 2 \cdot 1 & 1 \cdot 3 + 2 \cdot 2 \\ 3 \cdot (-1) + 4 \cdot 0 & 3 \cdot 1 + 4 \cdot 1 & 3 \cdot 3 + 4 \cdot 2 \end{pmatrix} \tag{5}$$

$$= \begin{pmatrix} -1 & 3 & 7 \\ -3 & 7 & 17 \end{pmatrix} \tag{6}$$

While matrix multiplication is not commutative (ie the order of multiplication matters), it is transitive. Thus it is easily shown that if you have three matrices with the appropriate dimensions, that

$$A(BC) = (AB)C \tag{7}$$

Ie, it does not matter if you multiply B and C first and then multiply by A or multiply A and B first and then by C. However the order (ie C on the left and A on the right) must be preserved.

From a matrix A one can always make a new matrix, called the transpose, A^T. This is defined by defining A^T elements as the ones in A with rows and columns interchanged. Thus

$$(A^T)_{ij} = A_{ji} \tag{8}$$

Note that

$$(AB)^T = B^T A^T \tag{9}$$

(Proof:

$$(AB)_{ij} = \sum_k A_{ik} B_{kj} \tag{10}$$

so

$$((AB)^T)_{ij} = (AB)_{ji} = \sum_k A_{jk} B_{ki} \tag{11}$$

And

$$(B^T A^T)_{ij} = \sum_k B^T_{ik} A^T_{kj} = \sum_k B_{ki} A_{jk} = \sum_k A_{jk} B_{ki} \tag{12}$$

(Note that last step in interchanging the $B_{ki}A_{jk}$ term is legal since those are just ordinary numbers for which the order of multiplication does not matter) Note that for any matrix, one can always multiply A by its transpose, since A^T has the same number of columns as A has rows and the same number of rows as A has columns.

One very useful matrix is the identity matrix. This is always a square matrix (ie has the same number of rows and columns.) All of its elements are zero except the diagonal ones, which are 1. If I is the identity matrix, then

$$\begin{aligned} I_{ij} &= 0 \qquad if\ i \neq j \\ I_{ij} &= 1 \qquad if\ i = j \end{aligned} \tag{13}$$

Its primary feature is that the product of the identity matrix (with the appropriate dimensions) with any other matrix gives that other matrix back. Eg, if A is an nxm matrix, and I is an nxn matrix, then

$$IA = A \tag{14}$$

Ie, it acts as far as multiplication of matrices is concerned, as the identity element (as the number 1 acts as far as ordinary multiplication). Note that if A is not a square matrix, then the identity multiplying from the left is a different matrix from the identity multiplying from the right.

Finally for square matrices, (eg, A an nxn matrix) there is the concept of an inverse. The matrix A^{-1}, the multiplicative inverse of A (note that this is **never** (well hardly ever) written as $\frac{1}{A}$) is the matrix which when multiplied by A gives the identity matrix.

$$A^{-1}A = I \tag{15}$$

where I has the same dimensions as does A.

It is easy to show that the left and right inverses must be the same. Let us say that the matrix B obeys

$$AB = I \tag{16}$$

Multiply both sides from the left by A^{-1}. We get

$$A^{-1}AB = IB = B \tag{17}$$

on the left hand side and

$$A^{-1}I = A^{-1} \tag{18}$$

on the right hand side which give us $B = A^{-1}$.

Rotations

Define the "vector" (an 3x1 matrix) X by

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{19}$$

We can define the length of the vector X by

$$length^2(X) = |X| = X^T X = x^2 + y^2 + z^2 \tag{20}$$

Now, consider the general linear transformation,

$$X' = UX \tag{21}$$

where U is a 3x3 matrix. This makes the three coordinates $x'y'z'$ to be general linear functions of xyz. What properties must U have so that the vector X' has the same length as X for arbitrary X? The answer is easy. The length squared of X' is

$$length^2(X') = X'^T X' = (UX)^T UX = X^T U^T UX \tag{22}$$

If this is always supposed to equal the length of X, then we can do that if

$$U^T U = I \tag{23}$$

since then

$$X^T U^T UX = X^T IX = X^T X = length^2(X) \tag{24}$$

Ie, this is a necessary and sufficient condition on the matrix U that it preserve lengths.

Thus any rotation, any transformation of the coordinates of an object, which preserves the distance of that object, can be represented by such a matrix. These are called orthogonal or rotation matrices.

Note that these matrices form what is called a group. Namely if U and V are such matrices, then so is UV the product.

$$(UV)^T UV = V^T U^T UV = V^T IV = V^T V = I \tag{25}$$

Ie, if U and V obey the orthogonality condition, then so does UV. In terms of rotations, this means that if you rotate, and then you rotate in some other way, the end result is the same as if you had just done one rotation (usually different from the either of the rotations).