Physics 139B/171. General Relativity/Quantum Mechanics.
Professor Dine

Fall, 2009. Handout: Vectors: "Everything You Need to Know"

This note is a concise summary of the things you need to know about vectors, matrices, and rotations. If you master what's here, I promise you will never be scared of a cross product or a curl again.
The exercises are intended to give you some practice. Please work them as you go along. You do not need to hand them in.

What is a vector?

The simplest definition is an object with magnitude and direction. But we want to be more precise. We define a vector by how it transforms under coordinate rotations. This is a very crucial notion in physics, where we believe the underlying laws of nature do not care about how we choose our coordinate axes, on the one hand, yet we often have to choose coordinates to write precise equations and do calculations on the other.

A vector is any object which transforms under rotations like the coordinates of a point. So how do the coordinates of a point transform? We can represent the position vector in various ways:

$$\vec{x} = x\hat{i} + y\hat{j} + z\hat{k} \tag{1}$$

or

$$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{2}$$

or as

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{3}$$

We can also just refer to the components, (x, y, z) or (x_1, x_2, x_3) or simply x_i. This way of writing the components of vectors with an index allows writing very compact, simple formulas. For example, the dot product of two vectors, $\vec{x}$ and $\vec{y}$ can be written:

$$\vec{x} \cdot \vec{y} = \sum_{i=1}^{3} x_i y_i. \tag{4}$$

Now we do something which may be new to you. Einstein realized that one could save a lot of writing if one adopted the convention, known as the "Einstein Summation Convention", that, unless otherwise stated, one sums over repeated indices. So, for example,

$$x_i y_i = \sum_{i=1}^{3} x_i y_i. \tag{5}$$

We will use this a lot in what follows. You should practice with this a few times.

To describe rotations, start with a rotation about the z axis. This has no effect on the z (x_3) coordinate, If we rotate the axes counterclockwise by an infinitesimal angle θ, the coordinates of the point (x, y) in the new coordinate system are:

$$x' = x\cos(\theta) + y\sin(\theta) \qquad y' = -x\sin(\theta) + y\cos(\theta). \tag{6}$$

Exercise: (a.) Draw a picture and check this formula. Check that the length of the vector is the same in both coordinate systems, as it should be.

We can write this in terms of a matrix. Call

$$O = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{7}$$

Then, *using the summation convention*:

$$x'_i = O_{ij}x_j. \tag{8}$$

The numbers, O_{ij} are the elements of the matrix O; $O_{11} = \cos(\theta)$, $O_{12} = -\sin(\theta)$, etc. Note that i is not repeated on either side of the equation, so there is no sum. So it is really three equations, one for $i = 1$, one for $i = 2$, etc.

Exercise: (b.) Write this equation out with the summation symbol. Write out the three equations, and verify that this is just the usual rule for multiplying a vector by a matrix.

So our compact index notation has given us a way of representing rotations of coordinates. We considered a very special rotation, but you can convince yourself that any rotation, however complicated, can be written as a product of rotation matrices like this. In other words, rotate first about the z axis; then rotate about the *new* x axis (say), and so on. Your book shows you a simple way to write the general rotation matrix. We won't need it here, but if someday you are writing code for an experiment or for graphics software, you might need to do it (or at least use a program which does).

Exercise: (c.) Write the rotation matrix for a rotation by 45 degrees about the x axis.

Exercise: (d.) Show that the matrices for these two rotations are orthogonal, $O^TO = 1$. Show that the product of any two orthogonal matrices is an orthogonal matrix; from this argue that any rotation matrix is an orthogonal matrix.

Now a vector is any quantity which transforms like the coordinates under a rotation. In general, a vector will have x, y and z coordinates, and we can indicate it by $\vec{A}$, $A_x\hat{\imath} + A_y\hat{\jmath} + A_z\hat{k}$, or A_i. In mechanics, there are really only a few vectors which will interest us. some familiar ones are:

$$\vec{v} = \frac{d\vec{x}}{dt} \qquad \vec{p} = m\vec{v} \qquad \vec{a} = \frac{d^2\vec{x}}{dt^2}. \tag{9}$$

Exercise: (e.) Show that $\vec{v}$ is a vector, i.e. if $\vec{x}' = O\vec{x}$, $\vec{v}' = O\vec{v}$. Work with the index notation and the summation convention.

The Kronecker delta symbol

This is a very simple but useful device:

$$\delta_{ij} = 1 \text{ if } i = j, 0 \text{ otherwise}$$

Note that we can write the *elements* of the unit matrix as:

$$I_{ij} = \delta_{ij}. \tag{10}$$

To see how useful it is, let's first use the summation convention to summarize in a concise way the rules of matrix multiplication. If we take the product of two matrices, we can write it as:

$$(AB)_{ik} = A_{ij}B_{jk} \tag{11}$$

Exercise: (f.) Check this for 2×2 matrices.

For an orthogonal matrix, using the Kronecker delta, we have

$$O^T_{ik}O_{kj} = \delta_{ij}$$

.

Now consider what happens to the dot product of any two vectors under rotations.

$$\vec{A}' \cdot \vec{B}' = A^T O^T O B = A^T B = \vec{A} \cdot \vec{B} \tag{12}$$

or with our index notation:

$$A'_i B'_i = O_{ij}A_j O_{ik}B_k = A_j O^T_{ji} O_{ik} B_k = A_j B_j \tag{13}$$

Exercise: (g.) Justify each step; use the Kronecker delta symbol at the last stage.

The Cross Product

Levi-Civita symbol:

$$\epsilon_{ijk} = 1 \text{ if } i = 1, j = 2, k = 3; \text{otherwise by antisymmetry.}$$

So

$$\epsilon_{231} = \epsilon_{312} = 1; \epsilon_{132} = \epsilon_{321} = -1; \epsilon_{112} = \epsilon_{122} = \ldots = 0.$$

Exercise: (h.) Evaluate

$$\delta_{ij}x_j$$

(remember to use the summation convention)

Using the ϵ symbol and the summation convention, we can give a nice, alternative description of the cross product.

$$(\vec{A} \times \vec{B})_i = \epsilon_{ijk}A_j B_k.$$

Remember, with the summation convention this really means:

$$(\vec{A} \times \vec{B})_i = \sum_{j,k=1}^{3} \epsilon_{ijk}A_j B_k.$$

Exercise: (i.) Check the expression above for the cross product.

In this form, one can often manipulate cross products more easily than with the determinant notation. One needs just one more ingredient. This is an identity for the product of two ϵ tensors. Once one has this, all of the mysterious identities you find in textbooks are at your fingertips:

$$\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{ij}\delta_{jl}.$$

You can check this the hard way by doing the sums, or you can be more clever, and make some symmetry arguments. Or you can just believe me

With this, for example, we can derive the famous identity:

$$\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A} \cdot \vec{C}) - \vec{C}(\vec{A} \cdot \vec{B})$$

Start with

$$\begin{aligned}
[\vec{A} \times (\vec{B} \times \vec{C})]_i &= \epsilon_{ijk} A_j (\vec{B} \times \vec{C})_k \\
&= \epsilon_{ijk} A_j \epsilon_{klm} B_l C_m \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) A_j B_l C_m \\
&= B_i (A_j C_j) - C_i (B_j A_j)
\end{aligned}$$

This is, for the components of the vectors, just the identity!

Exercise: (j.) Repeat the derivation, putting in the summation signs everywhere explicitly.

The gradient

The gradient is a vector **operator**. If it acts on a function, it produces a vector function.

$$\vec{\nabla} = \hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}$$

Just like $\vec{x}$, we can represent the components of $\vec{\nabla}$ in terms of objects with indices: $\frac{\partial}{\partial x_i}$. Again, this can be written in a convenient shorthand: ∂_i.

With this, one can construct three interesting quantities:

1. The gradient of a function:

$$\vec{\nabla} f(\vec{x}) = \hat{i}\frac{\partial f}{\partial x} + \hat{j}\frac{\partial f}{\partial y} + \hat{k}\frac{\partial f}{\partial z} = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z})$$

   Note that this is a vector. Again, we can write it more simply using the index notation:

$$\frac{\partial f}{\partial x_i} = \partial_i f.$$

2. The divergence:

$$\begin{aligned}
\vec{\nabla} \cdot \vec{V} &= \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z} \\
&= \frac{\partial V_i}{\partial x_i} = \partial_i V_i.
\end{aligned}$$

   Note that this is a scalar.

3. The curl:

$$\vec{\nabla} \times \vec{V} = (\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}, \frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}, \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y})$$

   Note that this is a (pseudo) vector

$$(\vec{\nabla} \times \vec{V})_i = \epsilon_{ijk} \partial_j V_k \tag{14}$$

The directional derivative, or the change in a function in a particular direction, $\hat{n}$, is obtained from $\hat{n} \cdot \vec{\nabla} f$. Note that

$$df = \vec{dx} \cdot \vec{\nabla} f = |\vec{dx}| \hat{n} \cdot \vec{\nabla} f.$$

The gradient of a function points in the direction the which the function changes most rapidly. The differential of a function is related to the gradient. Many forces can be written as the gradient of potentials (the condition to do this involves the curl: $\vec{\nabla} \times \vec{F} = 0$). The same statement applies to the electrostatic potential.

Exercise: (k.) Show that the electric field due to a charge, q, at the origin can be written as:

$$\vec{E} = -\vec{\nabla}\phi$$

where

$$\phi(\vec{x}) = \frac{q}{\epsilon_o |\vec{x}|}.$$

Check that $\vec{\nabla} \times \vec{E} = 0$. If you want to do this first using the x, y and z components separately, that's fine, but then do the exercise with the index notation.

Finally, there is one more operation which can be performed with the gradient operator. By taking the dot product of $\vec{\nabla}$ with itself, we obtain the Laplacian:

$$\nabla^2 = \vec{\nabla} \cdot \vec{\nabla} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

This *operator* is a scalar. If it acts on a scalar, it gives another scalar. If it acts on a vector, it gives a vector.
Exercise: (l.) Show that the scalar potential satisfies,

$$\nabla^2 \phi = 0$$

if $r = |\vec{x}| \neq 0$.

Infinitesimal Rotations

Finally, we return to rotations. Suppose that θ is very small, $\theta = \delta$. Then

$$O \approx \begin{pmatrix} 1 & -\delta & 0 \\ \delta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{15}$$

Since O is nearly a unit matrix, we can rewrite it as:

$$O \approx I + o \tag{16}$$

Note that $o = -o^T$, so that $O^T O = I + \mathcal{O}(o^2)$. Also, we can write ω in a way which generalizes to a transformation about any axis. If we rotate by an infinitesimal angle, δ about an axis $\vec{\omega}$

$$(o)_{ij} = -\delta \ \epsilon_{ijk}\omega_k. \tag{17}$$

Exercise: (m.) Check the statements in the paragraph above.

Complex Vectors, Matrices

In quantum mechanics, we are interested in complex vectors and matrices. These are simple generalizations of many of the notions above.
Unit vectors: $\hat{e}_a$. General vector:

$$V = c_a \hat{e}_a \tag{18}$$

where c_a are complex numbers, and we are again using the summation convention.
Abstractly, operators act on vectors, transforming one vector into another. More concretely, the action of operators can be represented as matrices,

$$V'_a = M_{ab} V_b. \tag{19}$$

Some special matrices:

1. Adjoint of a matrix: $M^{\dagger}_{ab} = M^{*}_{ba}$.
2. Hermitian matrix: $M = M^{\dagger}$ (important property: real expectation values, $V^{\dagger}MV$ real; real eigenvalues).
3. Unitary matrices: $U^{\dagger} = U^{-1}$. Analog of orthogonal matrices for real vectors; like rotations, describe change of basis.

All of this has a translation in Dirac's bra/ket notation.

1. $V \leftrightarrow |\phi\rangle, \langle\phi|$.
2. $MV \leftrightarrow \hat{M}|\phi\rangle$
3. $\hat{M} = \hat{M}^{\dagger} : \langle\phi|\hat{M}|\phi'\rangle = \langle\phi'|\hat{M}|\phi\rangle^{*}$.
4. $U_{i\alpha} : \langle i|\alpha\rangle; \langle i|\alpha\rangle\langle\alpha|j\rangle = \delta_{ij}$.