Lecture Notes II: Mathematical Preliminaries

Cory Schillaci

July 16, 2013

Physics and math have a complex relationship. On one hand, it's important to make sure that you understand the essential physics in a non-mathematical way. On the other, math helps make physics precise. This is especially important in modern physics, where many concepts are entirely counterintuitive and therefore hard to express precisely using English (or any other language except mathematics). I'll try to keep the math manageable and fun. The following sections describe tools that will appear often as we discuss topics in this course. If you find yourself struggling with my explanations, do a Google search. All of these topics are fairly basic and you will find many useful websites.

1 Complex Numbers

Historically, the first numbers studied by mathematicians were the natural numbers. In time, zero was introduced, fractions were studied, negative numbers filled out the rational numbers, and the real numbers came to be understood. All of these seem like natural concepts with physical interpretations. Mathematicians had introduced complex numbers as early as the 16^{th} century, but it wasn't until the 20^{th} century that physicists came to realize that these numbers were just as essential to physics as the reals.

Quick Question: Do you know what all these different sets of numbers (natural, integer, rational, real) are?

Complex numbers are numbers of the form $z = x+iy$, where $i^2 = -1$. That's basically all you need to know to get started with complex numbers. Algebraic operations follow from the normal rules, for example

$$(3+4i)^2 = 9 + 24i + 16i^2 = 9 + 24i - 16 = -7 + 24i \tag{1}$$

We often refer to the *real part*, which is just $\mathrm{Re}\,(z) = x$, or the *imaginary part*, $\mathrm{Im}\,(z) = y$. Note that the imaginary part is just the real number y, not iy!

Quick Question: Given $z = -7 + 24i$, find $\mathrm{Re}\,(z)$ and $\mathrm{Im}\,(z)$.

There is one special operation called *complex conjugation*. To find the complex conjugate of any complex number, simply replace every i with a $-i$. We notate this with an asterisk as follows:

$$z = x + iy \rightarrow z^* = x - iy \tag{2}$$

One way in which the conjugate is useful is that any number times its conjugate is a real number:

$$zz^* = (x + iy)(x - iy) = x^2 + y^2 \tag{3}$$

Quick Question: Write $\frac{3}{4+i}$ in the $x + iy$ format. Hint: Multiply by one in a clever way.

Now here is one of the most famous results of complex analysis. It turns out the the exponential with a complex argument has an interesting and very important property,

$$e^{i\theta} = \cos\theta + i\sin\theta \tag{4}$$

This is called Euler's formula. I don't know any way to prove this without calculus, so you will have to take this as given at this point. When you learn about Taylor series, ask your teacher to show you the very simple proof!

Quick Question: What is the complex conjugate of $e^{i\theta}$?

Quick Question: What is $e^{i\pi} + 1$?

2 Matrices

Matrices are collections of numbers arranged in rows and columns like this:

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \tag{5}$$

The example above is a 3×3 matrix, because it has three rows and three columns[1]. You may have learned to use matrices as a tool to solve systems of linear equations in algebra. In this class we will use them extensively in quantum mechanics for a different purpose. For now, we just want to learn how to work with matrices.

Only a few shapes of matrices show up in physics. The first is the square matrix, any matrix with an equal number of rows and columns as in (5). Physicists also use matrices with any number of rows but only a single column (often called a *column vectors*) or matrices with an number of columns but only a single row (called *row vectors*). As we will see in Section 3, it would be better not to always call these vectors, as vectors also must be associated with a specific coordinate system to be written this way. For now, just ignore that complication.

[1]In general, an $n \times m$ matrix has n rows and m columns

When referring to specific elements of a matrix, we sometimes use a special notation. If we have a matrix A, the element of A in the i-th row and j-th column is indicated by A_{ij}. For example, in the matrix of (5), $A_{13} = c$ and $A_{21} = d$.

Quick Question: What is A_{31}?

2.1 Matrix operations

There are several important types of matrix operations to learn about for now. First, the transpose operation. If you have a matrix A as in (5), you get the transpose by turning all the rows into columns and vice versa like so:

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \rightarrow A^T = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix} \tag{6}$$

In index notation, we would write this as $A^T_{ij} = A_{ji}$.

Quick Question: Find the transpose of the row vector $\begin{pmatrix} 3 & 4 \end{pmatrix}$. What is the dimension of the transpose?

The hermitian conjugate, notated with the † symbol[2], is similar to the transpose, except that if the entries of the matrix are complex we take the complex conjugate as well:

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \rightarrow A^\dagger = \begin{pmatrix} a^* & d^* & g^* \\ b^* & e^* & h^* \\ c^* & f^* & i^* \end{pmatrix} \tag{7}$$

Next is addition of matrices. To add matrices, they must be of the same size. I can add two 3×3 matrices or two 3×1 matrices, but not a 3×3 and a 3×1. To add them, add together the corresponding elements in each like this:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix} \tag{8}$$

Quick Question: Find the sum $\begin{pmatrix} 1 \\ 6 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

Quick Question: Write the addition operation in index notation.

To multiply a matrix times a number, just multiply every entry by that number:

$$n \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} na & nb \\ nc & nd \end{pmatrix} \tag{9}$$

[2] $A^\dagger$ is typically said as "A dagger"

Now for the hard one: multiplying two matrices together. Formally, we can find each entry in the product matrix using the formula

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

where k runs over all the entries (the limits depend on the size of the matrices). Let's see how this pans out in practice. Say we want to multiply the following matrices:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{11}$$

$$B = \begin{pmatrix} e & f \\ g & h \end{pmatrix} \tag{12}$$

To find the the entry in the first column and first row of the product AB, we only need the first row of A and the first column of B:

$$A \to \begin{pmatrix} a & b \\ & \end{pmatrix}, B \to \begin{pmatrix} e & \\ g & \end{pmatrix} \tag{13}$$

Now we step across A and down B, multiplying the entries at each step and then taking the sum to find the upper left element of the product, $(AB)_{11}$:

$$AB = \begin{pmatrix} ae + bg & ? \\ ? & ? \end{pmatrix} \tag{14}$$

Quick Question: Fill in the the question marks in (14).

Quick Question: Find the product AB of $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Finally, the *determinant* of a matrix is found in the following way. First, we define the determinant of a 2×2 matrix to be,

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \tag{15}$$

The determinant of a larger matrix is defined recursively. First pick a row (or column), any one will do. Let's use the first row of the following matrix, indicated by the bold letters:

$$A = \begin{pmatrix} \mathbf{a} & \mathbf{b} & \mathbf{c} \\ d & e & f \\ g & h & i \end{pmatrix} \tag{16}$$

Start from the left (top) and moving across the row (down the column), strike out the row and column containing that element. For step one, that gives us:

$$A = \begin{pmatrix} & & \\ & e & f \\ & h & i \end{pmatrix} \tag{17}$$

Multiply the element we are focusing on by the the determinant of the remaining matrix to get $a(ei-hf)$. Now we have to add a sign depending on where our element is according to the following scheme:

$$\begin{pmatrix} + & - & + & \dots \\ - & + & - & \dots \\ + & - & + & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{18}$$

So the first element gets added in directly. Now we have $\det A = a(ei-hf)+\dots$. Now move across the row and focus on A_{12}. Removing its row and column leaves:

$$A = \begin{pmatrix} d & & f \\ g & & i \end{pmatrix} \tag{19}$$

From (18) we see that this element corresponds to a minus sign, so the determinant is now $\det A = a(ei-hf) - b(di-gf) + \dots$. Iterating once more we arrive at the end of the row to discover that

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a(ei-fh) - b(di-fg) + c(dh-eg) \tag{20}$$

You can just use the result (20) if you prefer when working with 3×3 matrices, but the procedure I outlined above works recursively for any size of matrix. I also find it easier to remember the procedure than to memorize (20).

2.2 Special types of matrices

When studying quantum mechanics, we will use certain special types of matrices extensively. In no particular order, here are their definitions.

The *identity matrix* is the matrix with one on the diagonals and zero everywhere else. This matrix is often denoted I, and when multiplying with any other matrix M satisfies $IM = MI = M$.

A *unitary* matrix is any matrix M that satisfies $M^\dagger M = I$.

A matrix M is called *hermitian* if $M^\dagger = M$. In other words M is hermitian if it is it's own hermitian conjugate.

3 Vectors

At the most basic level, a vector is an object with a direction and a magnitude. The most common representation of a vector is an arrow with a length proportional to the vector's magnitude. Figure 1 shows this representation of a vector.

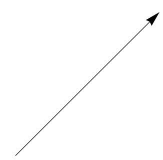

Figure 1: The vector $\vec{V}$.

What makes vectors really special is that, as in Figure 1, they exist independently of a coordinate system. In order to discuss them mathematically, we have to pick a coordinate system to write them in. This is called a choice of basis. Once the choice is made, the vector can be notated in terms of the basis several ways. The first is to write it as a row vector or column vector as defined in Section 2.1. For example, the vector $\vec{V} = \begin{pmatrix} 3/4 & 3/4 \end{pmatrix} = \begin{pmatrix} 3/4 \\ 3/4 \end{pmatrix}$ points from the origin to the point $(3/4, 3/4)$ in the x-y plane. You can also write the same vector as $\vec{V} = \frac{3}{4}\hat{x} + \frac{3}{4}\hat{y}$. It means the same thing. See Figure 2 for an example of a choice of basis.

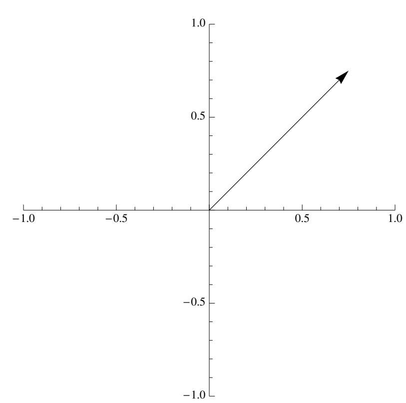

Figure 2: The vector $\vec{V} = (3/4, 3/4) = \frac{3}{4}(\hat{x} + \hat{y})$.

We could, of course, always pick a different coordinate system (basis). This makes us write things differently. Let's take a look at Figure 3, where I have chosen a coordinate system rotated by $\pi/4$ (or 90°). The vector's orientation was left unchanged, but the notation in terms of the new coordinates is different. If we use a prime to denote the new coordinates, we now have $\vec{V} = \frac{3}{4}(-\hat{x}' + \hat{y}')$. Now here is the tricky part. Since it is more normal to draw the x-axis horizontally and the y-axis vertically, we could also draw this as in Figure 4. This is sort of confusing, but also really common. We'll see examples in our study of quantum mechanics that may make everything a little more clear.

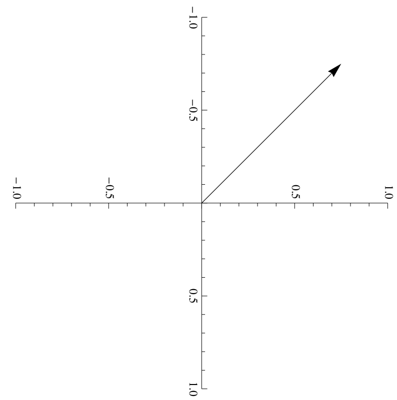

Figure 3: Here we show the same vector in a different coordinate system.

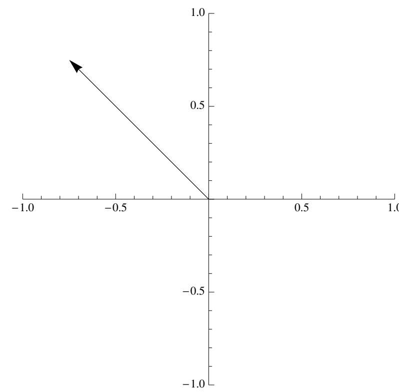

Figure 4: Now we show the vector of Figure 3 with the coordinate system oriented as in Figure 2. Doing this, one must be careful that the relationship between the two bases is explicitly defined!

Since vectors have a length and a direction, we could also specify them in terms of these two quantities rather than the (x, y) point of their tips. If you're familiar with polar coordinates, this is exactly the same thing. Every vector can now be specified as $\vec{V} = (r, \theta)$. The length is r and θ specifies the direction as you can see in Figure 5. To convert between cartesian (x and y) coordinates to polar coordinates, you can use the relations:

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned} \tag{21}$$

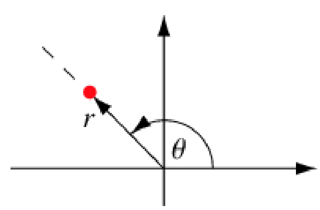

Figure 5: A vector in polar coordinates. From http://mathworld.wolfram.com/PolarCoordinates.html

Quick Question: Solve (21) for r and θ in terms of x and y.

One final thing for this section. The *magnitude* or *norm* of a vector $\vec{V}$ is often denoted as $|\vec{V}| = V$. You should have determined how to find this length, since r is exactly the magnitude of the vector.

3.1 Binary operations with vectors

There are a few mathematical operations we will need to be able to perform with vectors. In analogy with normal numbers, vectors can be added or multiplied. Vector addition is easier, so let's start there.

Suppose we have two vectors, $\vec{U} = a\hat{x} + b\hat{y}$ and $\vec{W} = c\hat{x} + d\hat{y}$, where a, b, c, and d are arbitrary real numbers. The sum of the two vectors is written as $\vec{U} + \vec{W} = (a+c)\hat{x} + (b+d)\hat{y}$. Note that the sum does not mix the x components with the y components at all, and it produces a new vector. We could also find this result by writing the vectors as column vectors and adding using the rules for matrices.

The sum of two vectors can be understood by looking at Figure 6. Take the first vector in the sum and place its tail at the origin (0,0). Now put the tail of the next vector at the end of the first. The sum points from the origin to the tip of the second vector. If we were to switch the order of addition, we would still get the same result.

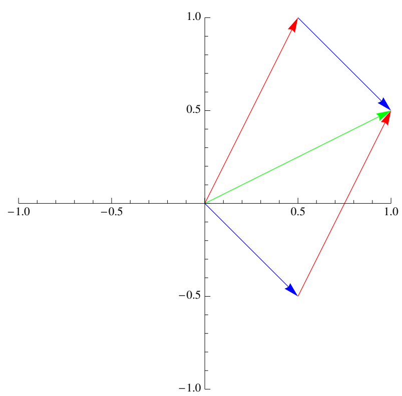

Figure 6: Visualizing the vector sum. The vectors added here are $\vec{U} = \frac{1}{2}\hat{x} + \hat{y}$ (red in the electronic version) and $\vec{W} = \frac{1}{2}(\hat{x} - \hat{y})$ (blue). Work out the sum for yourself to make sure the green line points to the correct location.

Quick Question: What do you think is the correct prescription for adding vectors in three dimensions? How about in four dimensions? Tip: In three dimensions we would use the symbol $\hat{z}$ for the new direction.

Quick Question: How does one subtract vectors?

Multiplication is a little trickier. There are two kind of multiplication for vectors, the *dot product*[3] and the *cross product*. We only need the dot product in this class, so let's ignore the cross product since there is a lot to learn already! The dot product takes two vectors and makes a number (or *scalar*) out of them. Mechanically, the dot product is given by $\vec{U} \cdot \vec{W} = ab + cd$. You can also find this by representing both vectors as column vectors then finding the product $U^T V$.

[3]This can also be called a *scalar product* or an *inner product*.

Geometrically, the dot product can also be found using the magnitudes of the two vectors and the angle θ between them using the relationship $\vec{U} \cdot \vec{W} = |\vec{U}||\vec{W}|\cos\theta$. From this you can see that the dot product essentially measures how much the vectors overlap, see Figure 7. Vectors at right angles will have a dot product of zero and are called *orthogonal* (a fancy name for perpendicular).

Quick Question: What is the meaning of the dot product of a vector with itself?

Quick Question: What is the dot product $\hat{x} \cdot \hat{y}$?

Later on we will see generalizations of the inner product. The definition I have just introduced is known as the *Euclidean* inner product, because it reproduces the standard Euclidean geometry in flat space.

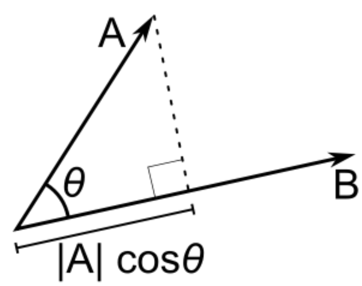

Figure 7: The dot product gives the magnitude of one vector times the overlap of the second vector with the first.

3.2 Complex numbers as vectors

There is a nice analogy between the complex numbers and vectors in the plane. The notation $z = x + iy$ suggests that the complex number z could be thought of as a two component vector pointing to the point (x, y). When drawn on a two dimensional coordinate system, called the *complex plane*, the horizontal axis represents the real part and the vertical part represents the imaginary part[4].

Quick Question: What does complex conjugation look like in the complex plane?

Addition of complex numbers is by component, just as for vectors. Multiplication proceeds quite differently, however. The inner product from vectors gave us a real scalar as the result. However, there is in general no analogous prescription for the complex numbers that would accomplish this. The one special case is when finding the norm of a complex number. We even notate it similarly: $|z|^2 = z^*z = x^2+y^2$. You can see that the modification from the vector case is essentially to take the product using the hermitian conjugate (rather than just the transpose) of the column vector to make the row vector. This will have a direct analogy in quantum mechanics!

[4]This picture is sometimes called an *Argand diagram*

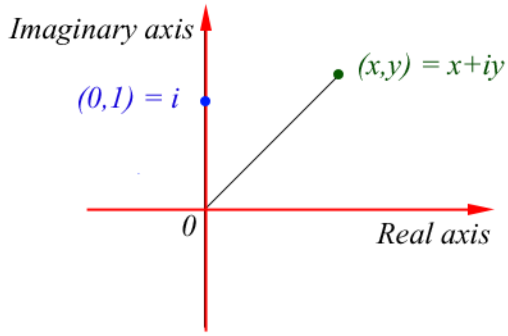

Figure 8: A complex number visualized in the complex plane.

We of course have to have a general prescription for multiplying or dividing complex numbers. The best way to do this is actually in polar coordinates! Going back to Euler's formula (4), you can easily reproduce the prescription of (21).

$$z = x + iy = re^{i\theta} = r(\cos\theta + i\sin\theta) \rightarrow \begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases} \tag{22}$$

With this form, multiplication and especially division are much easier than with the cartesian version. Addition and subtraction have become somewhat more difficult, however!

Quick Question: Let $z = \frac{1+i}{\sqrt{2}}$ and $w = -i$. Write these numbers in polar coordinates. Find $z + w$, zw, and w/z.

Quick Question: What is $4e^{i\frac{15\pi}{2}}$ in Cartesian coordinates.

Note that although we can think of these representations in the complex plane as 'cartesian' and 'polar,' it's not necessary. In particular, the Euler identity (4) follows directly from the definitions of the exponential, sine, and cosine functions.

4 Eigenvalues and eigenvectors

An *eigenvector* V of a square matrix A is a special non-zero column vector that satisfies a relation of the type

$$AV = \lambda V \tag{23}$$

where λ is a scalar constant called the *eigenvalue*. In words, we would say the matrix acting on the eigenvector gives a scalar multiple of the original vector. The *eigensystems* of matrices have a very special role in quantum mechanics, so we want to learn how to find them.

The very simplest case is a diagonal matrix. For such a matrix, each element A_{ii} on the diagonal corresponds to an eigenvalue with an eigenvector V that has a non-zero element only in the i-th row. Here's an example:

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{24}$$

You might note from this example that any scalar multiple of an eigenvector is also an eigenvector with the same eigenvalue.

What about finding the eigensystem of a more general matrix? There are a lot of subtleties to this for different kinds of matrices. Luckily, we will only need to find the eigenvalues of Hermitian matrices in quantum mechanics. Let's focus exclusively on these.

Here are a few properties of the eigenvalues of a Hermitian matrix:

i) The eigenvalues of a hermitian matrix are always real.

ii) Every $n \times n$ hermitian matrix has n orthogonal eigenvectors.

How do we find the eigenvalues of a Hermitian matrix? Start be rewriting (23) in the form

$$(A - \lambda I)V = 0 \tag{25}$$

I will not prove it here, but this equation has a solution $V \neq 0$ if and only if $\det(A - \lambda I) = 0$. Let's choose a definite form for the matrix,

$$A = \begin{pmatrix} a & ce^{i\theta} \\ ce^{-i\theta} & d \end{pmatrix} \tag{26}$$

where a and b must be real. Then

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} a - \lambda & ce^{i\theta} \\ ce^{-i\theta} & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - c^2 \\ &= \lambda^2 - (a + d)\lambda + ad - c^2 = 0 \end{aligned} \tag{27}$$

This is a quadratic equation for λ so the two solutions are

$$\lambda = \frac{a + d \pm \sqrt{(a + d)^2 - 4(ad - c^2)}}{2} = \frac{a + d \pm \sqrt{a^2 + d^2 + 4c^2 - 2ad}}{2} \tag{28}$$

With the eigenvalues known, we want to solve for the eigenvectors. The explicit formula is terribly ugly if we try to stay this general, so instead let's pick a matrix that we will see again later.

$$S = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{29}$$

Plugging into (28) shows that the eigenvalues of S are $\lambda = \pm 1$. To find the eigenvector corresponding to $\lambda = +1$, we simply solve

$$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = (+1) \begin{pmatrix} a \\ b \end{pmatrix} \tag{30}$$

$$\begin{pmatrix} -ib \\ ia \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \tag{31}$$

$$\tag{32}$$

Evidently the eigenvector with $\lambda = +1$ is $\begin{pmatrix} 1 \\ i \end{pmatrix}$.

Quick Question: What is the eigenvector with $\lambda = -1$?

The general procedure for finding the eigenvalue of any Hermitian matrix A is very similar. Simply follow these steps:

1. Solve $\det(A - \lambda I) = 0$ for the eigenvalues λ. If the matrix is $n \times n$ you should get n solutions.

2. For each eigenvalue λ, let the elements of a vector V be arbitrary and solve $AV = \lambda V$ for these elements. This will generally need to be done n times.