Quantum Mechanics
Day 1

Today
I. What is QM? Syllabus
II. Could QM by different?
III. Transformations

I. What is Quantum Mechanics?

See the syllabus for the start of an answer to this question.

II. Could Quantum Mechanics Be Different?

What would happen if quantum mechanics were not as it is? We'll spend some time with this question during this first week. Let's overview some of the unusual features of the quantum operating system; we'll meet them again and again.

- The outcome of a measurement of a quantum system is often, but not always, one of a discrete set of possibilities. (It's like digital computing.)
- If you setup and run an experiment multiple times, then, despite having taken every care to set it up in exactly the same way, you will often get different results. In sharp contrast to the saying that defines insanity, it is now quite reasonable to do exactly the same thing and to expect different results.
- Quantum theory tells you how to compute the probabilities of these measurement outcomes. It has nothing to say about the exact outcome you'll get during an individual trial run.

<ins>Ex.</ins> Since Quantum Mechanics is often digital, let's look at a simplest quantum digital system, one with just two states; it's called a qubit.

There are many different ways to realize a qubit. One nice one is an optical interferometer—here with just a single photon in it.

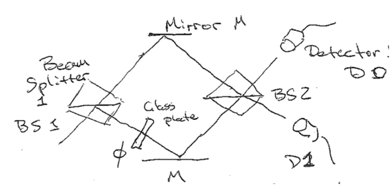

Figure 1: A Mach-Zehnder Interferometer.

The outcome of the experiment at detector $D0$ is either a click—a photon was measured—or, no click and no photon was measured. Just two possible outcomes. These outcomes can occur with different probabilities, p and $1 - p$.

We often say that probabilities can't be negative—this is very sensible, what would it mean for there to be a -20% chance of snow? Let's restrict consideration to real number for the moment. Certainly the probability of all possible outcomes should satisfy

$$p_1 + \cdots + p_N = \sum_i p_i = 1.$$

In other words, of all possible outcomes, certainly one must happen.

But, this uses the 1-norm, essentially the sum of absolute values. (Absolute values since the probabilities are necessarily positive.) What if we also allowed the 2-norm, that is, a norm like that used in the pythagorean theorem?

We can do this if we interpret the squared value $|\alpha|^2$ as a probability. Take (α, β) as variables, now we want

$$|\alpha|^2 + |\beta|^2 = 1.$$

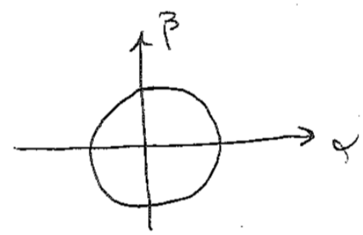

Figure 2: The plane of values of α and β. The set $|\alpha|^2 + |\beta|^2 = 1$ is the unit circle in this plane for real α and β.

In the first few lectures we will explore the relative strengths of the 1-norm and 2-norm formalisms. For example, we can ask "Why not just forget about α and β and only consider $|\alpha|^2$ and $|\beta|^2$?"

III. Transformations

Any answers to this question? This would amount to returning to the 1-norm formalism. The difference between the two formalisms is striking when you consider transformations!

Although many of you have taken linear algebra it is likely that you haven't used it in the way that physicists do most often. For this reason, I pause the discussion of quantum mechanics completely for a bit and want to just discuss transformations using matrices. The Euclidean geometry mentioned above is perfect for these purposes, and in many respects can be thought of as a classical analog of what we are going to do in the quantum theory.

Put coordinates (x, y) on the standard Euclidean plane. We can describe the points of this plane using several different common notations. For example you might write $\vec{r} = x\hat{x} + y\hat{y}$, or you might use the graphical notation, or you might write $|\mathbf{r}\rangle$. We will soon start switching between some of these various notations, but for now, let's display the coordinates of the point as a 2×1 array

$$\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

This means exactly the same thing as $\vec{r} = x\hat{x} + y\hat{y}$, but will be much more convenient for computations. For example, we can now perform rotations on vectors using the rules of matrix multiplication. To figure out which matrix represents rotation by angle θ, we first ask what rotating $\hat{x}$ and $\hat{y}$ gives by looking at the geometry (see figure below)

$$\mathsf{R}\hat{x} = \mathsf{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

and

$$\mathsf{R}\hat{y} = \mathsf{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}.$$

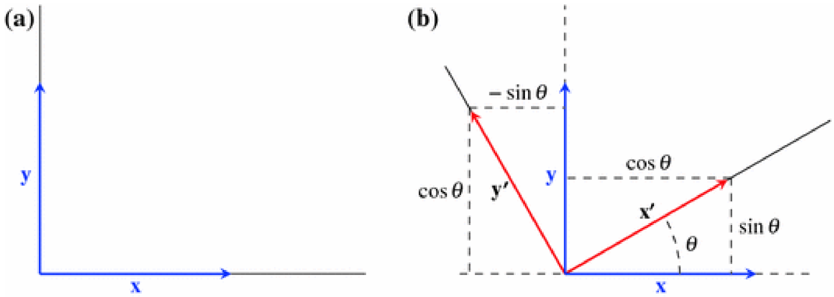

Figure 3: Geometry of rotations for $\hat{x}$ and $\hat{y}$.

On the other hand, any matrix R can be written as

$$R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix},$$

and so direct matrix multiplication gives

$$R\hat{x} = R\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} R_{11} \\ R_{21} \end{pmatrix} \quad \text{and} \quad R\hat{y} = R\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} R_{12} \\ R_{22} \end{pmatrix}.$$

Setting the two expressions for $R\hat{x}$ equal and those for $R\hat{y}$ equal gives

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

This is the matrix that rotates any vector by an angle θ in the counter-clockwise direction. Having found this rotation matrix, we can now rotate $\vec{r}$ to get the new vector $\vec{r}'$,

$$\vec{r}' = R\vec{r} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}.$$

This all works very well, but it took a bit of work to find R. Is there a more efficient way to do it? It turns out that there is and the key is to ask what property of the vector $\vec{r}$ is invariant under rotations. A little thought and some sketches will convince you that the length of a vector is invariant under rotations.

We can check this directly by computing the length of the vector $\vec{r}'$

$$\begin{aligned} r'^2 &= \vec{r}' \cdot \vec{r}' = \tilde{\vec{r}}' \vec{r}' \\ &= (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta) \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix} \\ &= (x\cos\theta - y\sin\theta)^2 + (x\sin\theta + y\cos\theta) = x^2 + y^2 = r^2! \end{aligned}$$

This leads us to our first definition of **orthogonal transformations**: they are those transformations that preserve the 2-norm. We have just shown that rotations are one type of orthogonal transformation.

There is a second, and efficient, way of doing this computation and of viewing orthogonal transformations. We can write the calculation in a more abstract way

$$\vec{r}' \cdot \vec{r}' = \widetilde{(\mathsf{R}\vec{r})}(\mathsf{R}\vec{r}) = \tilde{\vec{r}}\tilde{\mathsf{R}}\mathsf{R}\vec{r}$$

and so, if the right hand side is going to be equal to $\vec{r} \cdot \vec{r}$ it must be that

$$\boxed{\tilde{\mathsf{R}}\mathsf{R} = \mathsf{I},} \tag{1}$$

where

$$\mathsf{I} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is the identity matrix. Equation (1) gives us our second definition of **orthogonal transformations**, namely those whose matrix representations satisfy Eq. (1). Notice that this equation can also be written as

$$\boxed{\tilde{\mathsf{R}} = \mathsf{R}^{-1}.} \tag{2}$$

With this detour into transformation theory complete, we can return to the question of what would be different in physics if we worked with a 1-norm.