The transition from discrete to continuous spaces

(Version: June 13, 2020)

What this note is about

The old Porland and Denver groups spent several months studying quantum theory in a discrete setting. That is, in a finite vector space, where states are vectors and operators are matrices. The mathematics here (basic linear algebra) is pretty easy. The online group is now looking at this same material on Sundays.

We are now starting to look at quantum mechanics, which is a continuous theory. The states are functions and the operators are are things like derivatives. The math here (calculus, differential equations, and Fourier transforms) is significantly more difficult.

And yet (I claim) the underlying *theory* is in fact *exactly the same* once you strip out the mathematical complexities.

In order to do this, we need to understand the relationship between the kind of mathematical space we've been using for quantum information theory and the kind of space we'll now be using for quantum mechanics. Explaining the transition between these two spaces is the purpose of this write-up.

Formal vs informal descriptions

In what follows I am going to make various claims about the relationship of objects and procedures used in a finite discrete vector space vs an infinite continuous one. There are two ways to go about discussing these claims.

(1) The formal approach

The formal (axiomatic) approach would be to list out the postulates (axioms) of (for example) vector spaces. Then we would take a specific set of functions and apply all the axioms. Proving that these functions fulfilled all of the vector space axioms, we would have shown that they are elements of a vector space – i.e., vectors. This is how we would go about *proving* the claims.

This is a good thing to do, and if you want to give it a shot, take a look at the vector space axioms, and see what happens when you try to apply them to a function. You can find one way of breaking down the axioms at:

https://github.physicscafe.org/2020-QM-Series/120-Vector_Space_Axioms.pdf

(If anybody actually wants to try this, it might be a good idea if I helped you get started. There are some important details that I haven't mentioned :-)

We are *not* going to follow the formal approach here. I won't be saying any more about it in this note.

(2) The informal approach

What we *are* going to do is to take an informal approach to each claim about a relationship between the two spaces. I'm not going to try to *prove* that they are true. I'm simply going to explain *how* you can look at them as being true and hopefully provide some feeling for *why* you would want to do this, in practical terms.

Thinking about a function as a vector

In this little section I'm not going to worry about bras vs kets (or about conjugation) and I'll just write a vector as a list of numbers in a row. This is just to save space and has no other significance.

$$f = (f_1, f_2, f_3, ...f_n)$$

So the vector f has n elements: f_1 through f_n.

Now when it comes to functions, we often think of them as a "machine" that inputs the independent variable and outputs the resulting value:

$$x \rightarrow \boxed{f} \rightarrow f(x)$$

$$\text{if } \; f(x) = x^2 \; \text{ then } \; 3 \rightarrow \boxed{f} \rightarrow 9$$

But here we're going to take a different view of a function. We're going to think of it simply as a list of numbers corresponding to the "outputs." If $f(x) = x^2$ then:

$$f = (\ldots \; 4 \; \ldots \; 9 \; \ldots \; 16 \; \ldots)$$

The reason for the ellipses is that f doesn't just operate on integers, but rather on any real number. So there are actually a *continuously infinite* number of elements on this list. We can consider the independent variable to be the index into this list.

$$f(2) = 4, \quad f(2.5) = 6.25, \quad f(3) = 9, \text{ and so on.}$$

So we have both functions and vectors as essentially the same thing. It's just that:

1. We typically use subscripts to index into vectors: f_n and "arguments" to index into functions: $f(x)$.
2. Usually the index to a vector is an integer and the index to a function is a real number.

Inner Products

Here's how we've been computing the inner product of two vectors:

$$\langle\phi|\psi\rangle = \begin{pmatrix} \phi_1^* & \phi_2^* & \dots \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \dots \end{pmatrix} = \phi_1^*\psi_1 + \phi_2^*\psi_2 + ...$$

What we get is a scalar that results from adding the products of all the corresponding individual elements of the two vectors. This could be summarized, for two vectors with N elements, as:

$$\sum_{k=1}^{N} \phi_k^*\psi_k$$

But suppose that our vectors are functions, and have an infinite number of elements. We can try writing:

$$\sum_{k=1}^{\infty} \phi_k^*\psi_k$$

But there's a problem. A sum like this is going to be infinite. And of course we want a finite result. We could try to address the situation by multiplying each product by an infinitesimal, in hopes that the resulting sum would be finite:

$$\sum_{k=1}^{\infty} \phi_k^*\psi_k\, dk$$

But there's another problem, which is that we are going to be summing over a *continuously* infinite number of elements. You probably see where I'm heading here. We *have* a mathematical procedure to sum a continuously infinite number of elements multiplied by an infinitesimal. It's called an integral:

$$\int_1^{\infty} \phi^*(k)\psi(k)\, dk$$

And this is why an integral is an inner product. It's not that *every* integral is an inner product. But when we do an inner product in function space it will be an integral.

Projection

Projection is the process we use to find out "how much" of one vector lies along another vector. That is, $\langle\phi|\psi\rangle$ tells us how much of $|\psi\rangle$ lies along $|\phi\rangle$. We can do the same thing with functions. The projection $\int \phi(x)^*\psi(x)\,dx$ tells us "how much" of $\psi(x)$ lies along $\phi(x)$.

Normalization and Orthogonality

A vector is normalized (has a length of one) if its inner product with itself is equal to one: $\langle\psi|\psi\rangle = 1$. So if we're working with functions, then since an inner product is an integral, the normalization criterion is:

$$\int_a^b \psi^*(x)\psi(x)\,dx = 1$$

Similarly, we know that two vectors are orthogonal if one has a zero projection onto the other (an inner product of zero): $\langle\phi|\psi\rangle = 0$. So the test for orthogonal functions is:

$$\int_a^b \phi^*(x)\psi(x)\,dx = 0$$

If you haven't run into these concepts before, ideas like how much of one function "lies along" another or two functions being "orthogonal" to one another might sound meaningless. But (in an abstract sense) functions are vectors, and they do have these relationships. These concepts come into play when we talk about bases and change of basis for functions.

Bases

If functions are vectors, then the space they live in is a vector space, and there must be a basis out of which we can construct every function in the space.

Let's start with an example. Take the space of functions with a one-dimensional, finite domain (a line segment), where the functions all smoothly come to zero at the limits of the domain:

$$f(x):\ a \le x \le b,\ f(a) = 0,\ f(b) = 0.$$

In a space like this, appropriately constructed sine functions form a basis, out of which we can construct any function in the space. *See the ...*

The "number" of basis functions can be discrete or continuous

In function space there are always an infinite number of basis functions. But that infinity can be either discrete (like the integers) or continuous (like the real numbers).

Don't confuse this with the number of *elements* that a given function (vector) has. That number is always continuous. What we're talking about here is *how many functions* you need to form a basis.

When a function ranges over a finite domain: $a \leq x \leq b$, we typically have a discrete infinity of basis functions. When the function is defined over all space: $-\infty \leq x \leq +\infty$, then we typically have a continuous infinity of basis functions.

Change of basis

We do a change of basis with a series of projections (inner products).

Say we have a set of basis vectors $\{\, |b_n\rangle\}$ and a vector $|\psi\rangle$ written in the default basis. In order to write $|\psi\rangle$ in the b basis, we project it onto each of the b basis vectors. Each projection yields one of the elements of $|\psi\rangle$ in the b basis:

$$
\begin{aligned}
c_1 &= \langle b_1|\psi\rangle \\
c_2 &= \langle b_2|\psi\rangle \\
&\dots \\
|\psi\rangle &= c_1|b_1\rangle + c_2|b_2\rangle + \dots
\end{aligned}
$$

We can neatly write this procedure in summation notation as:

$$|\psi\rangle = \sum_n \langle b_n|\psi\rangle |b_n\rangle$$

Now simply replace the inner product with an integral, change to function notation, and you have:

$$\sum_n \left(\int b_n^*(x)\psi(x)\,dx \right) b_n(x)$$

This is called a *Fourier series*. It's usually written as two separate lines, corresponding more closely to the way I first did the projections above, using the c variables for the weights:

$$c_n = \int b_n^*(x)\psi(x)\,dx$$

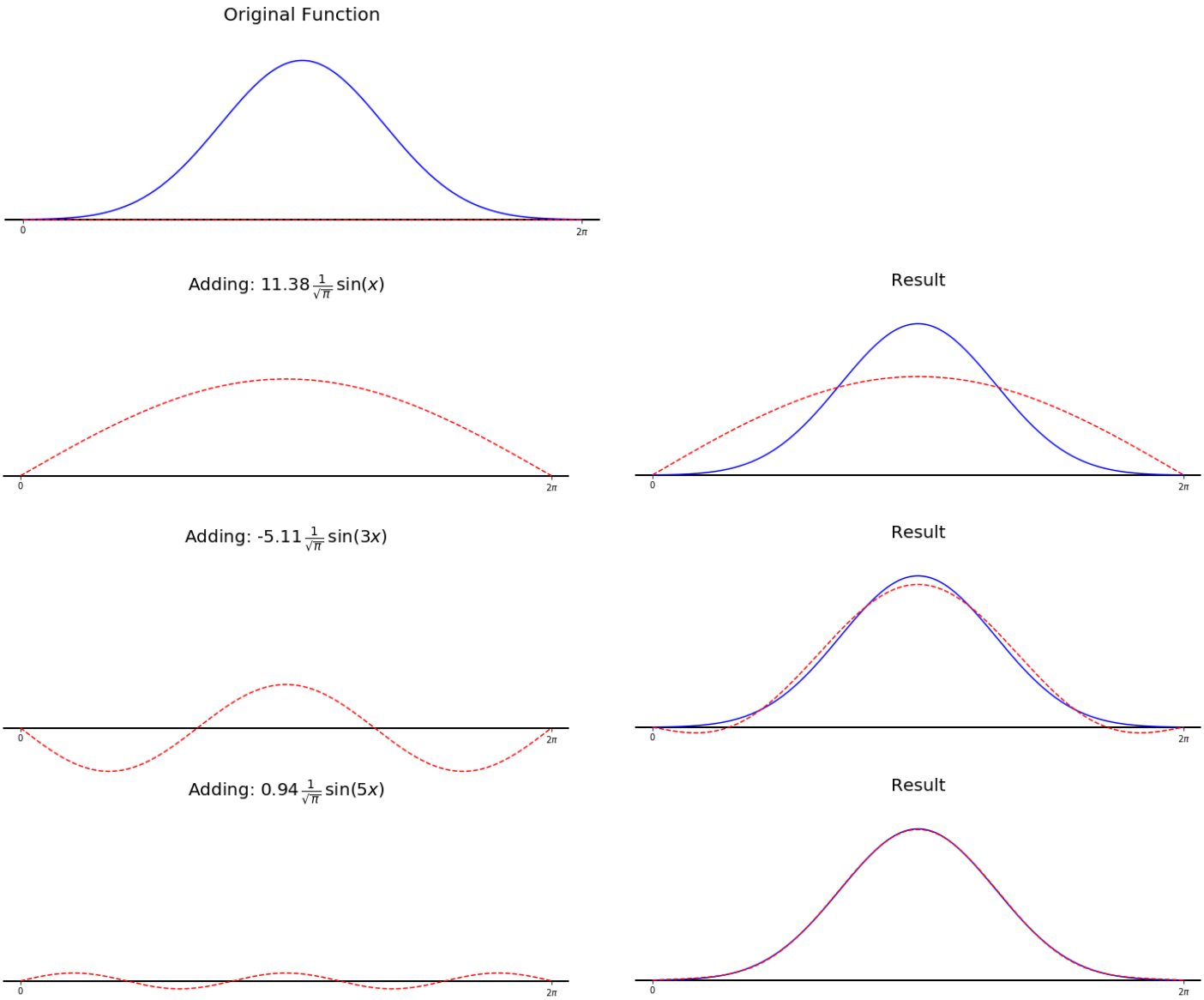
$$\psi(x) = \sum_n c_n\, b_n(x)$$

In the event that you have a *continuously* infinite number of basis functions you'll get a *Fourier transform*. This means you'll also have to use an integral to sum up the weighted basis functions. In this case people don't typically use n as a subscript, since n is so strongly associated with the integers:

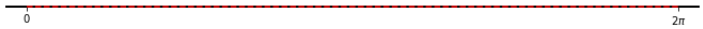
$$c_k = \int b_k^*(x)\psi(x)\,dx$$

$$\psi(x) = \int c_k\, b_k(x)\,dk$$

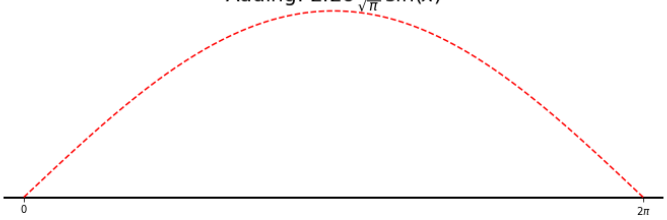
Examples with sine functions



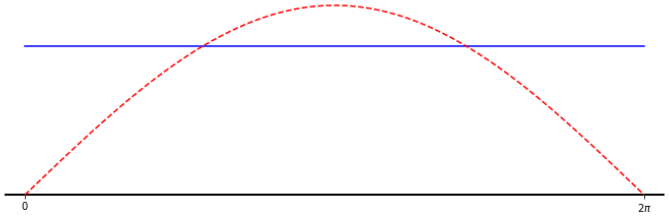
Original Function



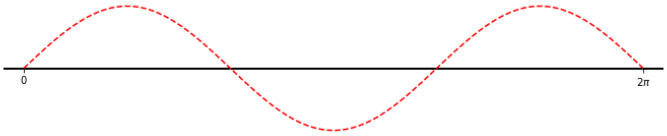
Adding: $2.26\frac{1}{\sqrt{\pi}}\sin(x)$



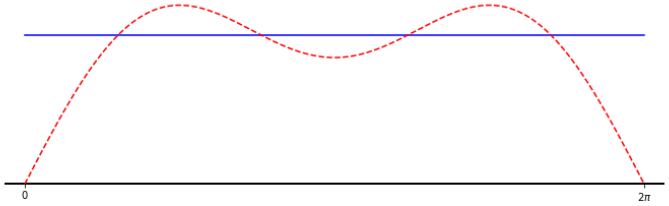
Result



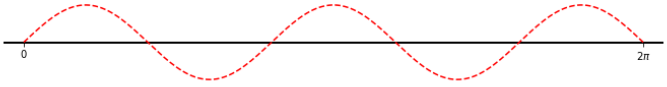
Adding: $0.75\frac{1}{\sqrt{\pi}}\sin(3x)$



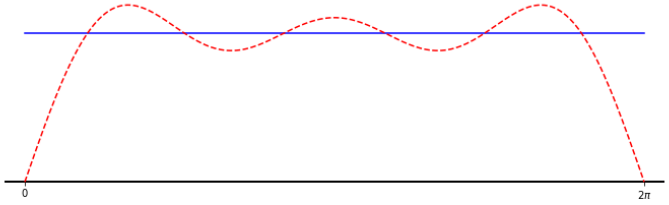
Result



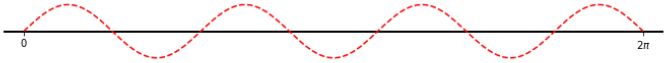
Adding: $0.45\frac{1}{\sqrt{\pi}}\sin(5x)$



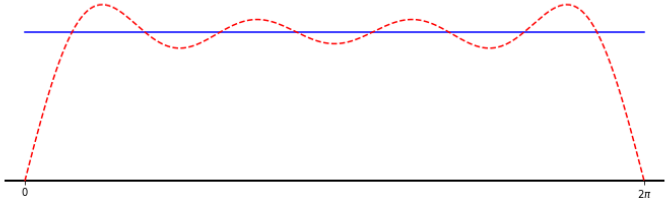
Result



Adding: $0.32\frac{1}{\sqrt{\pi}}\sin(7x)$



Result



Probability vectors and functions

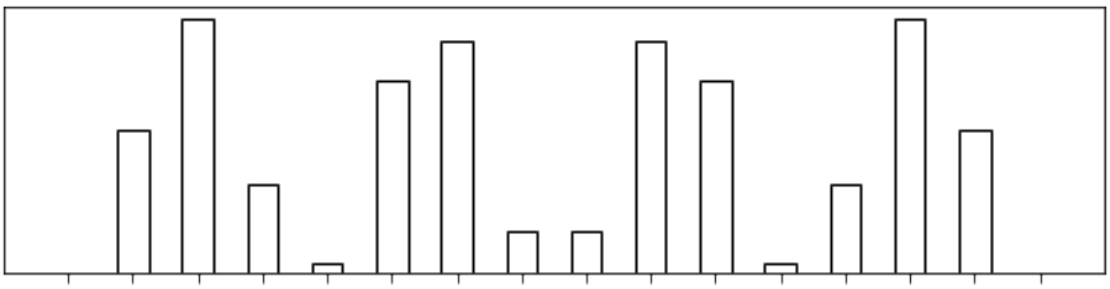
A quantum state, or state vector, is represented by a ket, which on the board we write as a column vector. Here I'm writing it as a row instead. Writing it as a row has no mathematical significance. It's still a ket, not a bra. The individual elements have not been conjugated.

$$|\psi\rangle = (\psi_1, \psi_2, \psi_3, ...)$$

As you may remember, the individual elements of the state vector are called *probability amplitudes.* They are not probabilities, but they are related to them. The probability of a measurement resulting in the state represented by ψ_n is given by the "modulus squared" of ψ_n or $|\psi_n|^2$. This is the same thing as $\psi_n \psi_n^*$.

Now you could imagine a vector of probabilities, rather than a vector of probability amplitudes. Lets call this vector $\vec{P}$. In other words $\vec{P} = (|\psi_1|^2, |\psi_2|^2, ...)$.

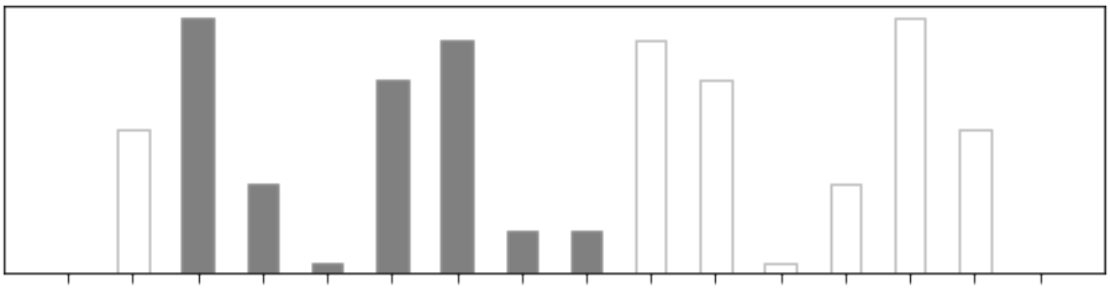
All of the elements of $\vec{P}$, being probabilities, are positive real numbers. If $\vec{P}$ happened to have 16 elements, we could plot them like this:



We can think of $\vec{P}$ as a (discrete) *probability function.*

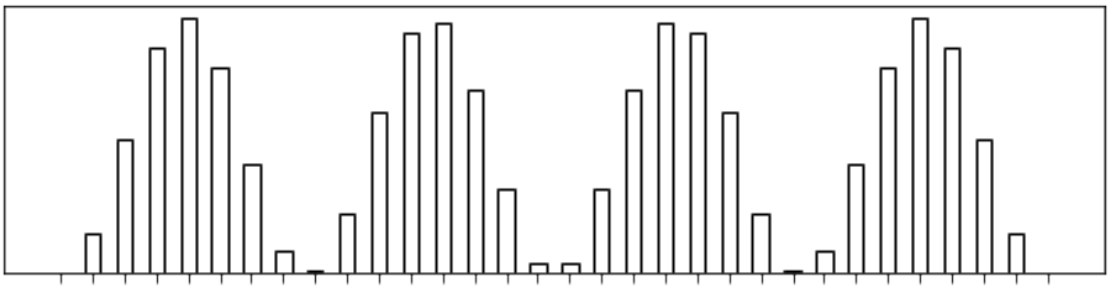
If we want to find the probability of getting any of the results 2 through 8, we simply add the probabilities:

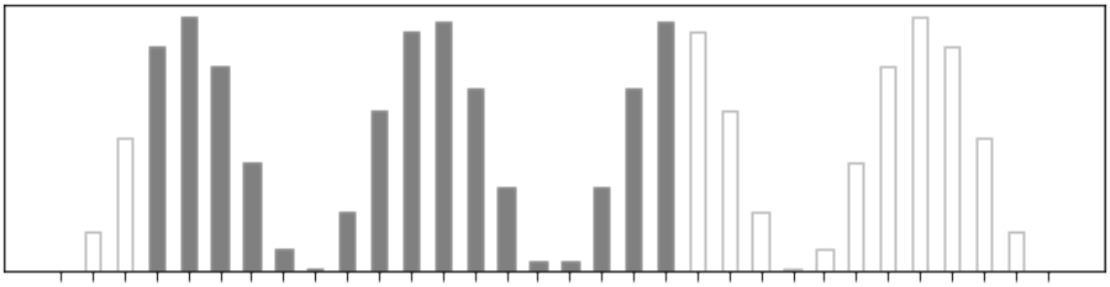
$$Prob(2,8) = \sum_{2}^{8} P_n$$



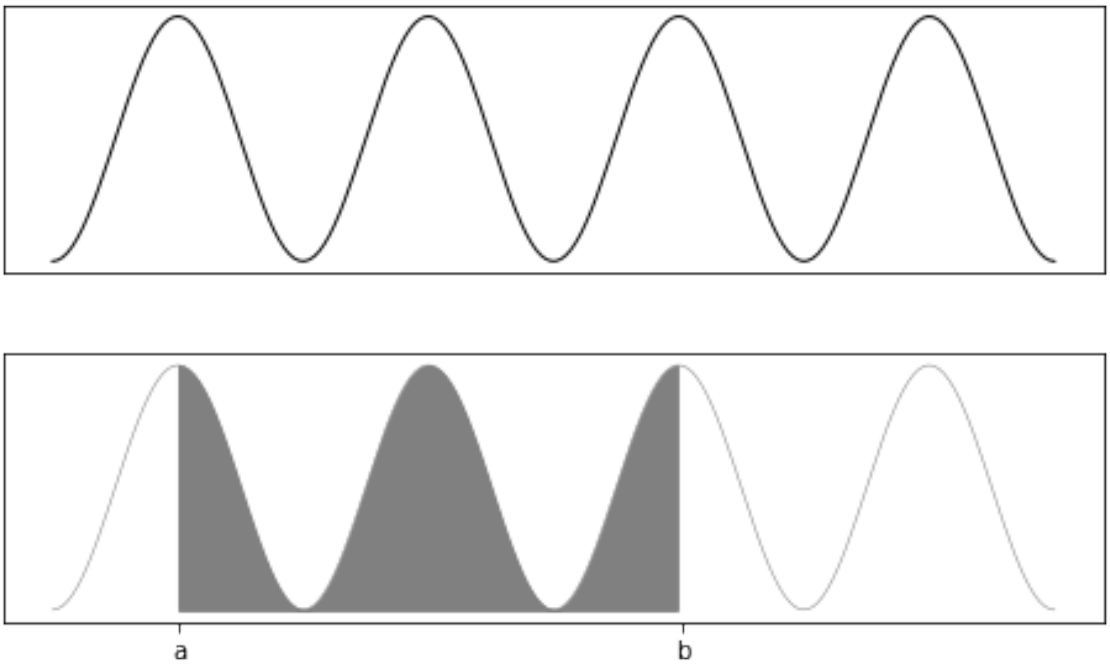
Probability density functions

As the size of $\vec{P}$ grows, it starts to look more and more like a continuous function:





And, if our "probability" vector actually has an infinite number of elements, then:



But there's a catch. With a continuous number of elements we can no longer do a simple summation. The summation of an infinite number of elements will be infinity. We need to do an integral over the desired range of what we now call the function P(x):

$$\int_a^b P(x)\,dx$$

$P(x)$ is no longer a probability function. It is now what we call a *probability density function*. The height of the function at each individual point no longer represents a finite probability, but rather the (infinitesimal) "density" of the probability at that point.

The default basis and the Dirac delta

In a discrete vector space, the "default" or "primary" basis is a set of vectors which have a single 1 component, with the other components zero. For example in a "two bit" (dimension four) vector space we have:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Under most circumstances, we write out the components of all vectors in this default basis. So projection onto the default basis simply "picks out" the individual components of the vector.

An example. We have the state for two bits in a quantum computer:

$$\alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$$

If we write it as a column vector and project is onto the four primary basis vectors (called the "classical basis" in QC) then:

$$\begin{pmatrix} 1\,0\,0\,0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} = \alpha, \quad \begin{pmatrix} 0\,1\,0\,0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} = \beta, \quad \begin{pmatrix} 0\,0\,1\,0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} = \gamma, \quad \begin{pmatrix} 0\,0\,0\,1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} = \delta$$

So what is the default basis in continuous function space? The "components" of the functions are simply the function values (the points on the line that we plot when we draw the function). What kind of mathematical object will "pick out" these points when we do a projection?

Bearing in mind that a projection is an integral, we are looking for something like this:

$$\int \delta_k f(x)\, dx = f(k)$$

where δ_k is the kth vector of the standard basis.

There is such an object δ and it is called the Dirac delta. You might think the deltas as a (continuously infinite) set of basis vectors, each of which have one infinite component and all the rest of the components zero:

$$\ldots \begin{pmatrix} \infty \\ 0 \\ 0 \\ 0 \end{pmatrix} \ldots \begin{pmatrix} 0 \\ \infty \\ 0 \\ 0 \end{pmatrix} \ldots \begin{pmatrix} 0 \\ 0 \\ \infty \\ 0 \end{pmatrix} \ldots \begin{pmatrix} 0 \\ 0 \\ 0 \\ \infty \end{pmatrix} \ldots$$

When one of these vectors gets multiplied by the dx in the integral, the infinity times the infinitesimal results in an infinitely high, infinitesimally thin "spike" *of unit area* which picks out the one number desired. But if that visualization doesn't work for you, you can just think of the object δ_k as being *defined* by the integral above.

Caveat: I've used a slightly non-standard notation above. The Dirac delta is normally defined as $\delta(x) = 0$ *for* $x \neq 0$ *and the unit one spike when* $x = 0$*, or* $\int \delta(x-k)f(x)\,dx = f(k)$*. So my* δ_k *is equivalent to* $\delta(x-k)$ *in the standard notation.*