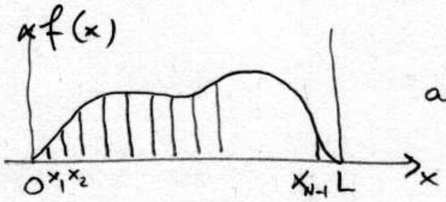
I.10. Infinite number of dimensions

Almost all operators that appear in Quantum Mechanics have an infinite number of eigenvalues and eigenvectors. In most cases the number of eigenvalues is not even countable. So we need to generalize what we have learned so far accordingly.

Let us develop the necessary tools by studying a familiar problem.

Consider a string clamped at $x=0$ and $x=L$, whose displacement in $0 \le x \le L$ is given by $f(x,t)$:



at some time t.

Suppose we want to communicate $f(x)$ at some time t to a person on the moon. We can divide the interval $0 \le x \le L$ into N equal parts, measure the displacement at $x_i = i\frac{L}{N}$ $(i = 1, 2, \ldots, N-1$ — $f(x_N) = f(L) = 0)$ and transmit the $N-1$ numbers $f(x_i)$ by wireless. We refer to this discretization by $f_N(x)$. (Obviously, to get a very precise description of $f(x)$, N should be large.) Let us interpret the ordered N-tuple $\{f_N(x_1), f_N(x_2), \ldots, f_N(x_N)\}$ as the components of a ket $|f_N\rangle$ in a vector space $\mathbb{V}^N(\mathbb{R})$:

$$|f_N\rangle \leftrightarrow \begin{pmatrix} f_N(x_1) \\ \vdots \\ f_N(x_N) \end{pmatrix}$$

(we have added $f_N(x_N) = 0$ to make the counting easier)

The basis vectors in this space that correspond to this decomposition are

$$|x_i\rangle \longleftrightarrow \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i^{th} \text{ position} \\ \\ \\ \\ \end{matrix}$$

They correspond to a discrete function which is $f(x) = 1$ at $x = x_i$ and zero everywhere else. They are orthonormalized and complete:

$$\underbrace{\langle x_i | x_j \rangle = \delta_{ij}}_{\text{orthonormality}} \quad , \quad \underbrace{\sum_{i=1}^{N} |x_i\rangle\langle x_i| = \hat{I}}_{\text{completeness}}$$

The discrete function $|f_N\rangle$ is then represented in this basis as

$$\boxed{|f_N\rangle = \sum_{i=1}^{N} f_N(x_i)\,|x_i\rangle}$$

Imagine a space containing N mutually orthogonal axes, one axis for each point x_i. Along each axis is a unit vector $|x_i\rangle$. $|f_N\rangle$ is then the vector obtained by scaling each unit vector by the value of the function at x_i and adding those vectors like you add 3-vectors, only in a much larger-dimensional space.

Each different shape $g(x), h(x)$ etc of the string between $x=0$ and $x=L$ will, in this discrete approximation, be represented by a different vector $|g_N\rangle$, $|h_N\rangle$ etc. in this space, with components $g_N(x_i)$, $h_N(x_i)$ along the axes $|x_i\rangle$. These vectors form a vector space.

We can define an inner product on this vector space as

$$\langle f_N | g_N \rangle = \sum_{i=1}^{n} f_N(x_i)\, g_N(x_i)$$

(If the functions $f(x)$ were complex, i.e. our vector space were $\mathbb{V}^N(\mathbb{C})$, we would generalize this to $\langle f_N | g_N \rangle = \sum_{i=1}^{N} f_N^*(x_i) g_N(x_i)$.)

Two functions are called <u>orthogonal</u> if $\langle f_N | g_N \rangle = 0$.

What happens if, instead of discretising the function at a finite (although possibly very large) number of points x_i, we want to specify the function exactly, i.e. at every real value of x? The set of points $0 \le x \le L$ is uncountably infinite. Correspondingly our vectors $|f\rangle \equiv |f_\infty\rangle$ would live in an infinite-dimensional vector space. Each vector $|f\rangle$ in this space would describe a different function $f(x)$ in the infinite-dimensional basis $\{|x\rangle \mid 0 \le x \le L\}$, $f(x) = \langle x | f \rangle$, and vice versa. Note how $|f\rangle$ stands for the abstract function, whereas its basis representation ("components") $\langle x | f \rangle = f(x)$ gives the concrete values of this function at each x. This is like when we talk conceptually about "the exponential function $e^{\cdot}$", without working out its value at any specific value of x.

If we take the above definition for the inner product which says $\langle f_N | f_N \rangle = \sum_{i=1}^{N} \left(f_N(x_i) \right)^2$, and let $N \to \infty$, we have a problem because the sum blows up. Clearly, what we should do, as we subdivide the interval

$0 \le x \le L$ ever more finely, is that we should weight each term with the decreasing bin-width $\Delta_N = \frac{L}{N}$:

$$\langle f_N | g_N \rangle = \Delta_N \sum_{i=1}^{N} f_N(x_i)\, g_N(x_i)$$

As we let $N \to \infty$ this reduces to the standard definition of the integral:

$$\boxed{\begin{aligned} \langle f | g \rangle &= \int_0^L f(x)\, g(x)\, dx \\ \langle f | f \rangle &= \int_0^L f^2(x)\, dx \quad (\text{"norm" of } f) \end{aligned}}$$

$\left(\int_0^L f^*(x)\, g(x)\, dx \text{ for complex fcts.}\right)$

$\left(\int_0^L |f(x)|^2 dx \text{ for complex } f\right)$

What about the basis vectors $|x\rangle$? We want orthogonality,

$$\langle x | x' \rangle = 0 \quad \text{for } x \neq x'$$

and normalization. But how do we normalize? $\langle x | x \rangle = 1$ does not work (we'll see in a minute why not). Let's <u>derive</u> the correct answer. To do so we start with the other necessary requirement of completeness. Let us now work with complex function on the interval $[a, b]$, such that

$$\langle f | g \rangle = \int_a^b f^*(x)\, g(x)\, dx \qquad (\text{inner product}), \quad f(x) = \langle x | f \rangle$$

$$\langle x | x' \rangle = 0 \quad (x \neq x') \qquad (\text{orthogonality})$$

and

$$\int_0^b |x'\rangle\langle x'|\, dx' = \hat{I} \qquad (\text{completeness})$$

Let us look at $f(x) = \langle x | f \rangle$:

$$f(x) = \langle x | f \rangle = \langle x | \hat{I} | f \rangle = \int_a^b \langle x | x' \rangle \langle x' | f \rangle\, dx' = \int_a^b \underbrace{\langle x | x' \rangle}_{\equiv \delta(x, x')} f(x')\, dx'$$

So $\langle x|x'\rangle$ must be a "function" that picks out from the integral over $f(x')$ only the value of f at point $x'=x$. $\delta(x,x')=1$ at $x=x'$ is, however, not strong enough to do so, since $x=x'$ is a point of zero measure. To see what we need, let's zero in on the region $x \approx x'$ where $\delta(x,x')$ obviously must act:

$$\lim_{\varepsilon \to 0} \int_{x-\varepsilon}^{x+\varepsilon} \delta(x,x')\, f(x')\, dx' = f(x)$$

(The rest of the integral doesn't contribute anyway since $\delta(x,x')=0$ when $x \neq x'$.) In the infinitesimal region $x-\varepsilon \le x' \le x+\varepsilon$, any reasonably smooth f can be approximated by its value at $x'=x$ and pulled out from the integral:

$$f(x) \int_{x-\varepsilon}^{x+\varepsilon} \delta(x,x')\, dx' = f(x) \Rightarrow \int_{x-\varepsilon}^{x+\varepsilon} \delta(x,x')\, dx' = 1$$

Clearly $\delta(x,x')$ cannot be finite at $x=x'$, since then its integral over an <u>infinitesimal interval</u> would also be <u>infinitesimal</u>. So $\delta(x,x')$ must be infinite at $x=x'$ and zero at $x \neq x'$ in such a way that its integral is finite and exactly 1. The "function" (mathematicians call it a "distribution") $\delta(x-x')$ (it clearly depends only on the difference between x and x') with properties

(1) $\delta(x-x') = 0$ for $x \neq x'$

(2) $\int_a^b dx'\, \delta(x-x') = 1$ for $a<x<b$

(3) $\int_a^b dx'\, f(x')\, \delta(x-x') = f(x)$ for $a<x<b$

is known as the "<u>Dirac δ-function</u>". The basis vectors

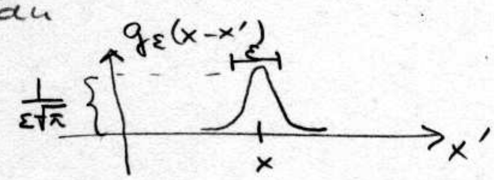
$|x\rangle$ must thus be orthonormalized as

$$\boxed{\langle x | x' \rangle = \delta(x - x')}$$

This will be needed any time the basis kets are labeled by a continuous index such as x. The δ-fct. is defined only in the context of integration by eq. (3) above (Eq. (2) is actually redundant – it follows from (3) for the special case $f(x) = 1 \ \forall x$.)

You may feel a little more comfortable with this weird distribution when you realize that you can think of it as the limit of a certain class of smooth, normal functions: Consider a Gaussian

$$g_\varepsilon(x - x') = \frac{1}{\sqrt{\pi \varepsilon^2}} e^{-\frac{(x-x')^2}{\varepsilon^2}}$$



As $\varepsilon \to 0$, this Gaussian gets narrower (i.e. more sharply peaked), while its peak value $g_\varepsilon(0) = \frac{1}{\varepsilon\sqrt{\pi}}$ gets bigger and bigger. The area under the Gaussian is 1 independent of ε. Once ε is so small that the function $f(x)$ no longer varies appreciably over the range where $g_\varepsilon(x-x')$ is significantly non-zero, we can pull $f(x)$ out of the integral and obtain

$$\lim_{\varepsilon \to 0} \int_a^b g_\varepsilon(x-x') f(x') dx' = \lim_{\varepsilon \to 0} f(x) \int_a^b g_\varepsilon(x-x') dx'$$

$$= f(x) \underbrace{\int_a^b g_\varepsilon(x-x') dx}_{=1 \text{ indep. of } \varepsilon} = f(x) \quad \text{as desired.}$$

The Gaussian $g_\varepsilon(x-x')$ is by no means the only smooth function that converges to $\delta(x-x')$ as $\varepsilon \to 0$. We will encounter many other representations of the Dirac δ.

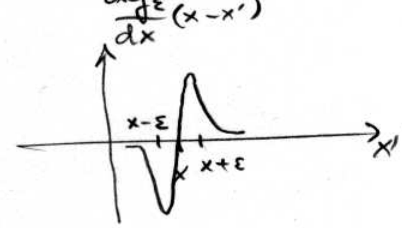
The δ-function is even:

$$\delta(x-x') = \langle x|x'\rangle = \langle x'|x\rangle^* = \left(\delta(x'-x)\right)^* = \delta(x'-x)$$

since it is real.

We can define the derivative of the δ-function

$$\delta'(x-x') \equiv \frac{d}{dx}\delta(x-x') = -\frac{d}{dx'}\delta(x-x')$$



as the limit of $\frac{d}{dx} g_\varepsilon(x-x')$ for $\varepsilon \to 0$;

In that limit, it basically becomes the difference between two δ-fcts. at $x+\varepsilon$ and $x-\varepsilon$:

$$\int \delta'(x-x')\, f(x')\, dx' \sim f(x+\varepsilon) - f(x-\varepsilon) \xrightarrow[\varepsilon\to 0]{} 2\varepsilon \frac{df}{dx'}\Big|_{x'=x}$$

The constant of proportionality works out as $\frac{1}{2\varepsilon}$, hence

$$\boxed{\int \delta'(x-x')\, f(x')\, dx' = \frac{df}{dx'}\Big|_{x=x'} = \frac{df}{dx}(x)} \qquad (*)$$

We can also get this by integration by parts:

$$\int_a^b \delta'(x-x')\, f(x')\, dx' = -\int_a^b \left[\frac{d}{dx'}\delta(x-x')\right] f(x')\, dx'$$

$$= \underbrace{-\delta(x-x')\, f(x')\Big|_{x'=a}^{x'=b}}_{0} + \int_a^b \delta(x-x')\, \frac{df}{dx'}(x')\, dx' = \frac{df}{dx'}\Big|_{x'=x}$$

$$= \frac{df(x)}{dx}$$

We can write eq. (*) as an operational prescription:

$$\boxed{\delta'(x-x') = \delta(x-x')\frac{d}{dx'}}$$

that applies when integrating it with any function (or combination of functions) $f(x')$ over an interval that includes the point x. Here $\frac{d}{dx'}$ acts as derivative operator on any function that accompanies the δ-fct. in the integrand.

Similarly for higher derivatives:

$$\boxed{\frac{d^n}{dx^n}\delta(x-x') = \delta(x-x')\frac{d^n}{dx^n}}$$

Another representation of $\delta(x-x')$:

Look at the Fourier transform of a function $f(x)$:

$$f(k) = \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} f(x) e^{-ikx}$$

and its inverse

$$f(x') = \int_{-\infty}^{\infty} \frac{dk}{\sqrt{2\pi}} e^{ikx'} f(k)$$

Insert the first into the second:

$$f(x') = \int_{-\infty}^{\infty} \frac{dk}{\sqrt{2\pi}} e^{ikx'} \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} e^{-ikx} f(x)$$

$$= \int_{-\infty}^{\infty} dx\, f(x) \left(\int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ik(x-x')} \right)$$

Clearly the factor in bracket acts exactly like a δ-function: It is even (switch $x \leftrightarrow x'$ and change integration variable $k \to -\tilde{k}$), satisfies condition (3) on p. 53 for <u>any</u> continuous function $f(x)$ and can be shown to vanish for $x \neq x'$.

Some more properties of $\delta(x-x')$:

(4) $\delta(ax) = \frac{\delta(x)}{|a|}$ (exercise!) (Consider $\int d(ax)\,\delta(ax)$ and use $\delta(x)=\delta(-x)$)

(5) $\delta(g(x)) = \sum_i \frac{\delta(x_i - x)}{|dg/dx|_{x=x_i}}$ where x_i are the zeroes of $g(x)$

(exercise!)

See Sec. 6.1.3, p. 295-299 in S.M. Lea, Math. for Physicists (Physics 834 web site)

<u>Example for (5)</u>: consider

$$I = \int \delta(\underbrace{E^2 - E_p^2}_{g(E)})\,F(E)\,dE = \int \delta\big((E+E_p)(E-E_p)\big)\,F(E)\,dE$$

The function $g(E)$ has 2 zeroes at $E = E_p$ and $E = -E_p$.

The derivative $\frac{dg}{dE} = 2E$. So we can split the integral in two parts:

$$I = \lim_{\varepsilon \to 0} \int_{-E_p-\varepsilon}^{-E_p+\varepsilon} F(E)\,\delta\big((E+E_p)\underbrace{(E-E_p)}_{\approx -2E_p}\big)\,dE + \lim_{\varepsilon \to 0} \int_{E_p-\varepsilon}^{E_p+\varepsilon} F(E)\,\delta\big(\underbrace{(E+E_p)}_{\approx 2E_p}(E-E_p)\big)\,dE$$

Near the zeroes of g, we can approximate $g(x)$ as a linear function, as indicated by $\approx$. Thus

$$I = \lim_{\varepsilon \to 0} \left[\int_{-E_p-\varepsilon}^{-E_p+\varepsilon} F(-E_p)\,\delta\big(\underbrace{-2E_p}_{\text{"a"}}(E+E_p)\big)\,dE + \int_{E_p-\varepsilon}^{E_p+\varepsilon} F(E_p)\,\delta\big(\underbrace{2E_p}_{\text{"a"}}(E-E_p)\big)\,dE \right]$$

$$\overset{(4)}{=} \frac{F(-E_p)}{|-2E_p|} \underbrace{\int \delta(E+E_p)\,dE}_{1} + \frac{F(E_p)}{|2E_p|} \underbrace{\int \delta(E-E_p)\,dE}_{1}$$

This is exactly the r.h.s. of (5), integrated over E. ◊