Functions as vectors

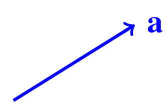
All problems in life can be solved with linear algebra. (almost)

Prereqs: vectors, matrices, calculus.
Unnecessary[1] rigor is avoided.

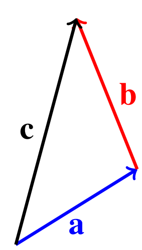
1 Vectors

1.1 What are they?

A picture is worth a thousand words.



Vectors can be added.



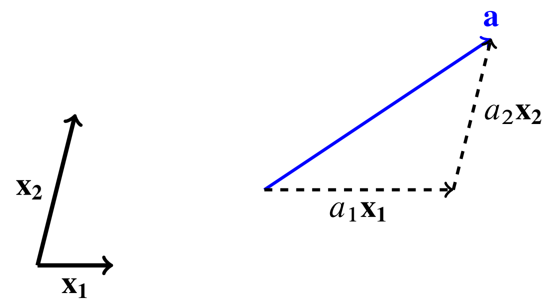
$\mathbf{c} = \mathbf{a} + \mathbf{b}$

Vectors exists independently of any coordinate system.
Now pick a coordinate system (pick a *basis*).

Basis $\{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}, \ldots, \mathbf{x_n}\}$

Representation $\mathbf{a} = a_1\mathbf{x_1} + a_2\mathbf{x_2} + \cdots + a_n\mathbf{x_n} = \sum_k a_k\mathbf{x_k} = (a_1, a_2, \ldots, a_n)$

[1] useless, intuition-inhibiting

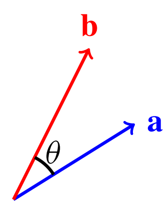


Different choices of basis vectors lead to different *representations*. Once a basis is chosen, the representation of any vector is unique.

1.2 Dot product for real vectors

Some properties:

- $\mathbf{a} \cdot \mathbf{a} > 0$ unless $\mathbf{a} = 0$ in which case $\mathbf{a} \cdot \mathbf{a} = 0$
- $\alpha \mathbf{a} \cdot \mathbf{b} = \mathbf{a} \cdot \alpha \mathbf{b} = \alpha(\mathbf{a} \cdot \mathbf{b})$
- $\mathbf{a} \cdot (\mathbf{b_1} + \mathbf{b_2}) = \mathbf{a} \cdot \mathbf{b_1} + \mathbf{a} \cdot \mathbf{b_2}$



Geometric interpretations: length or *norm* of $\mathbf{a}$ is $|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$. $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}| \cos\theta$ (projection of $\mathbf{a}$ onto $\mathbf{b}$ times length of $\mathbf{b}$, or vice versa).

1.3 Orthonormal basis

It is always possible to find a basis where all the basis vectors have length 1 and are *orthogonal* (perpendicular) to one another. A basis $\{\mathbf{x_1}, \ldots, \mathbf{x_n}\}$ where

$$\mathbf{x_i} \cdot \mathbf{x_j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{1.1}$$

is an *orthonormal basis*.

Dot products: let $\mathbf{a} = a_1\mathbf{x_1} + \cdots + a_n\mathbf{x_n}$ and $\mathbf{b} = b_1\mathbf{x_1} + \cdots + b_n\mathbf{x_n}$ be vectors and their representations in an orthonormal basis.

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \tag{1.2}$$

using (1.1).

We will only use orthonormal basis from now on, since it is always possible to convert a basis into an orthonormal one.

1.4 Matrix multiplication

$$C = \begin{pmatrix} c_{11} & \dots & c_{1j} & \dots & c_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{i1} & \dots & c_{ij} & \dots & c_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{m1} & \dots & c_{mj} & \dots & c_{mn} \end{pmatrix}$$

Suppose we have matrices A and B and we want to compute their product $C = AB$. Matrix multiplication is defined as

$$c_{ij} = \sum_k a_{ik} b_{kj} \tag{1.3}$$

If we used an orthonormal basis and we wanted to compute the dot product $\mathbf{a} \cdot \mathbf{b}$, we could put their representations into matrices and compute

$$\mathbf{a} \cdot \mathbf{b} = \begin{pmatrix} a_1 \dots a_n \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \cdots + a_n b_n \tag{1.4}$$

1.5 Dirac's notation

Dirac's notation for vectors distinguishes between row vectors and column vectors. Row vector:

$$\langle a| = \begin{pmatrix} a_1 \dots a_n \end{pmatrix} \tag{1.5}$$

Column vector:

$$|b\rangle = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \tag{1.6}$$

$\mathbf{a} \cdot \mathbf{b}$ is now written as $\langle a|b\rangle$.

In old notation, $\mathbf{a} = a_1 \mathbf{x_1} + \cdots + a_n \mathbf{x_n}$, and $a_k = \mathbf{x_k} \cdot \mathbf{a}$. In Dirac's notation, $|a\rangle = a_1 |x_1\rangle + \cdots + a_n |x_n\rangle$, and $a_k = \langle x_k|a\rangle$. We see

$$\begin{aligned} |a\rangle &= a_1 |x_1\rangle + \cdots + a_n |x_n\rangle \\ &= |x_1\rangle \langle x_1|a\rangle + \cdots + |x_n\rangle \langle x_n|a\rangle \\ &= (|x_1\rangle \langle x_1| + \cdots + |x_n\rangle \langle x_n|) |a\rangle \end{aligned}$$

from which we conclude

$$\begin{aligned} 1 &= |x_1\rangle \langle x_1| + |x_2\rangle \langle x_2| + \cdots + |x_n\rangle \langle x_n| \\ &= \sum_k |x_k\rangle \langle x_k| \end{aligned} \tag{1.7}$$

1.6 Complex vectors

When we allow vectors to be multiplied by complex numbers, we still want to keep the idea of the length of a vector. For real vectors, this was $\sqrt{\langle a|a\rangle}$, and we would like the same definition for complex vectors. But $a_1^2 + a_2^2 + \cdots + a_n^2$ is not guaranteed to be real, and we would like length to be a nonnegative real number. However, the quantity $|a_1|^2 = \overline{a_1}a_1$ is always real and nonnegative, where $\overline{a_1}$ refers to the complex conjugate of a_1. So we extend our definition of dot product to complex vectors by doing this:

$$\langle a|b\rangle = \overline{a_1}b_1 + \cdots + \overline{a_n}b_n \tag{1.8}$$

If the a_k's were real numbers, $\overline{a_k} = a_k$ and we recover our old formula. But now $\langle a|a\rangle$ is guaranteed to be a nonnegative real number for all complex vectors.

So now we have the following rule:

$$\text{If } |a\rangle = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \text{ then } \langle a| = \begin{pmatrix} \overline{a_1} \ldots \overline{a_n} \end{pmatrix} \tag{1.9}$$

and vice versa. Define $\overline{|a\rangle} = \langle a|$ and $\overline{\langle a|} = |a\rangle$. We see that

$$\begin{aligned} \overline{\langle a|b\rangle} &= \overline{\overline{a_1}b_1 + \cdots + \overline{a_n}b_n} \\ &= a_1\overline{b_1} + \cdots + a_n\overline{b_n} \\ &= \langle b|a\rangle \end{aligned} \tag{1.10}$$

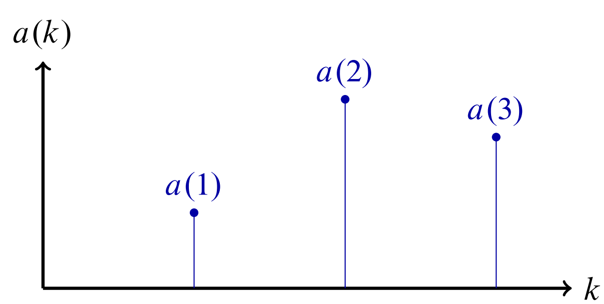
so *complex conjugation of a product is complex conjugation of individual parts of the product but everything written backwards.*[2]

2 Functions

2.1 Vectors as functions

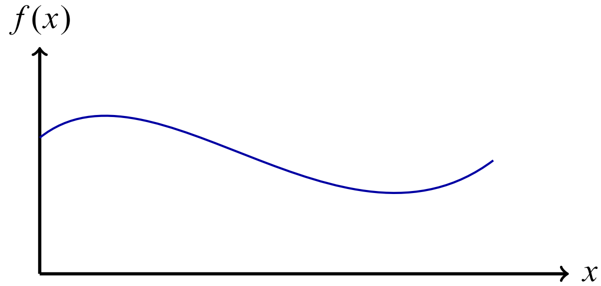
A function looks like f. A value, $f(x)$ is assigned to all x in the domain. $f(1.2) =$ blah, $f(\pi) =$ blahblah, etc.

Compare with vector representations: $a_1 =$ blah, $a_2 =$ blahblah. A value is assigned to every index. We can construct a function a where $a(1) = a_1$, $a(2) = a_2$, etc. So the function a is the representation of the vector $|a\rangle$ in some basis.



[2]Usually this is referred to as *adjoint* or *conjugate transpose* and written with † rather than a bar.

2.2 Functions as vectors



Now the reverse. From the function f we can construct a vector $|f\rangle$, so f is the representation of $|f\rangle$. The analogy:

$$
\begin{array}{ll}
\quad\quad |a\rangle & \quad\quad |f\rangle \\
\langle x_2|a\rangle = a_2 & \langle x_2|f\rangle = f(2) \\
\langle a|b\rangle = \sum_k \overline{a_k} b_k & \langle f|g\rangle = ???
\end{array}
$$

The problem is that the domain of f is continuous, so we can't just form a sum for $\langle f|g\rangle$ as we did for $\langle a|b\rangle$ because it would be an infinite sum that in general does not converge. Instead we shall do this

$$\langle f|g\rangle = \int \overline{f(x)} g(x)\, dx \tag{2.1}$$

This preserves the spirit of the dot product, the only change being the weight dx which converts infinite quantities to finite.

Some notation notes: $\langle x_2|f\rangle = f(2)$, $\langle x_\pi|f\rangle = f(\pi)$, etc., but to avoid the awkward notation of $\langle x_x|f\rangle = f(x)$, we shall just write $\langle x|f\rangle = f(x)$.

Continuing with (2.1)

$$
\begin{aligned}
\langle f|g\rangle &= \int \overline{f(x)} g(x)\, dx \\
&= \int \langle f|x\rangle \langle x|g\rangle \; dx \\
&= \langle f| \left(\int |x\rangle \langle x| \; dx \right) |g\rangle
\end{aligned}
$$

we see that

$$1 = \int |x\rangle \langle x| \; dx \tag{2.2}$$

analogous to (1.7).

2.3 Dirac's delta function

What is $\langle x|x_\pi\rangle$ or $\langle x|x_1\rangle$? It's going to be a function, just like $\langle x|f\rangle = f(x)$ is a function. So let's just write $\delta_\pi(x) = \langle x|x_\pi\rangle$, $\delta_2(x) = \langle x|x_2\rangle$, etc.

Using (2.2)

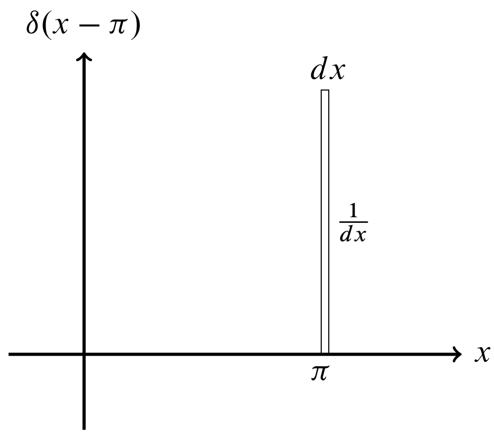
$$\begin{aligned}\overline{f(\pi)} = \langle f|x_\pi\rangle &= \int \langle f|x\rangle \langle x|x_\pi\rangle \; dx \\ &= \int \overline{f(x)}\delta_\pi(x)\, dx\end{aligned}$$

Since this works for any f, we must have [3]

$$\langle x|x_\pi\rangle = \delta_\pi(x) = \begin{cases} \frac{1}{dx} & \text{for } \pi - \frac{dx}{2} \le x \le \pi + \frac{dx}{2} \\ 0 & \text{everywhere else} \end{cases}$$

How small/What is dx? dx shall be a quantity that is smaller than the precision of your data collecting tools (ruler, weight scale, voltmeter, experimental apparati) or if you're thinking about it theoretically, smaller than anything you can imagine.

Notation notes: Instead of writing $\delta_\pi(x)$ or $\delta_2(x)$, we can just write $\delta_0(x-\pi)$ and $\delta_0(x-2)$. Then we can just drop the subscript 0 and just write $\delta(x)$ to mean $\delta_0(x)$ and $\langle x|x_\pi\rangle = \delta(x-\pi) = \delta_\pi(x)$, etc.



2.4 The differential operator

So we've established the notation $f(x) = \langle x|f\rangle$. How are we going to write $\frac{df}{dx} = f'(x)$? $f'(\pi) = (f(\pi + dx) - f(\pi))/dx$, which is $(\langle x_{\pi+dx}|f\rangle - \langle x_\pi|f\rangle)/dx$. Take a look at (2.3), where we have let D be the giant matrix

[3] Unnecessary rigor is avoided. It is possible to avoid defining the δ function completely and continue on without it, but the analogy with vectors in finite dimensional spaces is lost. Intuition is more important than rigor, and the δ function is a very useful idea.

$$
\begin{aligned}
D\,|a\rangle &= \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} \\
&= \begin{pmatrix} a_2 - a_1 \\ a_3 - a_2 \\ a_4 - a_3 \\ a_5 - a_4 \\ a_5 \end{pmatrix}
\end{aligned} \tag{2.3}
$$

Apart from the last element of $D\,|a\rangle$, all other elements are of the form $\langle x_k|\, D\,|a\rangle = \langle x_{k+1}|a\rangle - \langle x_k|a\rangle$. Comparing with $f'(\pi) = (\langle x_{\pi+dx}|f\rangle - \langle x_\pi|f\rangle)/dx$, we can imagine $\frac{d}{dx}$ to be a matrix that looks kinda like D in (2.3).

$$
f'(x) = \langle x|\, \frac{d}{dx}\, |f\rangle \tag{2.4}
$$

What about derivative of the delta function, like $\langle x|\, \frac{d}{dx}\, |x_\pi\rangle = \delta'(x-\pi)$? It's actually more useful to look at dot products involving the delta function's derivative.

$$
\begin{aligned}
\langle f|\, \frac{d}{dx}\, |x_\pi\rangle &= \int \langle f|x\rangle\, \langle x|\, \frac{d}{dx}\, |x_\pi\rangle\; dx \\
&= \int \overline{f(x)}\delta'(x-\pi)\, dx
\end{aligned}
$$

This can be integrated by parts.

$$
= \left[\overline{f(x)}\delta(x-\pi)\right] - \int \overline{f'(x)}\delta(x-\pi)\, dx
$$

The delta function is 0 at the limits of integration so

$$
\begin{aligned}
&= -\int \overline{f'(x)}\delta(x-\pi)\, dx \\
&= -\overline{f'(\pi)} \\
&= -\overline{\langle x_\pi|\, \frac{d}{dx}\, |f\rangle} \\
&= -\,\langle f|\, \overline{\frac{d}{dx}}\, |x_\pi\rangle
\end{aligned}
$$

So we have established the following 2 relations

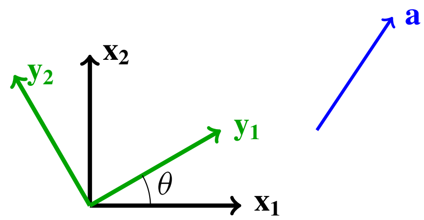
$$
\langle f|\, \frac{d}{dx}\, |x_\pi\rangle = -\overline{f'(\pi)} \tag{2.5}
$$

and

$$
\overline{\frac{d}{dx}} = -\frac{d}{dx} \tag{2.6}
$$

3 Change of basis

3.1 Rotations



Suppose we know the representation of a vector $|a\rangle$ in the x-basis and we want to know the representation in a rotated y-basis. We can use (1.7) to compute

$$\langle y_1|a\rangle = \sum_k \langle y_1|x_k\rangle \langle x_k|a\rangle \tag{3.1}$$

$$\langle y_2|a\rangle = \sum_k \langle y_2|x_k\rangle \langle x_k|a\rangle \tag{3.2}$$

This is just matrix multiplication.

$$\begin{pmatrix} \langle y_1|a\rangle \\ \langle y_2|a\rangle \end{pmatrix} = \begin{pmatrix} \langle y_1|x_1\rangle & \langle y_1|x_2\rangle \\ \langle y_2|x_1\rangle & \langle y_2|x_2\rangle \end{pmatrix} \begin{pmatrix} \langle x_1|a\rangle \\ \langle x_2|a\rangle \end{pmatrix}$$

The matrix $\langle y_i|x_j\rangle$ is the *transformation matrix*. The columns tell you the representation of each $|x_j\rangle$ in the y-basis. For rotations, we see from the picture that the transformation matrix is

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

3.2 Functions

Suppose we know $f(x) = \langle x|f\rangle$ and we have a new basis $|y\rangle$. The representation of $|f\rangle$ in the new basis will be a new function $\tilde{f}(y) = \langle y|f\rangle$. We can use (2.2) to compute

$$\begin{aligned} \tilde{f}(y) = \langle y|f\rangle &= \int \langle y|x\rangle \langle x|f\rangle \; dx \\ &= \int \langle y|x\rangle \, f(x) \, dx \end{aligned} \tag{3.3}$$

We have the transformation matrix $\langle y|x\rangle$, a function of two variables.

3.3 The Fourier basis

A particularly important class of basis changings is the Fourier transforms. These transforms all have the flavor of analyzing periodic things and frequency and sine waves and stuff and their transformation matrices all involve roots of unity.[4]

[4]From the point of view of groups, they are characters of a cyclic group. The concept of Fourier transforms can be generalized to more than just cyclic groups.

3.3.1 Discrete Fourier transform (DFT)

Suppose we have a vector $|a\rangle$ in an n-dimensional space and we know its representation in the x-basis. We define the y-basis with the transformation matrix

$$\langle y_j | x_k \rangle = \frac{1}{\sqrt{n}} \exp\left(\frac{-2\pi i}{n} kj\right) \tag{3.4}$$

The representation of $|a\rangle$ in the y-basis is given by

$$\begin{aligned} \tilde{a}_j = \langle y_j | a \rangle &= \sum_k \langle y_j | x_k \rangle \langle x_k | a \rangle \\ &= \frac{1}{\sqrt{n}} \sum_k \exp\left(\frac{-2\pi i}{n} kj\right) a_k \end{aligned} \tag{3.5}$$

which is the *discrete Fourier transform* of $(a_1, \ldots, a_n)$.

3.3.2 Fourier transform

Suppose we have a function $f(x)$, which is the representation of $|f\rangle$ in the x-basis. Define the y-basis[5] with the transformation matrix

$$\langle y | x \rangle = \frac{1}{\sqrt{2\pi}} e^{ixy} \tag{3.6}$$

The representation of $|f\rangle$ in the y-basis is given by

$$\begin{aligned} \tilde{f}(y) = \langle y | f \rangle &= \int \langle y | x \rangle \langle x | f \rangle \, dx \\ &= \frac{1}{\sqrt{2\pi}} \int e^{ixy} f(x) \, dx \end{aligned} \tag{3.7}$$

which is the *Fourier transform* of the function $f(x)$.

4 Variational calculus as multivariable calculus

4.1 The flavor of variational calculus

4.2 Euler-Lagrange equation as a gradient

4.3 Application to physics

4.3.1 The Lagrangian and Newton's second law

Let $L(x(t), \dot{x}(t), t) = \frac{1}{2} m \dot{x}^2 - U(x)$. The Euler-Lagrange equation yields $F = m\ddot{x} = ma$, a.k.a. Newton's second law. The quantity (kinetic energy − potential energy) is the *Lagrangian*.

[5] A basis for a restricted class of functions

4.3.2 The Hamiltonian and Hamilton's equations of motion

Computing the derivatives with chain rule and product rule

$$\begin{aligned} \frac{dL}{dt} &= \dot{x}\frac{\partial L}{\partial x} + \ddot{x}\frac{\partial L}{\partial \dot{x}} + \frac{\partial L}{\partial t} \\ \frac{d}{dt}\left(\dot{x}\frac{\partial L}{\partial \dot{x}}\right) &= \ddot{x}\frac{\partial L}{\partial \dot{x}} + \dot{x}\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} \end{aligned}$$

We can form

$$\frac{dL}{dt} - \frac{d}{dt}\left(\dot{x}\frac{\partial L}{\partial \dot{x}}\right) = \dot{x}\left(\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right) + \frac{\partial L}{\partial t}$$

which by Euler-Lagrange becomes

$$\frac{d}{dt}\left(L - \dot{x}\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial t}$$

The quantity $\dot{x}\frac{\partial L}{\partial \dot{x}} - L$ is the *Hamiltonian* (denoted H), which we see is a constant of motion as long as $\frac{\partial L}{\partial t} = 0$, i.e. L does not involve time explicitly.

Define $p = \frac{\partial L}{\partial \dot{x}}$, the *canonical momentum*. The Hamiltonian is $H = \dot{x}p - L$. Looking at its differential

$$\begin{aligned} dH &= p\,d\dot{x} + \dot{x}\,dp - dL \\ &= p\,d\dot{x} + \dot{x}\,dp - \left(\frac{\partial L}{\partial x}\,dx + p\,d\dot{x} + \frac{\partial L}{\partial t}\,dt\right) \\ &= \dot{x}\,dp - \frac{\partial L}{\partial x}\,dx - \frac{\partial L}{\partial t}\,dt \\ &= \dot{x}\,dp - \dot{p}\,dx - \frac{\partial L}{\partial t}\,dt \end{aligned}$$

where in the last step we have used Euler-Lagrange ($\dot{p} = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \frac{\partial L}{\partial x}$). We can read off some of H's partial derivatives

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial p} \\ \dot{p} &= -\frac{\partial H}{\partial x} \end{aligned} \tag{4.1}$$

which are *Hamilton's equations of motion*.

4.3.3 The Poisson bracket

Suppose we have a dynamical variable $b(x, p)$, which could be kinetic energy, position, momentum, angular momentum (if you extend this to 3 dimensions), whatever. Knowing the initial value of b

and $\frac{db}{dt}$ for all time would solve physics.

$$\begin{aligned}\frac{db}{dt} &= \frac{\partial b}{\partial x}\dot{x} + \frac{\partial b}{\partial p}\dot{p} \\ &= \frac{\partial b}{\partial x}\frac{\partial H}{\partial p} - \frac{\partial b}{\partial p}\frac{\partial H}{\partial x}\end{aligned}$$

using (4.1).

Define the *Poisson bracket*

$$[m,n] = \frac{\partial m}{\partial x}\frac{\partial n}{\partial p} - \frac{\partial n}{\partial x}\frac{\partial m}{\partial p} \tag{4.2}$$

Then we have

$$\frac{db}{dt} = [b, H] \tag{4.3}$$

If you know the Hamiltonian of the system and you know the initial value of any dynamical variable b (that is explicitly independent of time), you know b for all time (also see (4.5)). In other words, you win.

4.3.4 Quantum Poisson bracket and Schrödinger's equation

The passage of classical mechanics to quantum mechanics would require its own textbook. This section will therefore not make any sense except to people already familiar with basic quantum mechanics.

For simplicity, let's consider an isolated system (constant H). A nonconstant H would just mean we didn't take some interaction in our system into account properly.

Define the *quantum Poisson bracket* of two dynamical variables m and n[6]

$$[m,n] = \frac{mn - nm}{i\hbar} \tag{4.4}$$

where $\hbar$ is a constant of nature. As an experimentally verified fact of life (physical law), if we have a dynamical variable b explicitly independent of time (in the sense that its evolution is only a result of internal interactions)

$$\frac{db}{dt} = [b, H] \tag{4.5}$$

which is identical in form to (4.3) except we use the quantum Poisson bracket here.

Denote b_0 to be the initial value of b and b_t to be the value after time t.[7] It can be shown that b_t is related to b_0 by a unitary change of basis $T(t)$ independent of b (see Appendix A).

$$b_t = T^{-1}b_0T = \overline{T}b_0T \tag{4.6}$$

[6]Usually the notation $[m,n]$ refers to the commutator $mn - nm$. Usually, but not here.

[7]b_t is written instead of $b(t)$ to emphasize that b changes only as a result of physical interactions rather than mathematical stuffies

Plugging back into (4.5)

$$\frac{d\overline{T}}{dt}b_0T + \overline{T}b_0\frac{dT}{dt} = \frac{\overline{T}b_0TH}{i\hbar} - \frac{H\overline{T}b_0T}{i\hbar}$$

which can be written as

$$\overline{\overline{T}b_0\frac{dT}{dt}} + \overline{T}b_0\frac{dT}{dt} = \frac{\overline{\overline{T}b_0HT}}{i\hbar} + \frac{\overline{T}b_0HT}{i\hbar} \tag{4.7}$$

From (4.7) we can identify $i\hbar\frac{dT}{dt} = HT$ (see Appendix A for a less shaky approach; a counterexample to the argument $a + \overline{a} = b + \overline{b} \implies a = b$ is $0 + \overline{0} = i + \overline{i}$). If we write

$$T|\psi\rangle = |\psi_t\rangle \tag{4.8}$$

then

$$\begin{aligned} i\hbar\frac{dT}{dt}|\psi\rangle &= HT|\psi\rangle \\ i\hbar\frac{d}{dt}|\psi_t\rangle &= H|\psi_t\rangle \end{aligned} \tag{4.9}$$

which is the *Schrödinger equation*.

Last modified: Bryance Oyang, September 28, 2013

Appendix A Unhandwaiving the handwaiving

Proof by appendix.

Appendix B But seriously

We're trying to get the Schrödinger equation (assuming a constant H), expanding on an analogy with the classical equation of motion (4.3).

This is best done using diagrams (as is anything involving tensor products and products of tensors) but currently I don't feel like drawing them. So here it is using math symbols.

Start with (4.5).

$$\frac{db}{dt} = \frac{1}{i\hbar}(bH - Hb)$$

We can reshape b (a matrix) into a vector (sorta like MATLAB's `reshape` command). Then we'll have something like

$$\frac{db}{dt} = \frac{1}{i\hbar}(1 \otimes H^{\intercal} - H \otimes 1)b$$

where $H^\intercal$ is the transpose of H and $\otimes$ is tensor product.

We can recognize this as a differential equation of the form $\frac{d\mathbf{v}}{dt} = M\mathbf{v}$ with solution $\mathbf{v} = \exp(Mt)\mathbf{v_0}$.

$$\begin{aligned} b_t &= \exp\left(\frac{t}{i\hbar}(1 \otimes H^\intercal - H \otimes 1)\right) b_0 \\ &= \exp\left(\frac{t}{i\hbar} 1 \otimes H^\intercal\right) \exp\left(\frac{-t}{i\hbar} H \otimes 1\right) b_0 \end{aligned}$$

as $1 \otimes H^\intercal$ and $H \otimes 1$ commute. This can be written as

$$b_t = \exp\left(\frac{-t}{i\hbar} H\right) \otimes \exp\left(\frac{t}{i\hbar} H^\intercal\right) b_0$$

Reshaping this back into the original form where b was a matrix

$$\begin{aligned} b_t &= \exp\left(\frac{-t}{i\hbar} H\right) b_0 \exp\left(\frac{t}{i\hbar} H\right) \\ &= \overline{T} b_0 T \end{aligned}$$

letting $T = \exp\left(\frac{t}{i\hbar} H\right)$. We recognize T as the time evolution operator. We can compute its derivative

$$\frac{dT}{dt} = \frac{1}{i\hbar} H T$$

after which (4.8)–(4.9) should be clear.

I don't know how to do this when H is time-dependent (external influences are present), but (4.5) would have to be modified.