Hilbert Spaces

An abstract Hilbert space $\mathcal{H}$ is a complex vector space — of any dimension, finite or infinite — equipped with a positive-definite sesquilinear dot product; in quantum mechanics, this dot product is called the Dirac product. The 'complex vector space' means that we can add up elements of $\mathcal{H}$ and multiply them by complex numbers, thus any complex linear combination of any any vectors is a valid vector,

$$\forall \, |\phi\rangle, |\xi\rangle \in \mathcal{H}, \quad \forall \, \alpha, \beta \in \mathbf{C}: \quad \alpha \, |\phi\rangle + \beta \, |\xi\rangle \in \mathcal{H}. \tag{1}$$

Physically, this is the *superposition principle* of quantum mechanics: any complex linear combination of any allowed quantum states is itself an allowed quantum state.

The *sesquilinear* dot product means linear in one vector and antilinear in the other, like $(\mathbf{n}_1, \mathbf{n}_2) \to \mathbf{n}_1^* \cdot \mathbf{n}_2$. In Dirac's bracket notations for QM, the complex conjugation of ordinary vectors becomes the Hermitian conjugation between the bra-vector and the ket-vector for the same state $(|\psi\rangle)^\dagger = \langle\psi|$ and $(\langle\psi|)^\dagger = |\psi\rangle$; this Hermitian conjugation includes the complex conjugation of the coefficients, thus

$$(\alpha \, |\phi\rangle + \beta \, |\xi\rangle)^\dagger \;=\; \alpha^* \, \langle\phi| \;+\; \beta^* \, \langle\xi|. \tag{2}$$

The Dirac bracket is a bilinear product of a bra and a ket, thus

$$\langle\psi| \, (\alpha \, |\phi\rangle + \beta \, |\xi\rangle) \;=\; \alpha \, \langle\psi|\phi\rangle \;+\; \beta \, \langle\psi|\xi\rangle,$$
$$\text{but} \quad (\alpha \, |\phi\rangle + \beta \, |\xi\rangle)^\dagger \, |\psi\rangle \;=\; (\alpha^* \, \langle\phi| + \beta^* \, \langle\xi|) \, |\psi\rangle \tag{3}$$
$$=\; \alpha^* \, \langle\phi|\psi\rangle \;+\; \beta^* \, \langle\xi|\psi\rangle.$$

Also, this Dirac product is symmetric modulo complex conjugation,

$$\langle\phi|\xi\rangle \;=\; \langle\xi|\phi\rangle^*, \tag{4}$$

and the product of any state with itself is not only real but also positive definite,

$$\forall \, |\psi\rangle \neq 0: \quad \||\psi\rangle\|^2 \;\stackrel{\text{def}}{=}\; \langle\psi|\psi\rangle \;>\; 0. \tag{5}$$

Examples of Hilbert spaces

- $\mathbf{C}^N$, the N-dimensional complex vector space comprised of ordered lists $(\phi_1, \ldots, \phi_N)$ of N complex numbers. The vector structure here is obvious,

$$\alpha(\phi_1, \ldots, \phi_N) \; + \; \beta(\xi_1, \ldots, \xi_n) \; = \; (\alpha\phi_1 + \beta\xi_1, \ldots, \alpha\phi_N + \beta\xi_N), \tag{6}$$

  while the Dirac product generalizes the $\mathbf{n}_1^* \cdot \mathbf{n}_2$ of 2d polarization vectors for the EM waves,

$$\langle\phi|\xi\rangle \; = \; \phi_1^*\xi_1 \; + \; \cdots \; + \; \phi_N^*\xi_N \,. \tag{7}$$

- For a more interesting example, let's take the dimension N to infinity and look at the space of infinite lists of complex numbers $(\psi_1, \psi_2, \ldots)$. In this space, the Dirac bracket becomes an infinite series

$$\langle\psi|\xi\rangle \; = \; \sum_{i=1}^{\infty} \psi_i^*\xi_i \,. \tag{8}$$

- Finally, we may replace infinite but discrete lists of complex numbers $(\psi_1, \psi_2, \ldots)$ with complex-valued *wave-functions* $\psi(x)$ of a continuous argument x. The superposition here works by taking linear combinations of wave functions at each x,

$$[\alpha\psi + \beta\xi](x) \; = \; \alpha\psi(x) \; + \; \beta\xi(x), \tag{9}$$

  while the Dirac bracket obtains as an integral

$$\langle\psi|\xi\rangle \; = \; \int dx \, \psi^*(x)\xi(x). \tag{10}$$

Convergence issues

In infinite-dimensional Hilbert spaces, the infinite series (8) or the integrals (10) for the Dirac products do not always converge. The mathematicians and the physicists have different ways to handle the potential divergences here:

The mathematicians' approach is to restrict the Hilbert space to the vectors of finite norm, $\langle\psi|\psi\rangle < \infty$. For example, the space of infinite lists $(\psi_i, \psi_2, \ldots)$ of complex numbers is restricted to the

$$\ell_2 = \left\{ (\psi_i, \psi_2, \ldots) \text{ such that } \sum_{i=1}^{\infty} |\psi_i|^2 < \infty \right\} \tag{11}$$

while the space of wave-functions of a real coordinate x is restricted to the $L_2(\mathbf{R})$, the space of functions $\psi(x)$ such that $|\psi(x)|^2$ is Lebesgue-integrable with a finite integral,

$$\int\limits_{-\infty}^{+\infty} dx\, |\psi(x)|^2 < \infty. \tag{12}$$

Given such a restriction on the allowed elements of a Hilbert space, a theorem guarantees that for any two states $|\psi\rangle$ and $|\xi\rangle$ of finite norms, $\langle\psi|\psi\rangle < \infty$ and $\langle\xi|\xi\rangle < \infty$, the series or the integral for their Dirac product $\langle\psi|\xi\rangle$ is absolutely convergent.

From the mathematician's point of view, the wave-functions of infinite norm $\langle\psi|\psi\rangle = \infty$ simply do not belong to the Hilbert space.

On the other hand, the physicists frequently use the un-normalizable wave-functions as convenient limiting cases. For example, the quantum state $|x = x_0\rangle$ of a particle localized at a point x_0 has the un-normalizable wave function $\psi(x) = \delta(x - x_0)$. Physically, this state is the limit of states where the particle is restricted to a very small interval between $x_0 - \epsilon$ and $x_0 + \epsilon$. For finite $\epsilon > 0$ the wave-function is normalizable and can be rescaled to a unit norm,

$$\psi_\epsilon(x) = \frac{1}{\sqrt{2\epsilon}} \begin{cases} 1 & \text{if } |x - x_0| < \epsilon, \\ 0 & \text{otherwise,} \end{cases} \tag{13}$$

but the $\epsilon \to 0$ limit of such normalized wave functions is too awkward to use. Instead,

we rescale the wave-functions by $1/\sqrt{2\epsilon}$ before taking the $\epsilon \to 0$ limit, which gives an un-normalizable *generalized* wave-function,

$$\psi(x) \;=\; \lim_{\epsilon\to 0} \frac{1}{\sqrt{2\epsilon}} \psi_\epsilon(x) \;=\; \lim_{\epsilon\to 0} \frac{1}{2\epsilon} \begin{cases} 1 & \text{if } |x-x_0| < \epsilon, \\ 0 & \text{otherwise,} \end{cases} \;=\; \delta(x-x_0). \tag{14}$$

Likewise, the state of definite momentum $|p = \hbar k\rangle$ has un-normalizable wave function $\psi(x) = \exp(ikx)$. Physically, this state — which is spread over the whole coordinate space — is a limit of wave packets of *almost*-definite momenta near $\hbar k$ but dieing off for $x \to \pm\infty$. For example, the normalized Gaussian wave packets

$$\psi_\alpha(x) \;=\; \sqrt{\frac{2\alpha}{\pi}}\, e^{ikx}\, e^{-\alpha x^2}. \tag{15}$$

Again, the $\alpha \to 0$ limit of the normalized (to $\langle\psi|\psi\rangle = 1$) wave functions is too awkward to use, so instead we take the un-normalized limit

$$\psi(x) \;=\; \lim_{\alpha\to 0} \sqrt{\frac{\pi}{2\alpha}} \times \psi_\alpha(x) \;=\; e^{ikx}. \tag{16}$$

The Dirac products of the un-normalizable wave-functions are sometimes finite, sometimes infinite, and when you are not sure how to handle the infinity, you should regularize. Here are some examples:

$$\langle x = x_1 | x = x_2\rangle \;=\; \int dx\, \delta(x-x_1) \times \delta(x-x_2) \;=\; \delta(x_1 - x_2), \tag{17}$$

$$\langle x = x_0 | p = \hbar k\rangle \;=\; \int dx\, \delta(x-x_0) \times e^{ikx} \;=\; e^{ikx_0}, \tag{18}$$

$$\langle p = \hbar k_1 | p = \hbar k_2\rangle \;=\; \int dx\, \left(e^{ik_1 x}\right)^* \times e^{ik_2 x} \;=\; \int dx\, e^{i(k_2 - k_1)x}$$

where the integral diverges but can be regularized to

$$=\; 2\pi\, \delta(k_1 - k_2) \;=\; 2\pi\hbar\, \delta(p_1 = p_2). \tag{19}$$

Bases

In the ordinary 3d space, a basis is a set of 3 vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ such that any other vector $\mathbf{x}$ can be decomposed into a linear combination

$$\mathbf{x} = \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} \quad \text{for some real } \alpha, \beta, \gamma \tag{20}$$

Of particular interest are the orthogonal bases — where $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are $\perp$ to each other. For such a basis

$$\alpha = \frac{\mathbf{x} \cdot \mathbf{a}}{\mathbf{a}^2}, \quad \beta = \frac{\mathbf{x} \cdot \mathbf{b}}{\mathbf{b}^2}, \quad \gamma = \frac{\mathbf{x} \cdot \mathbf{c}}{\mathbf{c}^2} \quad \Longrightarrow \quad \mathbf{x} = \left(\frac{\mathbf{x} \cdot \mathbf{a}}{\mathbf{a}^2}\right)\mathbf{a} + \left(\frac{\mathbf{x} \cdot \mathbf{b}}{\mathbf{b}^2}\right)\mathbf{b} + \left(\frac{\mathbf{x} \cdot \mathbf{c}}{\mathbf{c}^2}\right)\mathbf{c}. \tag{21}$$

The same principle applies to the Hilbert spaces. A basis of a Hilbert space $\mathcal{H}$ is a set $\{|i\rangle, \ i = 1, 2, \ldots\}$ of linearly independent vectors $|i\rangle \in \mathcal{H}$ such that *any* vector $\psi \in \mathcal{H}$ can be decomposed into a linear combination

$$|\psi\rangle = \sum_i \psi_i |i\rangle \quad \text{for some coefficients } \psi_i \in \mathcal{H}. \tag{22}$$

If the basis is orthogonal — *i.e.*, $\langle i|j\rangle = 0 \ \forall i \neq j$, — then $\psi_i = \langle i|\psi\rangle / \langle i|i\rangle$ and therefore

$$|\psi\rangle = \sum_i |i\rangle \frac{\langle i|\psi\rangle}{\langle i|i\rangle}. \tag{23}$$

Note: in a space of finite dimension $d < \infty$, any set of d linearly-independent vectors forms a complete basis. But this does not work for the infinitely-dimensional Hilbert spaces: for $d = \infty$, linear independence of an infinite set of vectors does not guarantee its completeness as a basis. Instead, we must test for completeness by *proving* that eq. (23) holds for any $|\psi\rangle \in \mathcal{H}$.

In quantum mechanics, a basis is usually made of quantum states where some measurable quantity — often called an *observable* — has a definite value. For example, in the 2d Hilbert space of silver atom's spin states, we may use the basis $\{|m_z = +m_B\rangle, |m_z = -m_B\rangle\}$ where the z component of the atom's magnetic moment $\mathbf{m}$ has a definite value $+m_B$ or $-m_B$. But we may just as well use a basis where some other component of $\mathbf{m}$ has a definite value, such as $\{|m_x = +m_B\rangle, |m_x = -m_B\rangle\}$ or $\{|m_y = +m_B\rangle, |m_y = -m_B\rangle\}$. However, all these bases are bases of the same Hilbert space.

Note: according to the Stern–Gerlach experiment, different components of the atom's magnetic moment are *incompatible observables* — they cannot both have definite values in the same quantum state, and any measurement of one observable would *unavoidably* mess up the value of the other. Consequently, the basis of states where one component has definite values is different from the basis where the other component has definite values, and the same rule applies to other incompatible observables: *For any quantum system, incompatible observables give rise to different bases of the same Hilbert space.*

The best-known example of incompatible observables are the position x and the momentum p of a moving particle, so let's consider the basis $\{|x\rangle, \ x \in \mathbf{R}\}$ of states where the particle has a definite position x and the different basis $\{|p\rangle, \ p \in \mathbf{R}\}$ of states where the particle has a definite momentum p. Note that both of these bases are continuous rather than discrete, so we need to generalize the orthogonality condition $\langle i|j\rangle = 0$ for $i \neq j$ to a delta-function. In particular, for the states of definite position

$$\langle x_1|x_2\rangle \ = \ \delta(x_1 - x_2) \ \longrightarrow \ 0 \text{ for } x_1 \neq x_2. \tag{24}$$

Consequently, in the completeness condition (23) for the position basis, the sum becomes the integral

$$|\psi\rangle \ = \ \int dx \ |x\rangle \times \langle x|\psi\rangle. \tag{25}$$

To prove this formula to any state $|\psi\rangle$, let's use the wave-function formalism. The state $|x_0\rangle$ has wave-function $\psi_{x_0}(x) = \delta(x - x_0)$ hence

$$\langle x_0|\psi\rangle \ = \ \int dx \ \delta(x - x_0) \times \psi(x) \ = \ \psi(x_0), \tag{26}$$

or in other words, *the value of the wave-function at point x_0 is nothing but the Dirac product of the quantum state in question and the state for definite x_0,*

$$\psi(x_0) \ = \ \langle x_0|\psi\rangle. \tag{27}$$

Consequently, the RHS of eq. (25) becomes

$$\int dx_0 \ |x_0\rangle \times \psi(x_0) \tag{28}$$

with the wave function

$$\psi'(x) \;=\; \int dx_0\, \delta(x-x_0) \times \psi(x_0) \;=\; \psi(x). \tag{29}$$

Since this equality works for any wave-function $\psi(x)$¡ this proves the completeness (25) of the definite-position basis.

Now consider the basis of states of definite momenta $|p\rangle$. In my convention — which is different from the Sakurai's textbook — these states have wavefunctions

$$\psi_p(x) \;=\; e^{ipx/\hbar} \quad\Longrightarrow\quad \langle p_1|p_2\rangle \;=\; 2\pi\hbar\, \delta(p_1-p_2). \tag{30}$$

To compensate for the $2\pi\hbar$ factor in front of the delta-function, the integral in the completeness formula has measure $dp/2\pi\hbar$, thus

$$\text{any } |\psi\rangle \;=\; \int \frac{dp}{2\pi\hbar} \; |p\rangle \times \langle p|\psi\rangle\,. \tag{31}$$

Or in the wave-function formalism,

$$\text{any } \psi(x) \;=\; \int \frac{dp}{2\pi\hbar} \, e^{ipx/\hbar} \times \widetilde{\psi}(p) \tag{32}$$

where

$$\widetilde{\psi}(x) \;\overset{\text{def}}{=}\; \langle p|\psi\rangle \;=\; \int dx\, e^{-ipx/\hbar} \times \psi(x). \tag{33}$$

By analogy with the $\psi(x) = \langle x|\psi\rangle$, the $\widetilde{\psi}(p) = \langle p|\psi\rangle$ is often called the *momentum-space wave-function*. Mathematically, eqs. (33) and (32) are the direct and the inverse Fourier transforms, and a theorem guarantees that the inverse transform gets us back to the same $\psi(x)$ we have started from; physically, this proves that the states $|p\rangle$ of definite momentum form a complete basis of the Hilbert space.

Note: different QM textbooks use different normalizations of the wavefunctions of definite momentum states, for example $(2\pi\hbar)^{-1/2} \times \exp(ipx/\hbar)$, and this leads to different powers of

the $(2\pi\hbar)$ factors in eqs. (32) and (33), for example

$$\psi(x) \;=\; \int \frac{dp}{\sqrt{2\pi\hbar}} \, e^{ipx/\hbar} \times \widetilde{\psi}(p) \quad \text{for} \quad \widetilde{\psi}(p) \;=\; \int \frac{dx}{\sqrt{2\pi\hbar}} \, e^{-ipx/\hbar} \times \psi(x). \tag{34}$$

In my convention — which I shall follow in this class — a state of definite momentum has wavefunction $\psi_p(x) = \exp(ipx/\hbar)$ — or in three dimensions $\psi_{\mathbf{p}} = \exp(i\mathbf{p}\cdot\mathbf{x}/\hbar)$ — without any pre-exponential factors, hence

$$\langle p_1|p_2\rangle \;=\; 2\pi\hbar\,\delta(p_1 = p_2) \quad \text{or} \quad \langle \mathbf{p}_1|\mathbf{p}_2\rangle \;=\; (2\pi\hbar)^3\delta^{(3)}(\mathbf{p}_1 - \mathbf{p}_2), \tag{35}$$

and the Fourier transforms between the coordinate-space and the momentum-space wavefunctions work according to eqs. (32) and (33) in 1d, or in 3D according to

$$\psi(\mathbf{x}) \;=\; \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3} \, e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \times \widetilde{\psi}(\mathbf{p}) \quad \text{for} \quad \widetilde{\psi}(\mathbf{p}) \;=\; \int d^3\mathbf{x}\, e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar} \times \psi(\mathbf{x}). \tag{36}$$

Note that both definite-position and definite-momentum states are un-normalizable, and in general any state where a *continuous* observable has a definite value is un-normalizable. Consequently, in an orthogonal basis $\{|\xi\rangle\}$ labeled by some continuous variable ξ one generally has

$$\langle \xi_1|\xi_2\rangle \;=\; f(\xi_1) \times \delta(\xi_1 - \xi_2) \tag{37}$$

for some normalization factor f which may be ξ-dependent. In this case, the integral in the complete-decomposition formula has measure $d\xi/f(\xi)$, thus

$$|\psi\rangle \;=\; \int \frac{d\xi}{f(\xi)} \, |\xi\rangle \times \langle\xi|\psi\rangle \tag{38}$$

On the other hand, if some observable A takes discrete values $A_1, A_2, \ldots$ rather than continuous values, then the states with definite values of this observable are normalizable, and it's conventional to give such states unit norms. This gives us an *orthonormal* bases of

states $|A_i\rangle$ where

$$\langle A_i|A_j\rangle \;=\; \delta_{ij} \;=\; \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases} \tag{39}$$

For such an orthonormal basis, the decomposition formula (23) becomes particularly simple,

$$\psi \;=\; \sum_{i=1}^{\infty} |A_i\rangle \times \langle A_i|\psi\rangle \tag{40}$$

without any denominators. Also, for any complete orthonormal basis, we may calculate the Dirac products of any two states $|\psi_i\rangle$ and $|\psi_2\rangle$ in terms of the coefficients of their decompositions into the basis states,

$$\langle \psi_1|\psi_2\rangle \;=\; \sum_{i=1}^{\infty} \langle \psi_1|A_i\rangle \times \langle A_i|\psi_2\rangle \;=\; \sum_{i=1}^{\infty} \langle A_i|\psi_1\rangle^* \times \langle A_i|\psi_2\rangle \,. \tag{41}$$

For a continuously labeled basis $\{|\xi\rangle, \xi \in \mathbf{R}\}$, the sum here becomes an integral, for example

$$\langle \psi_1|\psi_2\rangle \;=\; \int d^3\mathbf{x} \, \langle \psi_1|\mathbf{x}\rangle \times \langle \mathbf{x}|\psi_2\rangle \;=\; \int d^3\mathbf{x} \, \psi_1^*(\mathbf{x}) \times \psi_2(\mathbf{x}) \tag{42}$$

or

$$\langle \psi_1|\psi_2\rangle \;=\; \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3} \, \langle \psi_1|\mathbf{p}\rangle \times \langle \mathbf{p}|\psi_2\rangle \;=\; \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3} \, \widetilde{\psi}_1^*(\mathbf{p}) \times \widetilde{\psi}_2(\mathbf{p}). \tag{43}$$

Or more generally,

$$\langle \psi_1|\psi_2\rangle \;=\; \int \frac{d\xi}{f(\xi)} \, \langle \psi_1|\xi\rangle \times \langle \xi|\psi_2\rangle \,, \tag{44}$$

where $f(x)$ is the same normalization factor as in $\langle \xi_1 \xi_2\rangle = f(\xi_1)\delta(\xi_1 - \xi_2)$.

Probabilities, Observables, and Operators

Born rule

In optics, when a light wave with a polarization vector $\vec{\mathcal{E}}$ goes through a polarizer in the direction (or an elliptic generalization of a direction) $\mathbf{n}$, its intensity is reduced by a factor

$$P \;=\; \frac{I'}{I} \;=\; \frac{\left|\mathbf{n}^* \cdot \vec{\mathcal{E}}\right|^2}{|\mathbf{n}|^2 \,|\vec{\mathcal{E}}^2|} \,. \tag{45}$$

In terms of the photon beam, P is the *probability* of a photon to go through the polarizer.

In quantum mechanics there is a similar *Born rule of probabilities:* if a quantum system is in some state $|\Psi\rangle$ and we test whether the system happens to be in some other state $|\Phi\rangle$, then the probability of the positive answer is

$$P(|\Psi\rangle \text{ is } |\Phi\rangle) \;=\; \frac{|\langle\Phi|\Psi\rangle|^2}{\langle\Phi|\Phi\rangle\,\langle\Psi|\Psi\rangle} \,. \tag{46}$$

Note the symmetry of this formula WRT $|\Phi\rangle \leftrightarrow |\Psi\rangle$; for this reason P is often called the *probability overlap* between the states $|\Phi\rangle$ and $|\Psi\rangle$, or simply the *overlap.*

To make sense as a probability, the overlap (46) of any two states must lie between zero and one,

$$\forall\,|\Psi\rangle, |\Psi\rangle \in \mathcal{H}: \quad 0 \;\le\; P(|\Psi\rangle \text{ is } |\Phi\rangle) \;\le\; 1. \tag{47}$$

The non-negativity $P \ge 0$ is obvious from the definition (46) (and the positivity of the norm^2), $\langle\Psi|\Psi\rangle > 0$ and $\langle\Phi|\Phi\rangle > 0$, while the $P \le 1$ limit stems from the Cauchy–Schwarz inequality

$$|\langle\Phi|\Psi\rangle|^2 \;\le\; \langle\Psi|\Psi\rangle\,\langle\Phi|\Phi\rangle \tag{48}$$

which I shall prove in a moment. Moreover, this inequality is saturated — and hence $P = 1$ — if and only if $|\Phi\rangle = \lambda\,|\Psi\rangle$ for some complex number λ. And that's why we say that the Hilbert space vectors $|\Psi\rangle$ and $\lambda\,|\Psi\rangle$ describe exactly the same physical state of the quantum system.

Proof: Given any two non-zero Hilbert space vectors $|\Phi\rangle$ and $|\Psi\rangle$, let

$$|\chi\rangle \;=\; |\Phi\rangle \;-\; |\Psi\rangle \times \frac{\langle\Psi|\Phi\rangle}{\langle\Psi|\Psi\rangle}, \tag{49}$$

then

$$
\begin{aligned}
\langle\chi|\chi\rangle \;&=\; \langle\Phi|\Phi\rangle \;-\; \langle\Phi|\Psi\rangle \times \frac{\langle\Psi|\Phi\rangle}{\langle\Psi|\Psi\rangle} \;-\; \left(\frac{\langle\Psi|\Phi\rangle}{\langle\Psi|\Psi\rangle}\right)^{*} \times \langle\Psi|\Phi\rangle \;+\; \left|\frac{\langle\Psi|\Phi\rangle}{\langle\Psi|\Psi\rangle}\right|^{2} \times \langle\Psi|\Psi\rangle \\
&=\; \langle\Phi|\Phi\rangle \;-\; \frac{\langle\Phi|\Psi\rangle\,\langle\Psi|\Phi\rangle}{\langle\Psi\Psi\rangle} \;-\; \frac{\langle\Phi|\Psi\rangle\,\langle\Psi|\Phi\rangle}{\langle\Psi\Psi\rangle} \;+\; \frac{\langle\Phi|\Psi\rangle\,\langle\Psi|\Phi\rangle}{\langle\Psi\Psi\rangle} \\
&=\; \langle\Phi|\Phi\rangle \;-\; \frac{\langle\Phi|\Psi\rangle\,\langle\Psi|\Phi\rangle = |\langle\Phi|\Psi\rangle|^{2}}{\langle\Psi\Psi\rangle}.
\end{aligned}
\tag{50}
$$

But like any Hilbert space vector, $|\chi\rangle$ should have a non-negative norm, $\langle\chi|\chi\rangle \geq 0$, hence the bottom line of eq. (49) mist be non-negative,

$$\langle\Phi|\Phi\rangle \;-\; \frac{|\langle\Phi|\Psi\rangle|^{2}}{\langle\Psi\Psi\rangle} \;\geq\; 0, \tag{51}$$

which immediately leads to the Cauchy–Schwarz inequality (48) and hence $P \leq 1$.

Moreover, if $P = 1$ then the bottom line of eq. (50) vanishes, and hence the LHS of that equation also vanishes, $\langle\chi|\chi\rangle = 0$. But the only vector in the Hilbert space with a zero norm is the zero vector, thus $|\chi\rangle = 0$ and therefore

$$|\Phi\rangle \;-\; |\Psi\rangle \times \frac{\langle\Psi|\Phi\rangle}{\langle\Psi|\Psi\rangle} \;=\; |\chi\rangle \;=\; 0. \tag{52}$$

In other words,

$$|\Phi\rangle \;=\; |\Psi\rangle \times \left(\text{complex number } \frac{\langle\Psi|\Phi\rangle}{\langle\Psi|\Psi\rangle}\right), \tag{53}$$

Quod erat demonstrandum.

Observables

An *observable* is some property of a quantum system one might measure in an experiment, for example position, momentum, or energy of some particle. To completely describe some observable A, we need to specify two things: (1) The *spectrum* of its allowed values $\{A_i, \; i = 1, 2, \ldots\}$; this spectrum may be discrete, or continuous, or even have both discrete and continuous parts, but to keep my notations simple I'll assume a discrete spectrum. (2) The complete list of states $|i\rangle = |A = A_i\rangle$ which have definite values of A. Moreover, this list of states almost automatically provides a complete orthogonal basis of the Hilbert space. Indeed, that's why the *interesting* bases of Hilbert spaces are usually comprised of states with definite values of some observable(s).

Let's start with the orthogonality. Obviously, no quantum state can have two different *definite* values of the same observable. So when a quantum state $|i\rangle$ with a definite value A_i of A is tested for being the state $|j\rangle$ with a *different* definite value $A_j \neq A_i$, this test has zero probability of succeeding, thus by the Born rule $\langle j|i\rangle = 0$: *The states with different definite values of the same observable are always orthogonal to each other.*

However, if the spectrum of A is degenerate and several distinct states $|i\rangle, |j\rangle, \ldots$ have the same definite value $A_i = A_j = \cdots$, then they are not *automatically* orthogonal to each other. But in this case, all linear combinations of these degenerate states also have the same definite value of A; such linear combinations span a subspace of the Hilbert space, and we may always choose an orthogonal basis $|i'\rangle, |j\rangle', \ldots$ of this subspace. The bottom line is, despite any degeneracies, we end up with a complete list of mutually orthogonal states with definite values of A.

Now that we have an orthogonal basis of states $|i\rangle = |A = A_i\rangle$, let's verify its completeness as a basis. Take any quantum state $|\Psi\rangle$ and consider its net probability of having *any* allowed value of A. Obviously this net probability is 1, thus

$$\sum_{A_i}^{\text{all allowed}} P(|\Psi\rangle \text{ has } A = A_i) \;=\; 1, \tag{54}$$

and hence

$$\sum_i P(|\Psi\rangle \text{ is } |i\rangle) \;=\; 1. \tag{55}$$

By the Born rule (46), eq. (55) means

$$\sum_i \frac{|\langle i|\Psi\rangle|^2}{\langle\Psi|\Psi\rangle\,\langle i|i\rangle} \;=\; 1 \tag{56}$$

and therefore

$$\langle\Psi|\Psi\rangle \;=\; \sum_i \frac{|\langle i|\Psi\rangle|^2}{\langle i|i\rangle}\,. \tag{57}$$

Now let

$$|\chi\rangle \;=\; |\Psi\rangle \;-\; \sum_i |i\rangle \times \frac{\langle i|\Psi\rangle}{\langle i|i\rangle} \tag{58}$$

and calculate

$$
\begin{aligned}
\langle\chi|\chi\rangle \;=\;& \langle\Psi|\Psi\rangle \;-\; \sum_i \langle\Psi|i\rangle \times \frac{\langle i|\Psi\rangle}{\langle i|i\rangle} \;-\; \sum_i \frac{\langle\Psi|i\rangle}{\langle i|i\rangle} \times \langle i|\Psi\rangle \\
&+\; \sum_i \sum_j \frac{\langle\Psi|i\rangle}{\langle i|i\rangle} \times \langle i|j\rangle \times \frac{\langle j|\Psi\rangle}{\langle j|j\rangle} \\
&\langle\!\langle \text{ using } \langle i|j\rangle = 0 \text{ for } i \neq j \,\rangle\!\rangle \\
\;=\;& \langle\Psi|\Psi\rangle \;-\; 2 \times \sum_i \frac{\langle\Psi|i\rangle\,\langle i|\Psi\rangle}{\langle i|i\rangle} \;+\; \sum_{i=j} \frac{\langle\Psi|i\rangle\,\langle j|\Psi\rangle}{\langle i|i\rangle = \langle j|j\rangle} \\
\;=\;& \langle\Psi|\Psi\rangle \;-\; \sum_i \frac{|\langle i|\Psi\rangle|^2}{\langle i|i\rangle} \\
\;=\;& 0 \qquad \langle\!\langle \text{ by eq. (57) }\rangle\!\rangle.
\end{aligned}
\tag{59}
$$

By positivity of the Hilbert space norm, this means $|\chi\rangle = 0$ and therefore

$$|\Psi\rangle \;=\; \sum_i |i\rangle \times \frac{\langle i|\Psi\rangle}{\langle i|i\rangle}\,. \tag{60}$$

Since we have started with a completely general Hilbert space vector $|\Psi\rangle$, this means that any such vector decomposes into a linear combination of the basic vectors $|i\rangle = |A = A_i\rangle$, and that makes the basis $\{|A = A_i\rangle\ i = 1, 2, \ldots\}$ complete.

13

Expectation values and operators

Suppose we have a complete description of some observable A — its spectrum, and the complete basis of states with definite A_i. What can we say about the value of A in some general quantum state $|\Psi\rangle$? Generically, $|\Psi\rangle$ does not have a *definite* value of A but only probabilities of having different values; specifically,

$$P(|\Psi\rangle \text{ has value } A) \;=\; \sum_{i:A_i=A} P(|\Psi\rangle \text{ is } |i\rangle) \;=\; \sum_{i:A_i=A} \frac{|\langle i|\Psi\rangle|^2}{\langle i|i\rangle\,\langle\Psi|\Psi\rangle} \,. \tag{61}$$

Given these probabilities, we may do statistical analysis. In particular, we may calculate the statistical mean value (AKA the *expectation value*)

$$
\begin{aligned}
\langle A\rangle_\Psi \;&=\; \sum_A A \times P(|\Psi\rangle \text{ has value } A) \;=\; \sum_i A_i \times P(|\Psi\rangle \text{ is } |i\rangle) \\
&=\; \sum_i A_i \times \frac{|\langle i|\Psi\rangle|^2}{\langle i|i\rangle\,\langle\Psi|\Psi\rangle} \,.
\end{aligned}
\tag{62}
$$

We may also calculate the expectation values of functions of A,

$$\langle F(A)\rangle_\Psi \;=\; \sum_i F(A_i) \times \frac{|\langle i|\Psi\rangle|^2}{\langle i|i\rangle\,\langle\Psi|\Psi\rangle} \,, \tag{63}$$

especially

$$\langle A^2\rangle_\Psi \;=\; \sum_i A_i^2 \times \frac{|\langle i|\Psi\rangle|^2}{\langle i|i\rangle\,\langle\Psi|\Psi\rangle} \tag{64}$$

and hence the RMS spread (AKA the standard deviation)

$$(\Delta A)_\Psi \;=\; \sqrt{\langle A^2\rangle_\Psi - (\langle A\rangle_\Psi)^2} \,. \tag{65}$$

A more compact way of calculating the expectation values like (62) or (63) involve linear operators in the Hilbert space. For any observable A with a known spectrum and a known

basis of states $|i\rangle$ where A have definite values A_i, we may construct a linear operator

$$\hat{A} = \sum_i |i\rangle \times \frac{A_i}{\langle i|i\rangle} \times \langle i| \tag{66}$$

which maps a general Hilbert space vector $|\Psi\rangle$ to another vector

$$\hat{A}|\Psi\rangle = \sum_i |i\rangle \times \frac{A_i}{\langle i|i\rangle} \times \langle i|\Psi\rangle . \tag{67}$$

In is easy to see that the operator $\hat{A}$ is indeed linear:

$$
\begin{aligned}
\hat{A}(\alpha |\phi\rangle + \beta |\xi\rangle) &= \sum_i |i\rangle \times \frac{A_i}{\langle i|i\rangle} \times \langle i| (\alpha |\phi\rangle + \beta |\xi\rangle) \\
&= \sum_i |i\rangle \times \frac{A_i}{\langle i|i\rangle} \times \big(\alpha \langle i|\phi\rangle + \beta \langle i|\xi\rangle\big) \\
&= \alpha \sum_i |i\rangle \times \frac{A_i}{\langle i|i\rangle} \times \langle i|\phi\rangle \; + \; \beta \sum_i |i\rangle \times \frac{A_i}{\langle i|i\rangle} \times \langle i|\xi\rangle \\
&= \alpha \hat{A}|\phi\rangle \; + \; \beta \hat{A}|\xi\rangle .
\end{aligned}
\tag{68}
$$

In terms of the operator $\hat{A}$, the expectation value of the observable A in any quantum state state Ψ obtains as simply

$$\langle A \rangle_\Psi = \frac{\langle \Psi| \hat{A} |\Psi\rangle}{\langle \Psi|\Psi\rangle} . \tag{69}$$

Indeed,

$$
\begin{aligned}
\frac{\langle \Psi| \hat{A} |\Psi\rangle}{\langle \Psi|\Psi\rangle} &= \frac{1}{\langle \Psi|\Psi\rangle} \sum_i \langle \Psi|i\rangle \frac{A_i}{\langle i|i\rangle} \langle i|\Psi\rangle = \sum_i A_i \times \frac{\langle \Psi|i\rangle \langle i|\Psi\rangle}{\langle \Psi|\Psi\rangle \langle i|i\rangle} \\
&= \sum_i A_i \times P(|\Psi\rangle \text{ is } |i\rangle) = \langle A \rangle_\Psi .
\end{aligned}
\tag{70}
$$

Eqs. (66) and (67) construct a linear operator for a variable with a known spectrum and known basis of states having those values. Conversely, given a linear operator we may reconstruct the spectrum and the state having definite values of the corresponding observable

by solving the eigenvalue/eigenstate problem

$$\hat{A} \left| A = A_i \right\rangle \;=\; A_i \times \left| A = A_i \right\rangle. \tag{71}$$

Indeed, a state $|i\rangle$ having definite $A = A_i$ automatically obeys this equation,

$$\hat{A} \left| i \right\rangle \;=\; \sum_j |j\rangle \times A_j \left(\frac{\langle j|i\rangle}{\langle j|j\rangle} = \delta_{ij} \right) \;=\; |i\rangle \times A_i. \tag{72}$$

Conversely, any eigenstate $|\Psi\rangle$ obeying $\hat{A} \left| \Psi \right\rangle = a \left| \Psi \right\rangle$ has

$$\langle \Psi | \hat{A} | \Psi \rangle \;=\; a \times \langle \Psi | \Psi \rangle, \quad \langle \Psi | \hat{A}^2 | \Psi \rangle \;=\; a \times \langle \Psi | \hat{A} | \Psi \rangle \;=\; a^2 \times \langle \Psi | \Psi \rangle, \tag{73}$$

hence

$$\langle A \rangle_\Psi \;=\; a, \quad \langle A^2 \rangle \;=\; a^2 \quad \Longrightarrow \quad \Delta a \;=\; 0; \tag{74}$$

in other words, the state Ψ has *definite* value $A = a$.

For example, in the wave-function picture of the Hilbert space, the momentum operator $\hat{p}$ works as a derivative

$$\hat{p}\psi(x) \;=\; -i\hbar \frac{d\psi}{dx}, \tag{75}$$

so the momentum eigenvalue equation becomes

$$-i\hbar \frac{d}{dx} \psi_p(x) \;=\; p \times \psi_p(x). \tag{76}$$

This equation has a solution for any real p — hence a continuous momentum spectrum from $-\infty$ to $+\infty$ — with the corresponding eigenfunctions giving the wave-functions of states with definite momenta p as $\psi_p(x) = \exp(ipx/\hbar)$.

Matrices of operators

The operators in Hilbert space are closely associated with infinite matrices (or finite matrices if the Hilbert space has a finite dimension). Take any linear operator $\hat{A}$ and any *unrelated* orthogonal basis $\{|i\rangle, \; i = 1, 2, \ldots\}$. Let's make a matrix of Dirac products $\langle i|\, \hat{A}\, |j\rangle$ for all possible i and j,

$$\|A\| \;=\; \begin{pmatrix} \langle 1|\, \hat{A}\, |1\rangle & \langle 1|\, \hat{A}\, |2\rangle & \cdots \\ \langle 2|\, \hat{A}\, |1\rangle & \langle 2|\, \hat{A}\, |2\rangle & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{77}$$

Then we may reconstruct the operator $\hat{A}$ from this matrix as

$$\hat{A} \;=\; \sum_{i,j} |i\rangle \, \frac{\langle i|\, \hat{A}\, |j\rangle}{\langle i|i\rangle\, \langle j|j\rangle} \, \langle j| \,, \tag{78}$$

so for any quantum state Ψ,

$$\hat{A}\, |\Psi\rangle \;=\; \sum_{i,j} |i\rangle \, \frac{\langle i|\, \hat{A}\, |j\rangle}{\langle i|i\rangle\, \langle j|j\rangle} \, \langle j|\Psi\rangle \,. \tag{79}$$

In an orthonormal (orthogonal and normalized) basis where $\langle i|j\rangle = \delta_{ij}$ these formula become simply

$$\hat{A} \;=\; \sum_{i,j} |i\rangle \, \langle i|\, \hat{A}\, |j\rangle \, \langle j| \,, \qquad \hat{A}\, |\Psi\rangle \;=\; \sum_{i,j} |i\rangle \, \langle i|\, \hat{A}\, |j\rangle \, \langle j|\Psi\rangle \,. \tag{80}$$

Moreover, in an orthonormal basis, the matrix of an operator product $\hat{A}\hat{B}$ is the matrix product of corresponding matrices, $\|AB\| = \|A\| \times \|B\|$. Indeed,

$$\|AB\|_{i,j} \;=\; \langle i|\, \hat{A}\hat{B}\, |j\rangle \;=\; (\langle i|\, \hat{A})(\hat{B}\, |j\rangle) \;=\; \sum_{k} \langle i|\, \hat{A}\, |k\rangle \times \langle k|\, \hat{B}\, |j\rangle$$
$$= \; \sum_{k} \|A\|_{i,k} \times \|B\|_{k,j} \;=\; \big(\|A\| \times \|B\|\big)_{i,j} \,. \tag{81}$$

In the *eigenbasis* of the operator $\hat{A}$ — *i.e.*, the basis of its eigenstates, in which the corresponding observable A has definite values = eigenvalues($\hat{A}$), — the matrix (77) of $\hat{A}$

becomes diagonal,

$$\langle i | \hat{A} | j \rangle = A_i \times \langle i | i \rangle \, \delta_{ij}, \tag{82}$$

or in an orthonormal basis

$$\langle i | \hat{A} | j \rangle = A_i \times \delta_{ij}. \tag{83}$$

That's why finding the eigenvalues and the eigenstates of an operator is often called *diagonalizing* the operator.

In the basis of states that do not have definite values of A, the matrix (77) becomes non-diagonal. But in any orthonormal basis the matrix $\|A\|$ has the same spectrum of eigenvalues as the operator $\hat{A}$. This is particularly convenient for the finite matrices (in a finite-dimensional Hilbert space) where the eigenvalues obtain as roots of the polynomial equation $\det(\|A\| - a) = 0$, and then the eigenvectors obtain by solving a bunch of linear equations.

For an example, consider the Stern–Gerlach experiment. Let $\hat{A} = \hat{m}_x$, the x component of the atom's magnetic moment, and let's use the basis $\{|Z+\rangle, |Z-\rangle\}$ where m_z rather than m_x has definite values $m_z = \pm m_B$. As you shall see in your homework (set#1, problem 1), the matrix of $\hat{m}_x$ in this basis is

$$\|m_x\| = m_B \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{84}$$

This matrix has eigenvalues $\pm 1 \times m_B$ with the corresponding eigenvectors

$$\frac{1}{\sqrt{2}} \begin{pmatrix} +1 \\ +1 \end{pmatrix} \quad \text{and} \quad \frac{1}{\sqrt{2}} \begin{pmatrix} +1 \\ -1 \end{pmatrix}. \tag{85}$$

Consequently, the operator $\hat{m}_x$ also has eigenvalues $\pm 1 \times m_B$ with the corresponding eigenstates — *i.e.*, states of definite m_x — being

$$|m_x = +m_B\rangle = \frac{|Z+\rangle + |Z-\rangle}{\sqrt{2}} \quad \text{and} \quad |m_x = -m_B\rangle = \frac{|Z+\rangle - |Z-\rangle}{\sqrt{2}}. \tag{86}$$

For continuous bases like $\{|x\rangle\}$, the matrix of an operator becomes the *integration kernel*

(or simply the *kernel*)

$$A(x, y) = \langle x| \hat{A} |y\rangle, \tag{87}$$

hence

$$\hat{A} : \psi(x) \rightarrow \langle x| \hat{A} |\Psi\rangle = \int dy \langle x| \hat{A} |y\rangle \times \langle y|\Psi\rangle = \int dy K(x, y) \times \psi(x). \tag{88}$$

For an operator diagonal in the position basis, — for example, the potential operator which acts in the wave-function language as

$$\hat{V} : \psi(x) \rightarrow V(x)\psi(x), \tag{89}$$

— the kernel becomes proportional to the delta-function,

$$\langle x| |\hat{V} |y\rangle = V(y) \times \langle x|y\rangle = V(y) \times \delta(x - y). \tag{90}$$

This is the kernel's analogue of diagonality.

Note: the derivatives of the delta functions vanish for $x \neq y$ but they do not count as diagonal. For example, the momentum operator $\hat{p}$ is non-diagonal in the position basis since the states $|x\rangle$ of definite x do not have definite momenta. However, the kernel of $\hat{p}$ in the position basis is a derivative of the delta-function,

$$\langle x| |\hat{p} |y\rangle = -i\hbar \frac{\partial}{\partial x} \delta(x - y) \tag{91}$$

so that

$$\langle x| \hat{p} |\Psi\rangle = \int dy (-i\hbar) \frac{\partial}{\partial x} \delta(x - y) \times \langle y|\Psi\rangle = -i\hbar \frac{\partial}{\partial x} \langle y|\Psi\rangle, \tag{92}$$

or in the wave-function language

$$\hat{p} : \psi(x) \rightarrow -i\hbar \frac{\partial}{\partial x} \psi(x). \tag{93}$$

Hermitian operators

The Hermitian conjugate of a linear operator $\hat{A}$ in some Hilbert space is another operator $\hat{A}^\dagger$ in the same space such that for any 2 quantum states $\langle\Phi|$ and $|\Psi\rangle$

$$\langle\Phi|\,\hat{A}^\dagger\,|\Psi\rangle \;=\; \langle\Psi|\,\hat{A}\,|\Phi\rangle^*. \tag{94}$$

This type of Hermitian conjugation is related to the Hermitian conjugation of complex matrices which combines transposition of the matrix with complex conjugation of its elements

$$\left(M^\dagger\right)_{i,j} \;=\; \left(M_{j,i}\right)^*. \tag{95}$$

Indeed, in any orthogonal basis the matrix of $\hat{A}^\dagger$ is the Hermitian conjugate of the matrix of $\hat{A}$,

$$\forall i,j: \quad \langle i|\,\hat{A}^\dagger\,|j\rangle \;=\; \langle j|\,\hat{A}\,|i\rangle^*. \tag{96}$$

Conversely, if the matrices of operators $\hat{A}$ and $\hat{A}^\dagger$ are Hermitian conjugates of each other, then the operators themselves are Hermitian conjugates. Indeed, given eq. (96), for any 2 quantum states $\langle\Phi|$ and $|\Psi\rangle$ we have

$$\begin{aligned}
\langle\Phi|\,\hat{A}^\dagger\,|\Psi\rangle &= \sum_{i,j}\langle\Phi|i\rangle\times\frac{\langle i|\,\hat{A}^\dagger\,|j\rangle}{\langle i|i\rangle\,\langle j|j\rangle}\times\langle j|\Psi\rangle \;=\; \sum_{i,j}\langle i|\Psi\rangle^*\times\frac{\langle j|\,\hat{A}\,|i\rangle^*}{\langle i|i\rangle\,\langle j|j\rangle}\times\langle\Psi|j\rangle^* \\
&= \left(\sum_{j,i}\langle\Psi|j\rangle\times\frac{\langle j|\,\hat{A}\,|i\rangle}{\langle j|j\rangle\,\langle i|i\rangle}\times\langle i|\Phi\rangle\right)^* \;=\; \langle\Psi|\,\hat{A}\,|\Phi\rangle^*
\end{aligned} \tag{97}$$

Of particular interest are the *Hermitian operators* which are equal to their own Hermitian conjugates, $\hat{A}^\dagger = \hat{A}$. In other words, the operator $\hat{A}$ is Hermitian if and only if

$$\forall\,|\Phi\rangle,|\Psi\rangle\in\mathcal{H}: \quad \langle\Phi|\,\hat{A}\,|\Psi\rangle \;=\; \langle\Psi|\,\hat{A}\,|\Phi\rangle^*. \tag{98}$$

In terms of matrices, a matrix of a Hermitian operator in any $\perp$ basis is Hermitian, and conversely if the matrix of an operator in any $\perp$ basis is Hermitian then the operator itself is Hermitian.

A useful criterion of Hermiticity is: $\hat{A}$ is Hermitian if and only if

$$\forall \left|\Psi\right\rangle \in \mathcal{H}: \quad \left\langle\Psi\right|\hat{A}\left|\Psi\right\rangle \text{ is real.} \tag{99}$$

Indeed, if $\hat{A}$ is Hermitian then

$$\left\langle\Psi\right|\hat{A}\left|\Psi\right\rangle^{*} \;=\; \left\langle\Psi\right|\hat{A}^{\dagger}\left|\Psi\right\rangle \;=\; \left\langle\Psi\right|\hat{A}\left|\Psi\right\rangle \quad \Rightarrow \text{ real.} \tag{100}$$

On the other hand, if eq. (99) holds for any $\left|\Psi\right\rangle$, then for any 2 states $\left|\Psi\right\rangle$ and $\left|\Phi\right\rangle$, pick any complex number λ and let $\chi = \left|\Phi\right\rangle + \lambda\left|\Psi\right\rangle$. By assumption (99),

$$\left\langle\chi\right|\hat{A}\left|\chi\right\rangle \;=\; \left\langle\Phi\right|\hat{A}\left|\Phi\right\rangle \;+\; |\lambda|^{2}\left\langle\Psi\right|\hat{A}\left|\Psi\right\rangle \;+\; \lambda\left\langle\Phi\right|\hat{A}\left|\Psi\right\rangle \;+\; \lambda^{*}\left\langle\Psi\right|\hat{A}\left|\Phi\right\rangle \tag{101}$$

must be real, and by the same assumption, the first two terms on the right hand side also must be real. Hence, the remaining 2 terms on the RHS

$$\lambda\left\langle\Phi\right|\hat{A}\left|\Psi\right\rangle \;+\; \lambda^{*}\left\langle\Psi\right|\hat{A}\left|\Phi\right\rangle \tag{102}$$

must add up to a real number. Furthermore, they must add up to a real number for any complex number λ, in particular for $\lambda = 1$ and $\lambda = i$, thus

$$\begin{aligned} \text{real } \left\langle\Phi\right|\hat{A}\left|\Psi\right\rangle \;+\; \left\langle\Psi\right|\hat{A}\left|\Phi\right\rangle &\implies \operatorname{Im}\left\langle\Phi\right|\hat{A}\left|\Psi\right\rangle \;=\; -\operatorname{Im}\left\langle\Psi\right|\hat{A}\left|\Phi\right\rangle \\ \text{and real } i\left\langle\Phi\right|\hat{A}\left|\Psi\right\rangle \;-\; i\left\langle\Psi\right|\hat{A}\left|\Phi\right\rangle &\implies \operatorname{Re}\left\langle\Phi\right|\hat{A}\left|\Psi\right\rangle \;=\; +\operatorname{Re}\left\langle\Psi\right|\hat{A}\left|\Phi\right\rangle, \end{aligned} \tag{103}$$

hence altogether

$$\left\langle\Phi\right|\hat{A}\left|\Psi\right\rangle \;=\; \left\langle\Psi\right|\hat{A}\left|\Phi\right\rangle^{*}. \tag{104}$$

Finally, eq. (104) must hold true for any 2 states $\left|\Phi\right\rangle$ and $\left|\Psi\right\rangle$, hence by the definition (98) the operator $\hat{A}$ is Hermitian.

Next, for any Hermitian operator, all its eigenvalues are real, and the corresponding bra-type and ket-type eigenvectors are conjugates to each other. Indeed, eq. (94) implies

that for any operator $\hat{A}$ and any ket state $|\Psi\rangle$,

$$\left(\hat{A}\,|\Psi\rangle\right)^{\dagger} \;=\; \langle\Psi|\,\hat{A}^{\dagger}, \tag{105}$$

hence if $|\Psi\rangle$ is an eigen-ket with eigenvalue a, $\hat{A}\,|\Psi\rangle = a\,|\Psi\rangle$, then $\langle\Psi|$ is the eigen-bra of the conjugate operator $\hat{A}^{\dagger}$ with the eigenvalue a^*,

$$\langle\Psi|\,\hat{A}^{\dagger} \;=\; \left(\hat{A}\,|\Psi\rangle\right)^{\dagger} \;=\; \left(a\,|\Psi\rangle\right)^{\dagger} \;=\; a^*\,\langle\Psi|\,. \tag{106}$$

Consequently, for an Hermitian operator $\hat{A} = \hat{A}^{\dagger}$

$$\langle\Psi|\,\hat{A}\,|\Psi\rangle \;=\; \langle\Psi|\,(\hat{A}\,|\Psi\rangle = a\,|\Psi\rangle) \;=\; a \times \langle\Psi|\Psi\rangle \tag{107}$$

and at the same time

$$\langle\Psi|\,\hat{A}\,|\Psi\rangle \;=\; \left(\langle\Psi|\,\hat{A}^{\dagger}\right)|\Psi\rangle \;=\; \left(a^*\,\langle\Psi|\right)|\Psi\rangle \;=\; a^* \times \langle\Psi|\Psi\rangle\,, \tag{108}$$

thus

$$a \times \langle\Psi|\Psi\rangle \;=\; a^* \times \langle\Psi|\Psi\rangle \quad \Longrightarrow \quad a = a^* \qquad \langle\!\langle\,\text{since } \langle\Psi|\Psi\rangle \neq 0\,\rangle\!\rangle. \tag{109}$$

Quod erat demonstrandum.

Physically, this means that the Hermitian operators correspond to the real-number-valued observables.

Next, the eigenstates of an Hermitian operator for different eigenvalues are always $\perp$ to each other,

$$\text{for} \quad \hat{A} = \hat{A}^{\dagger}, \quad \hat{A}\,|\Psi_1\rangle = a_1\Psi_1, \quad \hat{A}\,|\Psi_2\rangle = a_2\Psi_2, \quad \text{and } a_1 \neq a_2: \quad \text{always } \langle\Psi_1|\Psi_2\rangle \;=\; 0. \tag{110}$$

Indeed,

$$\langle\Psi_1|\,\hat{A}\,|\Psi_2\rangle \;=\; \langle\Psi_1|\,\left(\hat{A}\,|\Psi_2\rangle = a_2\,|\Psi_2\rangle\right) \;=\; a_2 \times \langle\Psi_1|\Psi_2\rangle \tag{111}$$

and at the same time

$$\langle \Psi_1 | \hat{A} | \Psi_2 \rangle \;=\; \big(\langle \Psi_1 | \hat{A} = a_1 \langle \Psi_1 | \big) | \Psi_2 \rangle \;=\; a_1 \times \langle \Psi_1 | \Psi_2 \rangle, \tag{112}$$

thus

$$a_2 \times \langle \Psi_1 | \Psi_2 \rangle \;=\; a_1 \times \langle \Psi_1 | \Psi_2 \rangle \tag{113}$$

and hence either $a_2 = a_1$ or $\langle \Psi_1 | \Psi_2 \rangle = 0$ or both.

Theorem: *For any Hermitian operator $\hat{H}$ there is a complete $\perp$ basis of the whole Hilbert space made made from eigenstates of $\hat{H}$.* The general proof of this theorem is way too complicated for this class, especially if one wants to be mathematically rigorous about the continuous spectra. Although if you are really interested, I recommend the *Functional Analysis* textbook by Michael Reed and Barry Simon (several volumes in the *Methods of Modern Mathematical Physics* series). Instead of proving the theorem, let me simply state a **Corollary:** *If an Hermitian operator $\hat{H}$ has a non-degenerate spectrum of eigenvalues, then the complete set of eigenstates of $\hat{H}$ is automatically a complete $\perp$ basis of the whole Hilbert space.* For example, consider the momentum operator for a particle moving in one dimension:

$$\hat{p}\psi(x) \;=\; -i\hbar \frac{d\psi}{dx}. \tag{114}$$

To check the Hermiticity of this operators, take any two wave-functions $\psi(x)$ and $\phi(x)$ and calculate the matrix elements

$$
\begin{aligned}
\langle \phi | \hat{p} | \psi \rangle \;&=\; -i\hbar \int dx \, \phi^*(x) \times \frac{d\psi(x)}{dx} \\
&\quad \langle\!\langle \text{ integrating by parts, assuming } \psi(x), \phi(x) \to 0 \text{ when } x \to \pm\infty \,\rangle\!\rangle \\
&=\; +i\hbar \int dx \, \frac{d\phi^*(x)}{dx} \times \psi(x) \\
&=\; \left(-i\hbar \int dx \, \psi^*(x) \times \frac{d\phi(x)}{dx} \right)^* \\
&=\; \langle \psi | \hat{p} | \psi \rangle^*.
\end{aligned}
\tag{115}
$$

Thus, the momentum operator is indeed Hermitian. Also, in 1d its spectrum is non-degenerate and runs continuously from $p \to -\infty$ to $p \to +\infty$. Consequently, the eigenstates of the momentum — the plane waves $\psi_p(x) = \exp(ipx/\hbar)$ — form a complete $\perp$ basis for the 1d particle's Hilbert space.

Functions of operators

There are two rather different ways to define a function f of an operator $\hat{A}$, but fortunately they end up yielding the same $f(\hat{A})$. The first approach is based on the fact that for any quantum observable A, a function $f(A)$ of that observable is an observable in its own rights. Specifically, the $f(A)$ has definite values for exactly the same states as A, and the values themselves are $f(\text{values of } A)$. In the operator language, this means that each eigenstate of $\hat{A}$ is also an eigenstate of $f(\hat{A})$ with the corresponding eigenvalues related by the function f,

$$\text{if} \quad \hat{A}\,|i\rangle \;=\; A_i\,|i\rangle \quad \text{then } f(\hat{A})\,|i\rangle \;=\; f(A_i)\,|i\rangle\,. \tag{116}$$

Therefore, when $\hat{A}$ has a complete orthogonal basis of eigenstates — which is automatic for any Hermitian $\hat{A}$ — we have

$$\hat{A} \;=\; \sum_i |i\rangle\,\frac{A_i}{\langle i|i\rangle}\,\langle i|\,, \tag{117}$$

and then for any function f

$$f(\hat{A}) \;=\; \sum_i |i\rangle\,\frac{f(A_i)}{\langle i|i\rangle}\,\langle i|\,. \tag{118}$$

For example, the potential energy $V(x)$ of a particle is a function of its position x, so the potential energy operator is

$$\hat{V} \;=\; V(\hat{x}) \;=\; \int dx\,|x\rangle\,V(x)\,\langle x|\,. \tag{119}$$

Hence, for any quantum state Ψ

$$\langle x|\,\hat{V}\,|\Psi\rangle \;=\; \int dx'\,\big(\langle x|x'\rangle = \delta(x - x')\big) \times V(x') \times \langle x'|\Psi\rangle \;=\; V(x) \times \langle x|\Psi\rangle\,, \tag{120}$$

Or in the language of position-space wave-functions $\psi(x)$,

$$\hat{V}\psi(x) \;=\; V(x) \times \psi(x). \tag{121}$$

The main disadvantage of this method of defining $f(\hat{A})$ is that it requires diagonalizing $\hat{A}$, *i.e.* finding a complete basis of $\hat{A}$'s eigenstates. Worse, we end up with eq. (118) for the

24

$f(\hat{A})$ in terms of that basis, and translating the action of $f(\hat{A})$ in terms of some other basis would take some extra work. For example, consider the kinetic energy of a particle as a function $K(p)$ of its momentum p, thus the kinetic energy operator

$$\hat{K} \;=\; K(\hat{p}) \;=\; \int \frac{dp}{2\pi\hbar} \, |p\rangle \, K(p) \, \langle p| \,. \tag{122}$$

In terms of the momentum-space wave-functions $\tilde{\psi}(p)$ this means

$$\hat{K}\tilde{\psi}(p) \;=\; K(p) \times \tilde{\psi}(p), \tag{123}$$

but translating this formula to the language of the coordinate-space wave-functions $\psi(x)$ takes a Fourier transform. For the non-relativistic kinetic energy $K = p^2/2m$ this Fourier transform is rather easy,

$$
\begin{aligned}
K(p) \times \tilde{\psi}(p) \;&=\; \frac{p^2}{2m} \int dx \, e^{-ipx/\hbar} \times \psi(x) \\
&=\; \frac{1}{2m} \int dx \left(p^2 \times e^{-ipx/\hbar} \;=\; -\hbar^2 \frac{d^2}{dx^2} e^{-ipx/\hbar} \right) \times \psi(x) \\
&=\; -\frac{\hbar^2}{2m} \int dx \left(\frac{d^2}{dx^2} e^{-ipx/\hbar} \right) \times \psi(x) \\
&\qquad \langle\!\langle \text{ integrating by parts, twice} \rangle\!\rangle \\
&=\; -\frac{\hbar^2}{2m} \int dx \, e^{-ipx/\hbar} \times \frac{d^2}{dx^2} \psi(x)
\end{aligned}
\tag{124}
$$

hence

$$\hat{K}\psi(x) \;=\; -\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} \,. \tag{125}$$

However, for a relativistic kinetic energy $K(p) = \sqrt{c^2p^2 + m^2c^4}$, the Fourier transform of the operator $\hat{K}$ to the coordinate-space basis is an ugly integro-differential mess

$$\hat{H}\psi(x) \;=\; 2\hbar c \int dy \, K_0(\mu|x-y|) \times \left(\mu^2\psi(y) - \frac{\partial^2\psi(y)}{\partial y^2} \right) \tag{126}$$

where $\mu = mc/\hbar$ and K_0 is the modified Bessel function (second kind, order 0).

The other approach to functions of operators is algebraic and it does not need the operator's eigenbasis; instead, it works in any basis for which we know the action of the original operator $\hat{A}$. We start with powers of A which we define via subsequent actions,

$$\hat{A}^2 \,|\Psi\rangle \;=\; \hat{A}(\hat{A}\,|\Psi\rangle), \quad \hat{A}^3\,|\Psi\rangle \;=\; \hat{A}(\hat{A}(\hat{A}\,|\Psi\rangle)), \quad etc., \; etc. \tag{127}$$

Next, for any polynomial function

$$f(a) \;=\; c_0 \;+\; c_1 \times a \;+\; c_2 \times a^2 \;+\; \cdots \;+\; c_n \times a^n, \tag{128}$$

we define

$$f(\hat{A}) \;=\; c_0 \;+\; c_1 \times \hat{A} \;+\; c_2 \times \hat{A}^2 \;+\; \cdots \;+\; c_n \times \hat{A}^n, \tag{129}$$

where the integer powers of $\hat{A}$ act as in eq. (127). For example, for the non-relativistic kinetic energy $K = p^2/2m$, once we know that in the coordinate basis the momentum operator acts as $\hat{p}\psi(x) \;=\; -i\hbar d\psi/dx$, we immediately get

$$\hat{K}\psi(x) \;=\; \frac{1}{2m}\left(-i\hbar\frac{d}{dx}\right)^2 \psi(x) \;=\; -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) \tag{130}$$

without going through any Fourier transforms. Finally, for non-polynomial but analytic functions $f(a)$, we expand f into a Taylor series

$$f(a) \;=\; \sum_{n=0}^{\infty} c_n \times a^n \tag{131}$$

and then generalize eq. (129) to an infinite power series

$$f(\hat{A}) \;=\; \sum_{n=0}^{\infty} c_n \times \hat{A}^n. \tag{132}$$

Or in terms of the action of this operator on the quantum states,

$$f(\hat{A})\,|\Psi\rangle \;=\; \sum_{n=0}^{\infty} c_n \times \hat{A}^n\,|\Psi\rangle. \tag{133}$$

But the devil is in the details: Actually summing up the infinite series (132) — or even (133) for a particular $|\Psi\rangle$ — can be rather difficult.

Whenever both approaches to functions of operators can be made to work, they always yield the same operator $f(\hat{A})$. To see that, let's diagonalize the operator $\hat{A}$, *i.e.* find an $\perp$ basis of its eigenstates. In this basis

$$\hat{A} = \sum_i |i\rangle \frac{A_i}{\langle i|i\rangle} \langle i|, \tag{134}$$

hence

$$\hat{A}^2 = \hat{A} \times \hat{A} = \sum_{i,j} |i\rangle \frac{A_i}{\langle i|i\rangle} \times \left(\langle i|j\rangle = \langle i|i\rangle \times \delta_{ij}\right) \times \frac{A_j}{\langle j|j\rangle} \langle j|$$
$$= \sum_{i=j} |i\rangle \frac{A_i \times A_i}{\langle i|i\rangle} \langle i|, \tag{135}$$

and likewise, for any integer power of $\hat{A}$ defined as operator products of n operators $\hat{A}$ we get

$$\hat{A}^n = \sum_i |i\rangle \frac{A_i^n}{\langle i|i\rangle} \langle i|. \tag{136}$$

Consequently, for any polynomial function $f(A)$ or even for a Taylor series (131), we have

$$f(\hat{A}) \text{ defined as } \sum_{n=0}^{\infty} c_n \times \hat{A}^n = \sum_{n=0}^{\infty} c_n \times \sum_i |i\rangle \frac{A_i^n}{\langle i|i\rangle} \langle i|$$
$$= \sum_i |i\rangle \frac{1}{\langle i|i\rangle} \left(\sum_{n=0}^{\infty} c_n \times A_i^n\right) \langle i| \tag{137}$$
$$= \sum_i |i\rangle \frac{1}{\langle i|i\rangle} f(A_i) \langle i|,$$

In other words, the algebraically defined operator $f(\hat{A})$ (*cf.* eq. (132)) is equal to the $f(\hat{A})$ defined via the eigenbasis method (*cf.* eq. (118)), *quod erat demonstrandum.*

As an example, consider the exponential function of the momentum operator of a particle living in one dimension,

$$\hat{T}_a = \exp\left(-\frac{ia}{\hbar}\hat{p}\right) \tag{138}$$

for some constant a. Algebraically,

$$\hat{T}^a = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-ia}{\hbar} \right)^n \hat{p}^n, \tag{139}$$

where in the coordinate basis $\hat{p}$ acts as $-i\hbar d/dx$. Consequently,

$$\begin{aligned}
\hat{T}_a \psi(x) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-ia}{\hbar} \right)^n (-i\hbar)^n \frac{d^n}{dx^n} \psi(x) \\
&= \sum_{n=0}^{\infty} \frac{(-a)^n}{n!} \frac{d^n}{dx^n} \psi(x) \\
&= \text{Taylor series expansion of } \psi(x-a) \text{ in powers of a} \\
&= \psi(x-a).
\end{aligned} \tag{140}$$

Physically, this operator $\hat{T}_a$ acts as a *passive translation* of the coordinate system by $-a$, $x \to x' = x - a$. In Hilbert space notations, eq. (140) becomes

$$\langle x | \hat{T}_a | \Psi \rangle = \langle x - a | \Psi \rangle. \tag{141}$$

In particular, for $|\Psi\rangle$ being the state $|y\rangle$ where the particle is located at point y, we have

$$\langle x | \hat{T}_a | y \rangle = \langle x - a | y \rangle = \delta((x-a) - y) = \delta(x - (y+a)) = \langle x | y + a \rangle, \tag{142}$$

and since this is true for all $\langle x|$, it follows that

$$\hat{T}_a | y \rangle = | y + a \rangle. \tag{143}$$

Thus the operator $\hat{T}_a$ *actively translates* the particle's position by $+a$ from y to $y + a$.

Now let's recalculate the operator $\hat{T}_a$ using the eigenbasis method. In the momentum

basis

$$\hat{p} \;=\; \int \frac{dp}{2\pi\hbar}\,|p\rangle\,p\,\langle p| \quad\Longrightarrow\quad \hat{T}_a \;=\; \int \frac{dp}{2\pi\hbar}\,|p\rangle \exp(-iap/\hbar)\,\langle p|\,. \tag{144}$$

In the coordinate basis, the matrix elements of this operator obtain as

$$
\begin{aligned}
\langle x|\,\hat{T}_a\,|y\rangle \;&=\; \int \frac{dp}{2\pi\hbar}\,\langle x|p\rangle \exp(-iap/\hbar)\,\langle p|y\rangle \\
&=\; \int \frac{dp}{2\pi\hbar}\,\exp(ipx/\hbar)\times \exp(-iap/\hbar)\times \exp(-ipy/\hbar) \\
&=\; \int \frac{dp}{2\pi\hbar}\,\exp\bigl(ip(x-a-y)/\hbar\bigr) \\
&=\; \delta(x-a-p) \\
&=\; \langle x|y+a\rangle\,,
\end{aligned}
\tag{145}
$$

exactly as we had for the algebraic method.

Compatible Observables

Two observables $\mathcal{A}$ and $\mathcal{B}$ are called *compatible* iff (if and only if) there is a basis of states where both $\mathcal{A}$ and $\mathcal{B}$ have definite values. Likewise, 3 or more observables $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$ are compatible iff there exist a basis of states $|i\rangle$ where they all have definite values $\mathcal{A}_i, \mathcal{B}_i, \mathcal{C}_i, \dots$. Example: for a particle moving in 3 dimensions, the 3 coordinates (x, y, z) of its position are compatible. Counterexample: the 3 components (m_x, m_y, m_z) of an atom's magnetic moment $\mathbf{m}$ are not compatible.

Theorem: *two real observables $\mathcal{A}$ and $\mathcal{B}$ are compatible if and only if the corresponding Hermitian operators $\hat{A}$ and $\hat{B}$ commute, i.e. $\hat{A}\hat{B} = \hat{B}\hat{A}$.* **Proof:** suppose the two observables are compatible, so there exist a basis $\{|i\rangle\}$ where both $\mathcal{A}$ and $\mathcal{B}$ have definite values A_i and B_i. In this basis, both operators $\hat{A}$ and $\hat{B}$ must be diagonal,

$$\hat{A} \;=\; \sum_i |i\rangle\,\frac{A_i}{\langle i|i\rangle}\,\langle i|\,, \quad \hat{B} \;=\; \sum_i |i\rangle\,\frac{B_i}{\langle i|i\rangle}\,\langle i|\,, \tag{146}$$

and since the diagonal matrices always commute with each other, the operators $\hat{A}$ and $\hat{B}$ should also commute with each other. Indeed the eigenbasis of an Hermitian operator must

be orthogonal, $\langle i|j \rangle = 0$ for $i \neq j$, hence

$$\hat{A}\hat{B} \;=\; \sum_{i,j} |i\rangle \frac{A_i}{\langle i|i\rangle} \big(\langle i|j\rangle = \delta_{ij}\langle j|j\rangle \big) \frac{B_j}{\langle j|j\rangle} \langle j| \;=\; \sum_{i=j} |i\rangle \frac{A_i B_i}{\langle i|i\rangle} \langle i| \tag{147}$$

and likewise

$$\hat{B}\hat{A} \;=\; \sum_{i=j} |i\rangle \frac{B_i A_i}{\langle i|i\rangle} |i\rangle \,, \tag{148}$$

and since the values A_i and B_i commute, $A_i B_i = B_i A_i$, the RHS of the last two equations are equal to each other so the LHS must also be equal, $\hat{A}\hat{B} = \hat{B}\hat{A}$.

For the other direction of the theorem, suppose the operator $\hat{A}$ is Hermitian and the operator $\hat{B}$ (Hermitian or not) commutes with $\hat{A}$, $\hat{A}\hat{B} = \hat{B}\hat{A}$. **Lemma:** *If $|i\rangle$ and $|j\rangle$ be eigenstates of $\hat{A}$ with different eigenvalues,*

$$\hat{A}|i\rangle \;=\; A_i |i\rangle \,, \quad \hat{A}|j\rangle \;=\; A_j |j\rangle \,, \quad A_i \neq A_j, \tag{149}$$

then the matrix element of $\hat{B}$ between these two states must vanish,

$$\langle i| \hat{B} |j\rangle \;=\; 0. \tag{150}$$

Proof:

$$\langle i| \hat{B}\hat{A} |j\rangle \;=\; \langle i| \hat{B}(\hat{A}|j\rangle = A_j |J\rangle) \;=\; A_j \times \langle i| \hat{B} |j\rangle$$
$$\big\| \quad \langle\!\langle \text{ since } \hat{A} \text{ and } \hat{B} \text{ commute} \rangle\!\rangle \tag{151}$$
$$\langle i| \hat{A}\hat{B} |j\rangle \;=\; (\hat{A}|i\rangle)^\dagger \hat{B} |j\rangle \;=\; (A_i |i\rangle)^\dagger \hat{B} |j\rangle \;=\; A_i \times \langle i| \hat{B} |j\rangle$$

hence for $A_i \neq A_j$ we must have $\langle i| \hat{B} |j\rangle = 0$, *quod erat demonstrandum.* **Corollary:** *if the Hermitian operator $\hat{A}$ has a non-degenerate spectrum of eigenvalues, then in the eigenbasis of $\hat{A}$ any other operators $\hat{B}$ which commutes with $\hat{A}$ is automatically diagonalized!*

In terms of the observables, this corollary means that *if an observable $\mathcal{A}$ has a non-degenerate spectrum, then any observable $\mathcal{B}$ that is compatible with $\mathcal{A}$ must be a function of $\mathcal{A}$, $\mathcal{B} = F(\mathcal{A})$.* Indeed, for any allowed value A_i of $\mathcal{A}$ there is a unique (up to an overall coefficient) eigenstate $|i\rangle$ of $\hat{A}$ where $\mathcal{A}$ has that value. Since $\hat{B}$ is diagonal in the same basis

$\{|i\rangle\}$, this state $|i\rangle$ is also an eigenvalue of $\hat{B}$, $\hat{B}|i\rangle = B_i|i\rangle$, which means that in that state the observable $\mathcal{B}$ has definite value B_i. Thus, any allowed value of $\mathcal{A}$ uniquely determines the value of $\mathcal{B}$,

$$A_i \ \rightarrow \ i \ \rightarrow \ B_i, \tag{152}$$

which makes $\mathcal{B}$ a function of $\mathcal{A}$.

Now, suppose $\hat{A}$ has a degenerate spectrum. Since $\hat{A}$ is Hermitian, it must have an $\perp$ basis of eigenstates, but now there may be several independent eigenstates for the same eigenvalue. Let's label these eigenstates $|a, \alpha\rangle$ where a is the eigenvalue of $\hat{A}$ and α is an additional label needed to resolve the degeneracy, thus

$$\hat{A}|a, \alpha\rangle \ = \ a \times |a, \alpha\rangle \tag{153}$$

Then for any operator $\hat{B}$ which commutes with $\hat{A}$

$$\langle a, \alpha|\hat{B}|a', \alpha'\rangle \ = \ 0 \quad \text{unless} \quad a = a', \tag{154}$$

but

$$\langle a, \alpha|\hat{B}|\text{same } a, \ \text{different } \alpha'\rangle \ \neq \ 0 \quad \text{is OK.} \tag{155}$$

In other words, in the eigenbasis of $\hat{A}$, the matrix of $\hat{B}$ is *block-diagonal*, for example

$$
A = \begin{pmatrix}
a_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & a_2 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & a_2 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & a_3 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & a_4 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & a_4 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & a_4 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & a_5 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\quad
B = \begin{pmatrix}
\star & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & \star & \star & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & \star & \star & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & \star & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & \star & \star & \star & 0 & \cdots \\
0 & 0 & 0 & 0 & \star & \star & \star & 0 & \cdots \\
0 & 0 & 0 & 0 & \star & \star & \star & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \star & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\tag{156}
$$

where the stars denote the allowed non-zero matrix elements of $\hat{B}$. In terms of (a, α) labels of $\hat{A}$'s eigenstates, each a corresponds to a diagonal block of $\hat{B}$'s matrix, while the matrix elements within the block are labeled by (α, α').

For each block a, the states $\{|a,\alpha\rangle\,,\ \text{fixed } a,\ \text{all } \alpha\}$ and their linear combinations form a subspace $\mathcal{H}_a$ of the Hilbert space $\mathcal{H}$. The operator $\hat{B}$ maps states in this subspace into other states in that subspace, Indeed,

$$\forall \sum_\alpha |a,\alpha\rangle \times \psi_\alpha \in \mathcal{H}_a\,,$$

$$
\begin{aligned}
\hat{B} \sum_\alpha |a,\alpha\rangle \times \psi_\alpha \ &= \ \sum_{a',\alpha} {\sum_{\alpha}}' |a',\alpha'\rangle \times \langle a',\alpha'|\,\hat{B}\,|a,\alpha\rangle \times \psi_\alpha \\
&\quad \langle\!\langle \text{ using } \langle a',\alpha'|\,\hat{B}\,|a,\alpha\rangle = 0 \text{ unless } a' = a \,\rangle\!\rangle \\
&= \ \sum_{\alpha'} \sum_\alpha |a,\alpha'\rangle \times \langle a,\alpha'|\,\hat{B}\,|a,\alpha\rangle \times \psi_\alpha \\
&= \ \sum_{\alpha'} |a,\alpha'\rangle \times (\text{a complex number}) \\
&\in \ \mathcal{H}_a\,,
\end{aligned}
\tag{157}
$$

so it may be treated as an operator $\hat{B}_a$ acting within that subspace. Also, if the operator $\hat{B}$ is Hermitian, then its restriction $\hat{B}_a$ on the block a is also Hermitian, so the general theorem about diagonalization of Hermitian operators may be applied to the operator $\hat{B}_a$ acting in the space $\mathcal{H}_a$. Consequently, there must exist a complete $\perp$ basis of the block $\mathcal{H}_a$ made of eigenstates $|a;b,\beta\rangle$ of $\hat{B}_a$,

$$\hat{B}_a \,|a;b,\beta\rangle \ = \ b\,|a;b,\beta\rangle\,. \tag{158}$$

Combining such bases of each block $\mathcal{H}_a$ then gives a complete $\perp$ basis

$$\{|a;b,\beta\rangle\,,\ \text{all } a,\ \text{all } b,\ \text{all } \beta\} \tag{159}$$

of the whole Hilbert space $\mathcal{H}$. Obviously, the operator $\hat{B}$ is diagonal in this basis,

$$\hat{B}\,|a;b,\beta\rangle \ = \ \hat{B}_a\,|a;b,\beta\rangle \ = \ b\,|a;b,\beta\rangle\,, \tag{160}$$

but the operator $\hat{A}$ is also diagonal in this basis since each $|a;b,\beta\rangle$ is a linear combination

of states $|a, \alpha\rangle$ from the same block:

$$
\begin{aligned}
|a; b, \beta\rangle &= \sum_\alpha |a, \alpha\rangle \times C(a, \alpha, b, \beta) \quad \text{for some coefficients } C(a, \alpha, b, \beta), \text{ hence} \\
\hat{A} |a; b, \beta\rangle &= \sum_\alpha \hat{A} |a, \alpha\rangle \times C(a, \alpha, b, \beta) \\
&= \sum_\alpha a \times |a, \alpha\rangle \times C(a, \alpha, b, \beta) \\
&= a \times \sum_\alpha |a, \alpha\rangle \times C(a, \alpha, b, \beta) = a \times |a; b, \beta\rangle .
\end{aligned}
\tag{161}
$$

Thus, both operators $\hat{A}$ and $\hat{B}$ have diagonal matrices in the same basis. Form the observable point of view, this means that both $\mathcal{A}$ and $\mathcal{B}$ have definite values in the same complete basis of the Hilbert space. By definition, this makes $\mathcal{A}$ and $\mathcal{B}$ compatible observables, which completes the Theorem's proof.

Basis building

The above theorem about compatible observables, commuting operators, and common eigenbases leads to following procedure for building complete $\perp$ basis of the Hilbert space:

1. Pick a convenient (and hopefully physically interesting) real observable $\mathcal{A}$ and diagonalize the corresponding Hermitian operator $\hat{A}$.

2. If the eigenvalue spectrum of $\hat{A}$ happens to be non-degenerate, then the eigenstates $|a\rangle$ are unique (up to rescaling by complex numbers), and together they make a complete $\perp$ basis $\{|a\rangle, \text{ all } a\}$. We are done.

3. Otherwise, — *i.e.* if the spectrum of $\hat{A}$ is degenerate, — pick another convenient observable $\mathcal{B}$ which is compatible with $\mathcal{A}$ but is independent of $\mathcal{A}$, *i.e.* $\mathcal{B}$ is not a function of $\mathcal{A}$. Then simultaneously diagonalize both $\hat{A}$ and $\hat{B}$.

4. If for any allowed combination (a, b) there is a unique simultaneous eigenstate $|a, b\rangle$, then together these eigenstates provide a good basis of the Hilbert space and our work is done.

5. Otherwise, pick a third observable $\mathcal{C}$ which is compatible with both $\mathcal{A}$ and $\mathcal{B}$ but is independent of them, $\mathcal{C} \neq F(\mathcal{A}, \mathcal{B})$. Then diagonalize all three operators $\hat{A}, \hat{B}, \hat{C}$ in the same common eigenbasis.

6. Again, if for all allowed eigenvalue combinations (a, b, c) there is a unique eigenstate $|a, b, c\rangle$, then the complete set of these eigenstates is the desired basis of the Hilbert space.

7. Otherwise, keep adding compatible and independent observables and diagonalizing all the corresponding commuting operators at the same time. Eventually, you would get enough observables to make all the eigenstates unique.

Example: for a 3D particle without internal structure or spin, the 3 coordinates $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ of its position are compatible and independent observables. Also, each simultaneous eigenstate $|x, y, z\rangle$ is unique, so the position basis $\{|x, y, z\rangle, \ (x, y, z) \in \mathbf{R}^3\}$ is a complete basis of the Hilbert space. Likewise, the 3 components $\hat{P}_x$, $\hat{P}_y$, $\hat{P}_x$ of the particle's momentum are compatible and independent observables, and each definite-momentum state $|p_x, p_y, p_z\rangle$ is unique, so the momentum basis $\{|p_x, p_y, p_z\rangle, \ (p_x, p_y, p_z) \in \mathbf{R}^3\}$ is also a good complete basis.

For some problems it's convenient to mix coordinates and momenta in different directions, for example use the basis of $|x, y, p_z\rangle$ states since $\hat{P}_z$ is compatible with $\hat{X}$ and $\hat{Y}$¿ However, one should never combine the coordinate and the momentum along the same axis since $\hat{X}$ and $\hat{P}_x$ — and likewise $\hat{Y}$ and $\hat{P}_y$, or $\hat{Z}$ and $\hat{P}_z$ — are incompatible with each other.

For a particle with non-zero spin — for example, an electron with spin $= \frac{1}{2}$, — the 3 coordinates of its position are compatible and independent, but they do not completely specify the quantum state. Instead, for each (x, y, z) there two distinct quantum states of an electron being at that position. To resolve this degeneracy, we need an additional observable for the electron spins' state, for example the z component $\hat{S}_z$ of the spin vector $\hat{\mathbf{S}}$, thus the complete basis comprises

$$|(x, y, z, s)\rangle \quad \text{for} \quad (x, y, z) \in \mathbf{R}^3 \quad \text{while} \quad s = \pm\frac{\hbar}{2}. \tag{162}$$

Alternatively, one may use

$$|(p_x, p_y, p_z, s)\rangle \quad \text{for} \quad (p_x, p_y, p_z) \in \mathbf{R}^3 \quad \text{while} \quad s = \pm\frac{\hbar}{2}. \tag{163}$$

or some other combination of coordinates, momenta, and spin observables.

Complex observables

Thus far we have focused on real-number-valued observables corresponding to Hermitian operators. The complex-number-valued observables are more complicated because their real and imaginary parts are not always compatible with each other. Indeed, let $\mathcal{A}$ be some kind of a complex observable so that the corresponding operator $\hat{A}$ is not Hermitian. Let

$$\hat{A}_1 \;=\; \frac{\hat{A} + \hat{A}^\dagger}{2} \;=\; \hat{A}_1^\dagger, \quad \hat{A}_2 \;=\; \frac{\hat{A} - \hat{A}^\dagger}{2i} \;=\; \hat{A}_2^\dagger, \quad \hat{A}_1 \;+\; i\hat{A}_2 \;=\; \hat{A}, \qquad (164)$$

in other words $\hat{A}_1$ is the Hermitian part of $\hat{A}$ corresponding to the real part of $\mathcal{A}$ while $i\hat{A}_2$ is the anti-Hermitian part of $\hat{A}$ corresponding to the imaginary part of the observable. If $\hat{A}_1$ and $\hat{A}_2$ commute with each other, $\hat{A}_1\hat{A}_2 = \hat{A}_2\hat{A}_1$, then there is common eigenbasis of the two operators where

$$\hat{A}_1 \left| a_1, a_2, \ldots \right\rangle \;=\; a_1 \left| a_1, a_2, \ldots \right\rangle \quad \text{and} \quad \hat{A}_2 \left| a_1, a_2, \ldots \right\rangle \;=\; a_2 \left| a_1, a_2, \ldots \right\rangle, \qquad (165)$$

and therefore

$$\begin{aligned}
\hat{A} \left| a_1, a_2, \ldots \right\rangle &\;=\; (\hat{A}_1 + i\hat{A}_2) \left| a_1, a_2, \ldots \right\rangle \;=\; (a_1 + ia_2) \left| a_1, a_2, \ldots \right\rangle, \\
\hat{A}^\dagger \left| a_1, a_2, \ldots \right\rangle &\;=\; (\hat{A}_1 - i\hat{A}_2) \left| a_1, a_2, \ldots \right\rangle \;=\; (a_1 - ia_2) \left| a_1, a_2, \ldots \right\rangle.
\end{aligned} \qquad (166)$$

Thus, a non-Hermitian operator $\hat{A}$ has an orthogonal eigenbasis if an only if its Hermitian and anti-Hermitian parts (164) commute with each other, $\hat{A}_1\hat{A}_2 = \hat{A}_2\hat{A}_1$.

To restate this criterion in terms of the $\hat{A}$ and $\hat{A}^\dagger$ operators rather than $\hat{A}_1$ and $\hat{A}_2$, let's define the *commutator bracket* of operators,

$$[\hat{A}, \hat{B}] \;\overset{\text{def}}{=}\; \hat{A}\hat{B} \;-\; \hat{B}\hat{A}. \qquad (167)$$

By definition, the commutator is bilinear

$$\begin{aligned}
[(\alpha\hat{A} + \beta\hat{B}), \hat{C}] &\;=\; \alpha[\hat{A}, \hat{C}] \;+\; \beta[\hat{B}, \hat{C}], \\
[\hat{A}, (\beta\hat{B} + \gamma\hat{C})] &\;=\; \beta[\hat{A}, \hat{B}] \;+\; \gamma[\hat{A}, \hat{C}],
\end{aligned} \qquad (168)$$

and antisymmetric

$$[\hat{A}, \hat{B}] \;=\; -[\hat{B}, \hat{A}] \quad \Longrightarrow \quad [\hat{A}, \hat{A}] \;=\; 0. \tag{169}$$

Less obviously, the commutator bracket obeys the Jacobi identity

$$[\hat{A}, [\hat{B}, \hat{C}]] \;+\; [\hat{B}, [\hat{C}, \hat{A}]] \;+\; [\hat{C}, [\hat{A}, \hat{B}]] \;=\; 0 \tag{170}$$

which you shall prove in a future homework.

For the moment, let's focus on the Hermitian and anti-Hermitian parts (164) of some operator $\hat{A}$. Using bi-linearity and antisymmetry of the commutator, we have

$$
\begin{aligned}
[\hat{A}_1, \hat{A}_2] \;&=\; \frac{1}{4i} \, [(\hat{A} + \hat{A}^\dagger), (\hat{A} - \hat{A}^\dagger)] \\
&=\; \frac{1}{4i} \left([\hat{A}, \hat{A}] + [\hat{A}^\dagger, \hat{A}] - [\hat{A}, \hat{A}^\dagger] - [\hat{A}^\dagger, \hat{A}^\dagger] \right) \\
&=\; \frac{1}{4i} \left(0 + [\hat{A}^\dagger, \hat{A}] + [\hat{A}^\dagger, \hat{A}] - 0 \right) \\
&=\; \frac{1}{2i} \, [\hat{A}^\dagger, \hat{A}],
\end{aligned}
\tag{171}
$$

hence $\hat{A}_1$ commutes with $\hat{A}_2$ if and only if $\hat{A}$ commutes with $\hat{A}^\dagger$. Therefore, *a non-hermitian operator $\hat{A}$ has a complete $\perp$ eigenbasis if an only if it commutes with its Hermitian conjugate,* $\hat{A}\hat{A}^\dagger = \hat{A}^\dagger \hat{A}$.

Generalizing this rule to multiple complex observables is straightforward: A complete $\perp$ basis of the Hilbert space where several complex observables $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, *etc.*, all have definite values exists if an only if all the corresponding operators $\hat{A}, \hat{B}, \hat{C}, \ldots$ and all their Hermitian conjugates $\hat{A}^\dagger, \hat{B}^\dagger, \hat{C}^\dagger, \ldots$ all commute with each other.

Functions of several operators

When two observables $\mathcal{A}$ and $\mathcal{B}$ are compatible, a function $\mathcal{F} = F(\mathcal{A}, \mathcal{B})$ of these observables is itself a well-defined observable: it has definite values $F(A_i, B_i)$ in the states $|i\rangle$ where both $\mathcal{A}$ and $\mathcal{B}$ have definite values A_i and B_i. Likewise, for several compatible observables $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$, any function of all these observable is itself a good observable.

In the operator languages, this translates into functions of multiple mutually commuting operators, for example the potential operator in 3 dimensions,

$$\hat{V} \; = \; V(\hat{x}, \hat{y}, \hat{z}). \tag{172}$$

Similar to the functions of a single operator, there is the eigenbasis method and the algebraic method for constructing functions of multiple (but commuting) operators, and whenever both methods are practical, they always produce the same result.

Let's start with the eigenbasis method. Suppose the operators $\hat{A}, \hat{B}, \hat{C}$ — as well as the conjugates $\hat{A}^\dagger, \ldots$ of all the non-Hermitian operators in this list — all commute with each other. In this case, there must exist a complete $\perp$ basis of the Hilbert space where all of these operators are diagonal,

$$\hat{A} \left| i \right\rangle \; = \; A_i \left| i \right\rangle, \quad \hat{B} \left| i \right\rangle \; = \; B_i \left| i \right\rangle, \quad \hat{C} \left| i \right\rangle \; = \; C_i \left| i \right\rangle, \quad etc., \; etc. \tag{173}$$

In this common eigenbasis, for any function $F(A, B, C, \ldots)$ we define the operator

$$F(\hat{A}, \hat{B}, \hat{C}, \ldots) \; \stackrel{\text{def}}{=} \; \sum_i \left| i \right\rangle \frac{F(A_i, B_i, C_i, \ldots)}{\langle i | i \rangle} \left\langle i \right|. \tag{174}$$

For example, using the position basis $\{ \left| x, y, z \right\rangle \}$ for a 3d particle, we define the potential operator as

$$\hat{V} \; = \; V(\hat{x}, \hat{y}, \hat{z}) \; = \; \int d^3(x, y, z) \left| x, y, z \right\rangle V(x, y, z) \left\langle x, y, z \right|, \tag{175}$$

so in the wave-function language it acts as

$$\hat{V} \psi(x, y, z) \; = \; V(x, y, z) \times \psi(x, y, z). \tag{176}$$

In the algebraic method, we start with product functions such as $f(x, y) = xy$ for which we let

$$f(\hat{x}, \hat{y}) \; = \; \hat{x} \hat{y}. \tag{177}$$

Note: an ordinary product of two real numbers x and y does not distinguish between xy and yx, so it does not tell us in which order to multiply the operators $\hat{x}$ and $\hat{y}$. So for general

operators $\hat{x}$ and $\hat{y}$ their product becomes ambiguous:

$$f(\hat{x}, \hat{y}) = \hat{x}\hat{y} \text{ or } \hat{y}\hat{x} ? \tag{178}$$

However, for the compatible observables the corresponding operators commute with each other, thus

$$f(\hat{x}, \hat{y}) = \hat{x}\hat{y} = \hat{y}\hat{x} \quad \text{is unambiguous.} \tag{179}$$

Likewise, for products of several operators, as long as all the operators commute with each other, the order of the product does not matter,

$$\hat{x}^2\hat{y} = \hat{x}\hat{y}\hat{x} = \hat{y}\hat{x}^2, \tag{180}$$

etc., etc. Thus, any product-of-powers function translates to product-of-powers of the corresponding operators, hence by linearity any polynomial function of two or more variables translates to a similar polynomial of commuting operators, for example

$$F(a, b, c) = \sum_{k,\ell,m} F_{k,\ell,m} \times a^k b^\ell c^m \quad \longrightarrow \quad F(\hat{A}, \hat{B}, \hat{C}) = \sum_{k,\ell,m} F_{k,\ell,m} \times \hat{A}^k \hat{B}^\ell \hat{C}^m. \tag{181}$$

Finally, for the non-polynomial but analytic functions we expand them into combined power series

$$F(a, b, c) = \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} F_{k,\ell,m} \times a^k b^\ell c^m, \tag{182}$$

and then construct the operator function $F(\hat{A}, \hat{B}, \hat{C})$ as a similar power series

$$F(\hat{A}, \hat{B}, \hat{C}) = \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \sum_{m=0}^{\infty} F_{k,\ell,m} \times \hat{A}^k \hat{B}^\ell \hat{C}^m. \tag{183}$$

As long as the operators $\hat{A}, \hat{B}, \hat{C}$ all commute with each other, the order of the operator product in each term in this series does not matter, so the series is well defined. But actually summing it up could be quite a challenge.