An Axiomatic Derivation of Schrödinger's Equation - INCOMPLETE

Chris Clark December 6, 2009

1 Introduction

This paper derives the Schrödinger equation in one-dimension from axioms based on probabilistic measurements and constraints imposed by classical expectations. The Lemmas found in the appendix are segregated based on the fact that they are purely mathematical and do not depend on any of the axioms presented below. Definitions can be found in a separate appendix. Where appropriate, the equal sign in an equation has been labelled with the number of the primary theorem (T), lemma (L), or definition (D) used. This paper also introduces a special notation to indicate the results of a measurement. The non-deterministic operator $\hat{M}^{\star}_{\hat{a}}$ represents the effect of the measurement of the quantity $\hat{a}$ on a state, which is to collapse the state to an eigenstate of $\hat{a}$. The non-deterministic function $M^{\star}_{\hat{a}}(|\psi\rangle)$ represents the value obtained by a measurement of the quantity $\hat{a}$. In both cases, the star denotes the fact that these expressions are non-deterministic so $\hat{M}^{\star}_{\hat{a}}|\psi\rangle \neq \hat{M}^{\star}_{\hat{a}}|\psi\rangle$ and $M^{\star}_{\hat{a}}(|\psi\rangle) \neq M^{\star}_{\hat{a}}(|\psi\rangle)$ in general.

2 Axioms

Axiom 1. *For all observables $\hat{a}$ and all states $|\psi\rangle$, measurement of the quantity $\hat{a}$ on state $|\psi\rangle$ results in a transformation to the state $\hat{M}^{\star}_{\hat{a}}|\psi\rangle = |a\rangle$ for some eigenstate $|a\rangle$ of $\hat{a}$.*

Axiom 2. *For all observables $\hat{a}$ and all states $|\psi\rangle$, the probability of obtaining a measured value of a upon measurement of the quantity $\hat{a}$ on state $|\psi\rangle$ is given by $P[M^{\star}_{\hat{a}}(|\psi\rangle) = a] = |\langle\psi|\, a\rangle|^2$, where $|a\rangle$ is an eigenstate of $\hat{a}$ with eigenvalue a.*

Axiom 3. *There exist observables $\hat{x}$ and $\hat{p}$, with no explicit time dependence, that satisfy $\frac{d}{dt}\langle\hat{x}\rangle_\psi = \langle\hat{p}\rangle_\psi/m$ and $\frac{d}{dt}\langle\hat{p}\rangle_\psi = -\langle V'(\hat{x})\rangle_\psi$ where m is a constant and V is an arbitrary function. We call these observables the position and momentum respectively.*

Axiom 4. *There exists an analytic operator-valued function of time duration $\hat{\tau}(\delta t) = \tau(\hat{x}, \hat{p}, \delta t)$ such that for all states $|\psi(t)\rangle$, $\hat{\tau}(\delta t)\,|\psi(t)\rangle = |\psi(t+\delta t)\rangle$. We call this operator-valued function the time-evolution operator (which is a slight abuse of nomenclature).*

Axiom 5. *All non-trivial operators that are conserved for all choices of V can be written as a function of a single conserved operator. We call this operator the Hamiltonian.*

3 Theorems

Theorem 1. *Let $\hat{a}$ be an observable and $|a\rangle$ be an eigenstate of $\hat{a}$ with eigenvalue a. Then $M^{\star}_{\hat{a}}(|a\rangle) = a$.*

Proof.

$$1 \overset{D17}{=} |\langle a|\, a\rangle|^2 \overset{A2}{=} P[M^{\star}_{\hat{a}}(|a\rangle) = a]$$

So the probability of the consequent being true is one, which completes the proof. □

Theorem 2. *Let $\hat{a}$ be an observable and let $|a_1\rangle$ and $|a_2\rangle$ be eigenstates of $\hat{a}$ with eigenvalues a_1 and a_2 respectively. Then $\langle a_1|\, a_2\rangle = \delta_{a_1 a_2}$.*

Proof.

$$|\langle a_1|\, a_2\rangle|^2 \overset{A2}{=} P[M^{\star}_{\hat{a}}(|a_1\rangle) = a_2]$$

$$\overset{T1}{=} P[a_1 = a_2] = \delta_{a_1 a_2}$$

□

Theorem 3. *Let $\hat{a}$ be an observable. Then $\sum_a |a\rangle\langle a| = \hat{I}$, where $|a\rangle$ is any eigenstate of $\hat{a}$ with eigenvalue a (such an eigenstate may be non-unique).*

Proof. Consider an arbitrary state $|\psi\rangle$ and an arbitrary vector in the span of eigenstates of $\hat{a}$, $|\psi_a\rangle = \sum_a c_a |a\rangle$. Let $|\psi_r\rangle = |\psi\rangle - |\psi_a\rangle$. Taking the inner product with $|a'\rangle$,

$$\langle a'|\, \psi_r\rangle = \langle a'|\, \psi\rangle - \sum_a c_a \langle a'|\, a\rangle$$

$$\overset{T2}{=} \langle a'|\, \psi\rangle - \sum_a c_a \delta_{a'a} = \langle a'|\, \psi\rangle - c_{a'}$$

Let the constants c_a be defined by $c_a = \langle a|\, \psi\rangle$ so that $\langle a|\, \psi_r\rangle = 0$ for all $|a\rangle$. Then

$$\langle\psi|\, \psi\rangle = \left(\sum_{a'} \langle a'|\, c^*_{a'} + \langle\psi_r|\right)\left(\sum_a c_a |a\rangle + |\psi_r\rangle\right)$$

$$= \sum_{a'} \sum_a c^*_{a'} c_a \langle a' | a \rangle + \langle \psi_r | \psi_r \rangle$$

$$\overset{T2}{=} \sum_{a'} \sum_a c^*_{a'} c_a \delta_{a'a} + \langle \psi_r | \psi_r \rangle$$

$$= \sum_a |c_a|^2 + \langle \psi_r | \psi_r \rangle$$

$$= \sum_a | \langle a | \psi \rangle |^2 + \langle \psi_r | \psi_r \rangle$$

$$\overset{A2}{=} \sum_a P[M^\star_{\hat{a}}(|\psi\rangle) = a] + \langle \psi_r | \psi_r \rangle$$

$$\overset{D20}{=} \sum_a P[|\hat{a}\hat{M}^\star_{\hat{a}} |\psi\rangle | = a] + \langle \psi_r | \psi_r \rangle$$

$$\overset{A1}{=} \sum_a P[|\hat{a} |a'\rangle | = a] + \langle \psi_r | \psi_r \rangle$$

$$= \sum_a P[a' = a] + \langle \psi_r | \psi_r \rangle$$

$$= 1 + \langle \psi_r | \psi_r \rangle$$

But by since $|\psi\rangle$ is a state, $\langle \psi | \psi \rangle = 1$, so $\langle \psi_r | \psi_r \rangle = 0$. This can only be true if $|\psi_r\rangle = 0$ by the definition of an inner product. Therefore,

$$|\psi\rangle = |\psi_a\rangle = \sum_a |a\rangle \langle a | \psi \rangle = \left(\sum_a |a\rangle \langle a| \right) |\psi\rangle$$

So by Definition 7, $\sum_a |a\rangle \langle a| = \hat{I}$. ◻

Theorem 4. *Let $\hat{\alpha}_1$ and $\hat{\alpha}_2$ be linear operators. If $\hat{\alpha}_1 |a\rangle = \hat{\alpha}_2 |a\rangle$ for all eigenstates $|a\rangle$ of some observable $\hat{a}$, then $\hat{\alpha}_1 = \hat{\alpha}_2$.*

Proof. For any state $|\psi\rangle$ we have

$$\hat{\alpha}_1 |\psi\rangle \overset{T3}{=} \hat{\alpha}_1 \sum_a |a\rangle \langle a | \psi \rangle$$

$$\overset{D9}{=} \sum_a \hat{\alpha}_1 |a\rangle \langle a | \psi \rangle = \sum_a \hat{\alpha}_2 |a\rangle \langle a | \psi \rangle$$

$$\overset{D9}{=} \hat{\alpha}_2 \sum_a |a\rangle \langle a | \psi \rangle \overset{T3}{=} \hat{\alpha}_2 |\psi\rangle$$

By Definition 7, $\hat{\alpha}_1 = \hat{\alpha}_2$. ◻

Theorem 5. *For all observables $\hat{a}$ and all states $|\psi\rangle$, $\langle \hat{a} \rangle_\psi = \langle \psi | \hat{a} |\psi\rangle$.*

Proof.

$$\langle \hat{a} \rangle_\psi \overset{D21}{=} \sum_a a P[M^\star_{\hat{a}}(|\psi\rangle) = a]$$

$$\overset{A2}{=} \sum_a a | \langle \psi | a \rangle |^2$$

$$\overset{L1}{=} \sum_a a \langle \psi | a \rangle \langle a | \psi \rangle$$

$$\overset{D14}{=} \sum_a \langle \psi | \hat{a} |a\rangle \langle a | \psi \rangle$$

$$\overset{D9}{=} \left\langle \psi \middle| \hat{a} \left(\sum_a |a\rangle \langle a| \right) \middle| \psi \right\rangle$$

$$\overset{T3}{=} \langle \psi | \hat{a} |\psi\rangle$$

◻

Theorem 6. *There exists a Hermitian operator $\hat{H} = H(\hat{x}, \hat{p})$, where H is analytic, and a real constant $\hbar$ such that for all states $|\psi\rangle$, $\hat{H} |\psi(t)\rangle = i\hbar \frac{d}{dt} |\psi(t)\rangle$.*

Proof. By Axiom 4, there exists an analytic operator-valued function $\hat{\tau}(dt)$ such that

$$\hat{\tau}(\delta t) |\psi(t)\rangle = |\psi(t + \delta t)\rangle$$

for all states $|\psi(t)\rangle$ where δt is any duration of time. If we let $\delta t = dt$, where dt is an infinitesimal duration of time and expand $\hat{\tau}(dt)$ as a power series in dt (which works because $\hat{\tau}(\delta t)$ is analytic), we obtain

$$\hat{\tau}(dt) = \hat{c}_0 + \hat{c}_1 dt + \hat{c}_2 dt^2 + \cdots = \hat{c}_0 + \hat{c}_1 dt$$

where terms of order dt^2 have been dropped because dt is infinitesimal. In the limit $dt \to 0$, we have

$$\hat{\tau}(0) |\psi(t)\rangle = |\psi(t)\rangle$$

So $\hat{\tau}(0)$ is the identity operator on states, which implies that $\hat{c}_0 = \hat{I}$. The norms of states are unity at all times, so

$$\langle \psi(t) | \psi(t) \rangle \overset{D17}{=} \langle \psi(t + dt) | \psi(t + dt) \rangle$$

$$= \left\langle \psi(t) \middle| (\hat{I} + \hat{c}_1^\dagger dt)(\hat{I} + \hat{c}_1 dt) \middle| \psi(t) \right\rangle$$

$$= \left\langle \psi(t) \middle| \hat{I} + (\hat{c}_1^\dagger + \hat{c}_1) dt \middle| \psi(t) \right\rangle$$

$$= \langle \psi(t) | \psi(t) \rangle + dt \left\langle \psi(t) \middle| \hat{c}_1^\dagger + \hat{c}_1 \middle| \psi(t) \right\rangle$$

Therefore

$$\left\langle \psi(t) \middle| \hat{c}_1^\dagger + \hat{c}_1 \middle| \psi(t) \right\rangle = 0$$

Since $\hat{c}_1^\dagger + \hat{c}_1$ is Hermitian, by Lemma 9 we have $\hat{c}_1^\dagger + \hat{c}_1 = 0$ or $\hat{c}_1^\dagger = -\hat{c}_1$. Define $\hat{H} = i\hbar\hat{c}_1$, where $\hbar$ is an unspecified real constant. Then $\hat{H}$ is a Hermitian operator because $\hat{H}^\dagger \overset{L4}{=} -i\hbar\hat{c}_1^\dagger = i\hbar\hat{c}_1 = \hat{H}$. So we have

$$\hat{\tau}(dt) = \hat{I} - i\hat{H}dt/\hbar$$

for some Hermitian operator $\hat{H}$. Now by Axiom 4, $\hat{\tau}(dt) = \tau(\hat{x}, \hat{p}, dt)$, and $\hat{H}$ is a constant with respect to time, so $\hat{H} = H(\hat{x}, \hat{p})$. Since τ is analytic, H must be also. Now we can write down an expression for the time derivative of states.

$$\begin{aligned}
\frac{d}{dt}\left|\psi(t)\right\rangle &= \lim_{dt\to 0}\frac{\left|\psi(t+dt)\right\rangle - \left|\psi(t)\right\rangle}{dt} \\
&\overset{A4}{=} \lim_{dt\to 0}\frac{\hat{\tau}(dt)\left|\psi(t)\right\rangle - \left|\psi(t)\right\rangle}{dt} \\
&= \lim_{dt\to 0}\frac{(\hat{I} - i\hat{H}dt/\hbar)\left|\psi(t)\right\rangle - \left|\psi(t)\right\rangle}{dt} \\
&= \lim_{dt\to 0}\frac{-\frac{i}{\hbar}\hat{H}dt\left|\psi(t)\right\rangle}{dt} \\
&= -\frac{i}{\hbar}\hat{H}\left|\psi(t)\right\rangle
\end{aligned}$$

Therefore $\hat{H}\left|\psi(t)\right\rangle = i\hbar\frac{d}{dt}\left|\psi(t)\right\rangle$ for some Hermitian operator $\hat{H} = H(\hat{x}, \hat{p})$ and some real constant $\hbar$. □

Theorem 7. *For all linear operators $\hat{\alpha}$ and for all states $\left|\psi(t)\right\rangle$, $\frac{d}{dt}\langle\hat{\alpha}\rangle_\psi = \frac{1}{i\hbar}\langle[\hat{\alpha}, \hat{H}]\rangle_\psi + \langle\frac{\partial\hat{\alpha}}{\partial t}\rangle_\psi$.*

Proof.

$$\begin{aligned}
\frac{d}{dt}\langle\hat{\alpha}\rangle_\psi &\overset{T5}{=} \frac{d}{dt}\left\langle\psi(t)\right|\hat{\alpha}\left|\psi(t)\right\rangle \\
&= \left(\frac{d}{dt}\left\langle\psi(t)\right|\right)\hat{\alpha}\left|\psi(t)\right\rangle + \left\langle\psi(t)\middle|\frac{\partial\hat{\alpha}}{\partial t}\middle|\psi(t)\right\rangle \\
&\quad + \left\langle\psi(t)\right|\hat{\alpha}\left(\frac{d}{dt}\left|\psi(t)\right\rangle\right)
\end{aligned}$$

Taking the Hermitian conjugate of the equation in Theorem 6 we obtain $\left\langle\psi(t)\right|\hat{H} = -i\hbar\frac{d}{dt}\left\langle\psi(t)\right|$, so

$$\begin{aligned}
&\overset{T6}{=} -\frac{1}{i\hbar}\left\langle\psi(t)\middle|\hat{H}\hat{\alpha}\middle|\psi(t)\right\rangle + \left\langle\psi(t)\middle|\frac{\partial\hat{\alpha}}{\partial t}\middle|\psi(t)\right\rangle \\
&\quad + \frac{1}{i\hbar}\left\langle\psi(t)\middle|\hat{\alpha}\hat{H}\middle|\psi(t)\right\rangle \\
&\overset{D8}{=} \frac{1}{i\hbar}\left\langle\psi(t)\middle|\hat{\alpha}\hat{H} - \hat{H}\hat{\alpha}\middle|\psi(t)\right\rangle + \left\langle\psi(t)\middle|\frac{\partial\hat{\alpha}}{\partial t}\middle|\psi(t)\right\rangle \\
&\overset{T5}{=} \frac{1}{i\hbar}\langle[\hat{\alpha}, \hat{H}]\rangle_\psi + \left\langle\frac{\partial\hat{\alpha}}{\partial t}\right\rangle_\psi
\end{aligned}$$

□

Theorem 8. *$[\hat{x}, \hat{H}] = i\hbar\hat{p}/m$ and $[\hat{p}, \hat{H}] = -i\hbar V'(\hat{x})$.*

Proof. By Axiom 3, there exist observables $\hat{x}$ and $\hat{p}$, with no explicit time-dependence, that satisfy

$$\frac{d}{dt}\langle\hat{x}\rangle_\psi = \langle\hat{p}\rangle_\psi/m$$

and

$$\frac{d}{dt}\langle\hat{p}\rangle_\psi = -\langle V'(\hat{x})\rangle_\psi$$

for all states $|\psi\rangle$. By Theorem 7, we can write

$$-\frac{1}{\hbar}\langle i[\hat{x}, \hat{H}]\rangle_\psi = \langle\hat{p}\rangle_\psi/m$$

and

$$-\frac{1}{\hbar}\langle i[\hat{p}, \hat{H}]\rangle_\psi = -\langle V'(\hat{x})\rangle_\psi$$

Observables and functions of observables are Hermitian, so the right hand side operators are Hermitian. By Lemma 10, the left hand side operators are Hermitian. So by Lemma 9,

$$[\hat{x}, \hat{H}] = i\hbar\hat{p}/m$$

and

$$[\hat{p}, \hat{H}] = -i\hbar V'(\hat{x})$$

□

Theorem 9. *For all states $|\psi\rangle$, $\frac{d}{dt}\langle i[\hat{x}, \hat{p}]\rangle_\psi = 0$.*

Proof.

$$\begin{aligned}
\frac{d}{dt}\langle i[\hat{x}, \hat{p}]\rangle_\psi &\overset{T7}{=} \frac{1}{\hbar}\langle[[\hat{x}, \hat{p}], \hat{H}]\rangle_\psi \\
&\overset{L12}{=} -\frac{1}{\hbar}\langle[[\hat{p}, \hat{H}], \hat{x}] + [[\hat{H}, \hat{x}], \hat{p}]\rangle_\psi \\
&\overset{T8}{=} -\frac{1}{\hbar}\langle[-i\hbar V'(\hat{x}), \hat{x}] + [-i\hbar\hat{p}/m, \hat{p}]\rangle_\psi \overset{D13}{=} 0
\end{aligned}$$

□

Theorem 10. *There exists a real constant $\hbar$ such that $[\hat{x}, \hat{p}] = i\hbar$.*

Proof. By Theorem 9,

$$\frac{d}{dt}\langle i[\hat{x}, \hat{p}]\rangle_\psi = 0$$

And by Lemma 10, $i[\hat{x}, \hat{p}]$ is Hermitian, which means $i[\hat{x}, \hat{p}]$ is a conserved quantity. But by Theorem 7, $\hat{H}$ is also a conserved quantity. So by Axiom 5, we have three possibilities: $i[\hat{x}, \hat{p}]$ can be written as a function of $\hat{H}$, $\hat{H}$ can be written as a function of $i[\hat{x}, \hat{p}]$, or one of these two operators is trivial. By Theorem 8, we see that $\hat{H}$ must depend on V, while $i[\hat{x}, \hat{p}]$ does not, so $\hat{H}$ cannot be written as a function of $i[\hat{x}, \hat{p}]$.

It remains to show that $i[\hat{x},\hat{p}]$ cannot be written as a function of $\hat{H}$ [MISSING STEPS]. Therefore one must be trivial. If $\hat{H}$ was trivial, then by Theorem 8, $\hat{p}$ would be zero. This contradicts Axiom 3 because the zero operator does not satisfy the definition of an observable, so $\hat{H}$ cannot be trivial. Therefore, $i[\hat{x},\hat{p}]$ is a trivial conserved operator, so we can write $i[\hat{x},\hat{p}] = -\hbar\hat{I}$ for some constant $\hbar$. Since $i[\hat{x},\hat{p}]$ is Hermitian, by Lemma 8, $\hbar$ must be real. Previously, the value of $\hbar$ was undetermined, but now we use this commutator to define its magnitude. Multiplying both sides by $-i$ we obtain $[\hat{x},\hat{p}] = i\hbar$ for some real constant $\hbar$. □

Theorem 11. *For all states* $|\psi\rangle$, $i\hbar\frac{d}{dt}|\psi(t)\rangle = \left(\frac{\hat{p}^2}{2m} + V(\hat{x})\right)|\psi(t)\rangle$.

Proof. By Theorem 6, $\hat{H} = H(\hat{x},\hat{p})$ for some analytic function H, and by Theorem 10, $[\hat{x},\hat{p}] = i\hbar$, so by Lemma 16 we can write,

$$\hat{H} = H(\hat{x},\hat{p}) = H_0(\hat{x}) + H_1(\hat{x})\hat{p} + H_2(\hat{x})\hat{p}^2 + \cdots$$

By Theorem 8,

$$\begin{aligned}\frac{i\hbar}{m}\hat{p} &\overset{T8}{=} [\hat{x},\hat{H}] = \left[\hat{x}, \sum_{k=0}^{\infty} H_k(\hat{x})\hat{p}^k\right] \\ &\overset{L11}{=} \sum_{k=0}^{\infty}[\hat{x}, H_k(\hat{x})\hat{p}^k] \\ &\overset{L13}{=} \sum_{k=0}^{\infty} H_k(\hat{x})[\hat{x},\hat{p}^k] \\ &\overset{L14}{=} \sum_{k=0}^{\infty} H_k(\hat{x})(i\hbar k\hat{p}^{k-1})\end{aligned}$$

Therefore, $H_2(\hat{x}) = \frac{1}{2m}$ while all other $H_k(\hat{x})$ are zero except perhaps $H_0(\hat{x})$ because it is multiplied by $k = 0$. So $\hat{H} = H_0(\hat{x}) + \hat{p}^2/2m$. Again using Theorem 8,

$$-i\hbar V'(\hat{x}) \overset{T8}{=} [\hat{p},\hat{H}] = [\hat{p}, H_0(\hat{x})]$$

By Lemma 15, $H_0(\hat{x}) = V(\hat{x})$, so

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x})$$

So by Theorem 6, $i\hbar\frac{d}{dt}|\psi(t)\rangle = \left(\frac{\hat{p}^2}{2m} + V(\hat{x})\right)|\psi(t)\rangle$ for all states $|\psi\rangle$. □

4 Appendix: Definitions

Definition 1. A vector $|\chi\rangle$ is an element of a Hilbert space.

Definition 2. The Hermitian conjugate of a vector, denoted $|\chi\rangle^\dagger$, is the complex conjugate of its transpose i.e. $|\chi\rangle^\dagger = |\chi\rangle^{T*}$. For aesthetics we write this as $\langle\chi|$.

Definition 3. The inner product of two vectors $|\chi_1\rangle$ and $|\chi_2\rangle$ is the matrix product $\langle\chi_1|\,|\chi_2\rangle$. For aesthetics we write this as $\langle\chi_1|\,\chi_2\rangle$. We omit the proof that this satisfies the mathematical axioms for an inner product.

Definition 4. Two vectors $|\chi_1\rangle$ and $|\chi_2\rangle$ are orthogonal if and only if $\langle\chi_1|\,\chi_2\rangle = 0$.

Definition 5. The norm of a vector, $|\,|\chi\rangle\,|$, is $\sqrt{\langle\chi|\,\chi\rangle}$. We omit the proof that this satisfies the mathematical axioms for a norm.

Definition 6. An operator $\hat{\alpha}$ is a mapping from vectors to vectors.

Definition 7. Two operators $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are equal if and only if $\hat{\alpha}_1|\chi\rangle = \hat{\alpha}_2|\chi\rangle$ for all vectors $|\chi\rangle$.

Definition 8. The sum/difference of two operators $\hat{\alpha}_1$ and $\hat{\alpha}_2$, written $\hat{\alpha}_1 + \hat{\alpha}_2$ or $\hat{\alpha}_1 - \hat{\alpha}_2$, is the operator defined by $(\hat{\alpha}_1 \pm \hat{\alpha}_2)|\chi\rangle = \hat{\alpha}_1|\chi\rangle \pm \hat{\alpha}_2|\chi\rangle$, for any vector $|\chi\rangle$.

Definition 9. A linear operator $\hat{\alpha}$ is an operator that satisfies $\hat{\alpha}(c_1|\chi_1\rangle + c_2|\chi_2\rangle) = c_1\hat{\alpha}|\chi_1\rangle + c_2\hat{\alpha}|\chi_2\rangle$ for all vectors $|\chi_1\rangle$ and $|\chi_2\rangle$ and all constants c_1 and c_2.

Definition 10. The Hermitian conjugate of an operator $\hat{\alpha}$, written $\hat{\alpha}^\dagger$, is the operator that satisfies $(\hat{\alpha}|\chi\rangle)^\dagger = \langle\chi|\,\hat{\alpha}^\dagger$.

Definition 11. An operator $\hat{\alpha}$ is Hermitian if and only if $\hat{\alpha}$ is a linear operator and $\hat{\alpha}^\dagger = \hat{\alpha}$.

Definition 12. An operator $\hat{\alpha}$ is unitary if and only if $\hat{\alpha}$ is a linear operator and $\hat{\alpha}^\dagger\hat{\alpha} = \hat{I}$.

Definition 13. The commutator of two operators $\hat{\alpha}_1$ and $\hat{\alpha}_2$, written $[\hat{\alpha}_1,\hat{\alpha}_2]$, is the operator $\hat{\alpha}_1\hat{\alpha}_2 - \hat{\alpha}_2\hat{\alpha}_1$.

Definition 14. An eigenvector of an operator $\hat{\alpha}$, written $|\alpha\rangle$, is a vector that satisfies $\hat{\alpha}|\alpha\rangle = \alpha|\alpha\rangle$, where α is a constant.

Definition 15. The eigenvalue of an eigenvector $|\alpha\rangle$ of an operator $\hat{\alpha}$ is the constant α that satisfies $\hat{\alpha}|\alpha\rangle = \alpha|\alpha\rangle$.

Definition 16. An observable $\hat{a}$ is a Hermitian operator such that all states are in the span of any set of vectors containing at least one eigenvector of $\hat{a}$ for every distinct eigenvalue of $\hat{a}$.

Definition 17. A state $|\psi(t)\rangle$ is a time-dependent vector with unit norm $\langle\psi(t)|\,\psi(t)\rangle = 1$ at all times. When time is not important we write $|\psi(t)\rangle$ as $|\psi\rangle$, even though the time dependence is still there.

Definition 18. An eigenstate of operator $\hat{\alpha}$ is a state that is an eigenvector of $\hat{\alpha}$.

Definition 19. A measurement corresponding to observable $\hat{a}$, written $\hat{M}^{\star}_{\hat{a}}$, is a non-deterministic pseudo-operator taking states to states in a probabilistic manner.

Definition 20. A measured value a corresponding to the observable $\hat{a}$ for a particular measurement is the eigenvalue of the eigenvector returned by that measurement, which can be expressed as $M^{\star}_{\hat{a}}(|\psi\rangle) = |\hat{a}\hat{M}^{\star}_{\hat{a}}\,|\psi\rangle\,|$.

Definition 21. The expectation value of observable $\hat{a}$ in state $|\psi\rangle$, $\langle\hat{a}\rangle_{\psi}$ is the probability-weighted sum of all possible measured values of $\hat{a}$. That is

$$\langle\hat{a}\rangle_{\psi} = \sum_{a} aP[M^{\star}_{\hat{a}}(|\psi\rangle) = a]$$

Definition 22. A conserved operator is a Hermitian operator $\hat{a}$ that satisfies $\frac{d}{dt}\langle\hat{a}\rangle_{\psi} = 0$ for all states $|\psi\rangle$. A conserved operator is called non-trivial if it is non-zero and not proportional to the identity operator.

5 Appendix: Lemmas

Lemma 1. *For any two vectors $|\chi_1\rangle$ and $|\chi_2\rangle$, $(\langle\chi_1|\,\chi_2\rangle)^* = \langle\chi_2|\,\chi_1\rangle$.*

Proof. For any vectors $|\chi_1\rangle$ and $|\chi_2\rangle$, it is always true that $|\chi_1\rangle^T\,|\chi_2\rangle = |\chi_2\rangle^T\,|\chi_1\rangle$ based on the definition of vectors. Therefore,

$$(\langle\chi_1|\,\chi_2\rangle)^* \overset{D2}{=} \Big(|\chi_1\rangle^\dagger\,|\chi_2\rangle\Big)^* \overset{D2}{=} |\chi_1\rangle^T\,|\chi_2\rangle^*$$

$$= (|\chi_2\rangle^*)^T\,|\chi_1\rangle \overset{D2}{=} |\chi_2\rangle^\dagger\,|\chi_1\rangle \overset{D2}{=} \langle\chi_2|\,\chi_1\rangle$$

□

Lemma 2. *If $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are two operators such that $\langle\chi_1|\,\hat{\alpha}_1\,|\chi_2\rangle = \langle\chi_1|\,\hat{\alpha}_2\,|\chi_2\rangle$ for all vectors $|\chi_1\rangle$ and $|\chi_2\rangle$ then $\hat{\alpha}_1 = \hat{\alpha}_2$.*

Proof. Suppose not; then by Definition 7 there exists a $|\chi_2\rangle$ such that $\hat{\alpha}_1\,|\chi_2\rangle \neq \hat{\alpha}_2\,|\chi_2\rangle$. Therefore,

$$(\hat{\alpha}_1 - \hat{\alpha}_2)\,|\chi_2\rangle \neq 0$$

So choose $|\chi_1\rangle = (\hat{\alpha}_1 - \hat{\alpha}_2)\,|\chi_2\rangle$. Then we have

$$0 \neq \langle\chi_1|\,\chi_1\rangle = \langle\chi_1|\,\hat{\alpha}_1 - \hat{\alpha}_2\,|\chi_2\rangle$$

$$= \langle\chi_1|\,\hat{\alpha}_1\,|\chi_2\rangle - \langle\chi_1|\,\hat{\alpha}_2\,|\chi_2\rangle$$

But this difference has to be zero by the hypothesis, so we have a contradiction. Therefore $\hat{\alpha}_1 = \hat{\alpha}_2$. □

Lemma 3. *Let $\hat{\alpha}_1$ and $\hat{\alpha}_2$ be any two operators. If $\langle\chi|\,\hat{\alpha}_1 = \langle\chi|\,\hat{\alpha}_2$ for all vectors $|\chi\rangle$, then $\hat{\alpha}_1 = \hat{\alpha}_2$.*

Proof. Let $|\chi'\rangle$ be an arbitrary vector. If

$$\langle\chi|\,\hat{\alpha}_1 = \langle\chi|\,\hat{\alpha}_2$$

for all vectors $|\chi\rangle$, then by multiplying both sides to the right by $|\chi'\rangle$ we find that

$$\langle\chi|\,\hat{\alpha}_1\,|\chi'\rangle = \langle\chi|\,\hat{\alpha}_2\,|\chi'\rangle$$

for all vectors $|\chi\rangle$ and $|\chi'\rangle$. So by Lemma 2, $\hat{\alpha}_1 = \hat{\alpha}_2$. □

Lemma 4. *Let $\hat{\alpha}$ be any operator and let c be a constant. Then $(c\hat{\alpha})^\dagger = c^*\hat{\alpha}^\dagger$.*

Proof. Let $|\chi\rangle$ be an arbitrary vector. Then

$$\langle\chi|\,(c\hat{\alpha})^\dagger \overset{D10}{=} (c\hat{\alpha}\,|\chi\rangle)^\dagger$$

$$\overset{D2}{=} (c\hat{\alpha}\,|\chi\rangle)^{T*} = c^*\,(\hat{\alpha}\,|\chi\rangle)^{T*}$$

$$\overset{D2}{=} c^*\,(\hat{\alpha}\,|\chi\rangle)^\dagger \overset{D10}{=} \langle\chi|\,(c^*\hat{\alpha}^\dagger)$$

Therefore, by Lemma 3, $(c\hat{\alpha})^\dagger = c^*\hat{\alpha}^\dagger$. □

Lemma 5. *Let $\hat{\alpha}_1$ and $\hat{\alpha}_2$ be any two operators. Then $(\hat{\alpha}_1 + \hat{\alpha}_2)^\dagger = \hat{\alpha}_1^\dagger + \hat{\alpha}_2^\dagger$.*

Proof. Let $|\chi\rangle$ be an arbitrary vector. Then

$$\langle\chi|\,(\hat{\alpha}_1 + \hat{\alpha}_2)^\dagger \overset{D10}{=} ((\hat{\alpha}_1 + \hat{\alpha}_2)\,|\chi\rangle)^\dagger$$

$$\overset{D8}{=} (\hat{\alpha}_1\,|\chi\rangle + \hat{\alpha}_2\,|\chi\rangle)^\dagger \overset{D2}{=} (\hat{\alpha}_1\,|\chi\rangle + \hat{\alpha}_2\,|\chi\rangle)^{T*}$$

$$= (\hat{\alpha}_1\,|\chi\rangle)^{T*} + (\hat{\alpha}_2\,|\chi\rangle)^{T*} \overset{D2}{=} (\hat{\alpha}_1\,|\chi\rangle)^\dagger + (\hat{\alpha}_2\,|\chi\rangle)^\dagger$$

$$\overset{D10}{=} \langle\chi|\,\alpha_1^\dagger + \langle\chi|\,\alpha_2^\dagger \overset{D8}{=} \langle\chi|\,(\hat{\alpha}_1^\dagger + \hat{\alpha}_2^\dagger)$$

So by Lemma 3, $(\hat{\alpha}_1 + \hat{\alpha}_2)^\dagger = \hat{\alpha}_1^\dagger + \hat{\alpha}_2^\dagger$. □

Lemma 6. *For all operators $\hat{\alpha}$, $(\hat{\alpha}^\dagger)^\dagger = \hat{\alpha}$.*

Proof. Let $|\chi_2'\rangle = \hat{\alpha}|\chi_2\rangle$.

$$(\langle\chi_1|\hat{\alpha}|\chi_2\rangle)^* = (\langle\chi_1|\chi_2'\rangle)^*$$

$$\overset{L1}{=} \langle\chi_2'|\chi_1\rangle \overset{D2}{=} |\chi_2'\rangle^\dagger|\chi_1\rangle$$

$$= (\hat{\alpha}|\chi_2\rangle)^\dagger|\chi_1\rangle \overset{D10}{=} \left\langle\chi_2\middle|\hat{\alpha}^\dagger\middle|\chi_1\right\rangle$$

Let $|\chi_1'\rangle = \hat{\alpha}^\dagger|\chi_1\rangle$ and take the complex conjugate of both ends,

$$\langle\chi_1|\hat{\alpha}|\chi_2\rangle = \left(\left\langle\chi_2\middle|\hat{\alpha}^\dagger\middle|\chi_1\right\rangle\right)^* = (\langle\chi_2|\chi_1'\rangle)^*$$

$$\overset{L1}{=} \langle\chi_1'|\chi_2\rangle \overset{D2}{=} |\chi_1'\rangle^\dagger|\chi_2\rangle$$

$$= (\hat{\alpha}^\dagger|\chi_1\rangle)^\dagger|\chi_2\rangle \overset{D10}{=} \left\langle\chi_1\middle|(\hat{\alpha}^\dagger)^\dagger\middle|\chi_2\right\rangle$$

So by Lemma 2, $(\hat{\alpha}^\dagger)^\dagger = \hat{\alpha}$. □

Lemma 7. *Let $\hat{\alpha}_1$ and $\hat{\alpha}_2$ be any two operators. Then $(\hat{\alpha}_1\hat{\alpha}_2)^\dagger = \hat{\alpha}_2^\dagger\hat{\alpha}_1^\dagger$.*

Proof. Let $|\chi\rangle$ be an arbitrary vector.

$$\langle\chi|(\hat{\alpha}_1\hat{\alpha}_2)^\dagger \overset{D10}{=} ((\hat{\alpha}_1\hat{\alpha}_2)|\chi\rangle)^\dagger$$

$$= (\hat{\alpha}_1(\hat{\alpha}_2|\chi\rangle))^\dagger \overset{D10}{=} (\hat{\alpha}_2|\chi\rangle)^\dagger\hat{\alpha}_1^\dagger \overset{D10}{=} \langle\chi|\hat{\alpha}_2^\dagger\hat{\alpha}_1^\dagger$$

Therefore, by Lemma 3, $(\hat{\alpha}_1\hat{\alpha}_2)^\dagger = \hat{\alpha}_2^\dagger\hat{\alpha}_1^\dagger$. □

Lemma 8. *If $\hat{a}$ is a Hermitian operator, then $\langle\psi|\hat{a}|\psi\rangle$ is real for all vectors $|\psi\rangle$.*

Proof. Let $|\psi'\rangle = \hat{a}|\psi\rangle$.

$$(\langle\psi|\hat{a}|\psi\rangle)^* = (\langle\psi|\psi'\rangle)^* \overset{L1}{=} \langle\psi'|\psi\rangle$$

$$\overset{D2}{=} |\psi'\rangle^\dagger|\psi\rangle = (\hat{a}|\psi\rangle)^\dagger|\psi\rangle \overset{D10}{=} \left\langle\psi\middle|\hat{a}^\dagger\middle|\psi\right\rangle \overset{D11}{=} \langle\psi|\hat{a}|\psi\rangle$$

□

Lemma 9. *If $\hat{a}_1$ and $\hat{a}_2$ are Hermitian operators that satisfy $\langle\psi|\hat{a}_1|\psi\rangle = \langle\psi|\hat{a}_2|\psi\rangle$ for all states $|\psi\rangle$, then $\hat{a}_1 = \hat{a}_2$.*

Proof. Let $\hat{\eta} = \hat{a}_1 - \hat{a}_2$. Then $\hat{\eta}$ is a Hermitian operator that satisfies

$$\langle\chi|\hat{\eta}|\chi\rangle = \langle\chi|\hat{a}_1|\chi\rangle - \langle\chi|\hat{a}_2|\chi\rangle = 0$$

for all vectors $|\chi\rangle$. Now any Hermitian operator can be diagonalized by a unitary matrix (proof omitted), so there exists a unitary matrix $\hat{U}$ such that $\hat{\eta}' = \hat{U}^{-1}\hat{\eta}\hat{U} = \hat{U}^\dagger\hat{\eta}\hat{U}$ is a diagonal matrix. Consider the matrix elements of $\hat{\eta}'$,

$$\langle\chi|\hat{\eta}'|\chi\rangle = \left\langle\chi\middle|\hat{U}^\dagger\hat{\eta}\hat{U}\middle|\chi\right\rangle = \langle\chi'|\hat{\eta}|\chi'\rangle = 0$$

where $|\chi'\rangle = \hat{U}|\chi\rangle$. By letting $|\chi\rangle$ run over all basis vectors, we obtain a set of equations showing that each element on the diagonal of $\hat{\eta}'$ must be zero. But $\hat{\eta}'$ is diagonal, so $\hat{\eta}' = 0$. Therefore, $\hat{\eta} = \hat{U}\hat{\eta}'\hat{U}^{-1} = 0$, which means $\hat{a}_1 = \hat{a}_2$. □

Lemma 10. *If $\hat{a}_1$ and $\hat{a}_2$ are Hermitian operators, then $i[\hat{a}_1, \hat{a}_2]$ is also a Hermitian operator.*

Proof.

$$(i[\hat{a}_1,\hat{a}_2])^\dagger \overset{L4}{=} -i\,([\hat{a}_1,\hat{a}_2])^\dagger$$

$$\overset{D13}{=} -i\,(\hat{a}_1\hat{a}_2 - \hat{a}_2\hat{a}_1)^\dagger$$

$$\overset{L5}{=} -i\left((\hat{a}_1\hat{a}_2)^\dagger - (\hat{a}_2\hat{a}_1)^\dagger\right)$$

$$\overset{L7}{=} -i\,(\hat{a}_2\hat{a}_1 - \hat{a}_1\hat{a}_2) \overset{D13}{=} i[\hat{a}_1,\hat{a}_2]$$

□

Lemma 11. *Let $\hat{\alpha}_0$, $\hat{\alpha}_1$, and $\hat{\alpha}_2$ be any operators. Then $[\hat{\alpha}_0, c_1\hat{\alpha}_1 + c_2\hat{\alpha}_2] = c_1[\hat{\alpha}_0,\hat{\alpha}_1] + c_2[\hat{\alpha}_0,\hat{\alpha}_2]$.*

Proof.

$$[\hat{\alpha}_0, c_1\hat{\alpha}_1 + c_2\hat{\alpha}_2] \overset{D13}{=} \hat{\alpha}_0(c_1\hat{\alpha}_1 + c_2\hat{\alpha}_2) - (c_1\hat{\alpha}_1 + c_2\hat{\alpha}_2)\hat{\alpha}_0$$

$$\overset{D8}{=} c_1\hat{\alpha}_0\hat{\alpha}_1 + c_2\hat{\alpha}_0\hat{\alpha}_2 - c_1\hat{\alpha}_1\hat{\alpha}_0 - c_2\hat{\alpha}_2\hat{\alpha}_0$$

$$\overset{D13}{=} c_1[\hat{\alpha}_0,\hat{\alpha}_1] + c_2[\hat{\alpha}_0,\hat{\alpha}_2]$$

□

Lemma 12. *Let $\hat{A}$, $\hat{B}$, and $\hat{C}$ be any operators. Then $[[\hat{A},\hat{B}],\hat{C}] + [[\hat{B},\hat{C}],\hat{A}] + [[\hat{C},\hat{A}],\hat{B}] = 0$*

Proof.

$$[[\hat{A},\hat{B}],\hat{C}] + [[\hat{B},\hat{C}],\hat{A}] + [[\hat{C},\hat{A}],\hat{B}]$$

$$\overset{D13}{=} [\hat{A}\hat{B} - \hat{B}\hat{A},\hat{C}] + [\hat{B}\hat{C} - \hat{C}\hat{B},\hat{A}] + [\hat{C}\hat{A} - \hat{A}\hat{C},\hat{B}]$$

$$\overset{D13}{=} (\hat{A}\hat{B} - \hat{B}\hat{A})\hat{C} - \hat{C}(\hat{A}\hat{B} - \hat{B}\hat{A})$$
$$+ (\hat{B}\hat{C} - \hat{C}\hat{B})\hat{A} - \hat{A}(\hat{B}\hat{C} - \hat{C}\hat{B})$$
$$+ (\hat{C}\hat{A} - \hat{A}\hat{C})\hat{B} - \hat{B}(\hat{C}\hat{A} - \hat{A}\hat{C})$$

$$\overset{D8}{=} \hat{A}\hat{B}\hat{C} - \hat{B}\hat{A}\hat{C} - \hat{C}\hat{A}\hat{B} + \hat{C}\hat{B}\hat{A}$$
$$+ \hat{B}\hat{C}\hat{A} - \hat{C}\hat{B}\hat{A} - \hat{A}\hat{B}\hat{C} + \hat{A}\hat{C}\hat{B}$$
$$+ \hat{C}\hat{A}\hat{B} - \hat{A}\hat{C}\hat{B} - \hat{B}\hat{C}\hat{A} + \hat{B}\hat{A}\hat{C}$$
$$= 0$$

□

Lemma 13. *Let $\hat{A}$, $\hat{B}$, and $\hat{C}$ be any operators. Then $[\hat{A},\hat{B}\hat{C}] = [\hat{A},\hat{B}]\hat{C} + \hat{B}[\hat{A},\hat{C}]$*

Proof.

$$[\hat{A},\hat{B}]\hat{C}+\hat{B}[\hat{A},\hat{C}]$$

$$\overset{D13}{=}(\hat{A}\hat{B}-\hat{B}\hat{A})\hat{C}+\hat{B}(\hat{A}\hat{C}-\hat{C}\hat{A})$$

$$\overset{D8}{=}\hat{A}\hat{B}\hat{C}-\hat{B}\hat{A}\hat{C}+\hat{B}\hat{A}\hat{C}-\hat{B}\hat{C}\hat{A}$$

$$=\hat{A}\hat{B}\hat{C}-\hat{B}\hat{C}\hat{A}\overset{D13}{=}[\hat{A},\hat{B}\hat{C}]$$

□

Lemma 14. *If $\hat{a}_1$ and $\hat{a}_2$ are operators that satisfy $[\hat{a}_1,\hat{a}_2]=c$ for some constant c, then $[\hat{a}_1,\hat{a}_2^k]=ck\hat{a}_2^{k-1}$ for all integers $k\geq 1$.*

Proof. We procede by induction on k. For the base case, $k=1$, we have $[\hat{a}_1,\hat{a}_2]=c$, which is true by the assumption of the theorem. Now assume that the result holds for $k=N-1$. Then for $k=N$ we have

$$[\hat{a}_1,\hat{a}_2^N]=[\hat{a}_1,\hat{a}_2\hat{a}_2^{N-1}]\overset{L13}{=}[\hat{a}_1,\hat{a}_2]\hat{a}_2^{N-1}+\hat{a}_2[\hat{a}_1,\hat{a}_2^{N-1}]$$

Now using the inductive hypothesis,

$$=c\hat{a}_2^{N-1}+\hat{a}_2\left(c(N-1)\hat{a}_2^{N-2}\right)=cN\hat{a}_2^{N-1}$$

So the theorem holds for $k=N$ and the induction is complete. □

Lemma 15. *If $\hat{a}_1$ and $\hat{a}_2$ are operators that satisfy $[\hat{a}_1,\hat{a}_2]=c$ for some constant c and $f(x)$ is an analytic function, then $[\hat{a}_1,f(\hat{a}_2)]=cf'(\hat{a}_2)$.*

Proof. Since $f(x)$ is an analytic function, we can write

$$f(x)=\sum_{k=0}^{\infty}f^{(k)}(0)x^k$$

where $f^{(k)}(0)=\frac{d^kf(x)}{dx}\Big|_{x=0}$. Therefore

$$[\hat{a}_1,f(\hat{a}_2)]=\left[\hat{a}_1,\sum_{k=0}^{\infty}f^{(k)}(0)\hat{a}_2^k\right]$$

$$\overset{L11}{=}\sum_{k=0}^{\infty}f^{(k)}(0)[\hat{a}_1,\hat{a}_2^k]\overset{L14}{=}c\sum_{k=0}^{\infty}f^{(k)}(0)k\hat{a}_2^{k-1}$$

Now,

$$f'(x)=\frac{d}{dx}\sum_{k=0}^{\infty}f^{(k)}(0)x^k=\sum_{k=0}^{\infty}f^{(k)}(0)kx^{k-1}$$

Multiplying by c and replacing x with $\hat{a}_2$ we get the same result, so $[\hat{a}_1,f(\hat{a}_2)]=cf'(\hat{a}_2)$. □

Lemma 16. *If $\hat{a}_1$ and $\hat{a}_2$ are operators that satisfy $[\hat{a}_1,\hat{a}_2]=c$ where c is a constant, then any analytic function $f(\hat{a}_1,\hat{a}_2)$ can be expanded in a power series as $f(\hat{a}_1,\hat{a}_2)=\sum_{k=0}^{\infty}f_k(\hat{a}_1)\hat{a}_2^k$ for some functions f_k.*

Proof. Since $f(\hat{a}_1,\hat{a}_2)$ is analytic, it can be expressed as a series expansion, but the terms do not necessarily have all the $\hat{a}_2$ operators on the right-hand end. It remains to show that any term made up of a product of $\hat{a}_1$ and $\hat{a}_2$ operators can be expressed as a sum of terms with all $\hat{a}_2$ operators at the right-hand end. We proceed by induction on the number of $\hat{a}_1$ and $\hat{a}_2$ operators in the term. The base case $n=0$ is true by setting f_0 appropriately. Now assume the inductive hypothesis holds for all n up to N. Let $\hat{a}^{(N)}$ represent an arbitrary term containing N operators. Then any term with $N+1$ operators can be written as $\hat{a}^{(N)}\hat{a}_1$ or $\hat{a}^{(N)}\hat{a}_2$. By the inductive hypothesis, we can write $\hat{a}^{(N)}$ in the required form, so the second case is done because it just adds an extra $\hat{a}_2$ to the end of each term. As for the first case, if $\hat{a}^{(N)}$ contains no instances of $\hat{a}_2$, then we are done. If it does contain an instance of $\hat{a}_2$, we can assume it is at the end by the inductive hypothesis: $\hat{a}^{(N-1)}\hat{a}_2\hat{a}_1=\hat{a}^{(N-1)}(\hat{a}_1\hat{a}_2-c)$. Now $\hat{a}^{(N-1)}\hat{a}_1$ can be re-written in the proper form by the inductive hypothesis so the first term is in the desired form and the second term is too, so all cases have been shown. □

6 Bonus Lemmas

Lemma 17. *If for all states $|\psi\rangle$, $|\psi\rangle$ is an eigenstate of the linear operator $\hat{\alpha}$, then $\hat{\alpha}\propto\hat{I}$.*

Proof. Consider two orthogonal states $|\psi_1\rangle$ and $|\psi_2\rangle$. Let their $\hat{\alpha}$ eigenvalues be α_1 and α_2 respectively. Now $\frac{1}{\sqrt{2}}(|\psi_1\rangle+|\psi_2\rangle)$ must also be an eigenstate, so there is a constant α_0 such that

$$\frac{1}{\sqrt{2}}\alpha_0\left(|\psi_1\rangle+|\psi_2\rangle\right)\overset{D14}{=}\frac{1}{\sqrt{2}}\hat{\alpha}\left(|\psi_1\rangle+|\psi_2\rangle\right)$$

$$\overset{D9}{=}\frac{1}{\sqrt{2}}\left(\hat{\alpha}\,|\psi_1\rangle+\hat{\alpha}\,|\psi_2\rangle\right)=\frac{1}{\sqrt{2}}\left(\alpha_1\,|\psi_1\rangle+\alpha_2\,|\psi_2\rangle\right)$$

Multiplying by $\langle\psi_1|$ to the left we get $\alpha_0=\alpha_1$. Multiplying by $\langle\psi_2|$ to the left we get $\alpha_0=\alpha_2$. Therefore $\alpha_1=\alpha_2$ for any pair of orthogonal states. Now let $|\psi_1\rangle$ and $|\psi_2\rangle$ be any two states, not necessarily orthogonal. Define

$$|\phi\rangle=|\psi_2\rangle-\langle\psi_1|\,\psi_2\rangle\,|\psi_1\rangle$$

Then $\langle\psi_1|\,\phi\rangle=\langle\psi_1|\,\psi_2\rangle-\langle\psi_1|\,\psi_2\rangle=0$, which means that $|\phi\rangle$ and $|\psi_1\rangle$ are orthogonal. So by what was just shown, $|\phi\rangle$ and $|\psi_1\rangle$ must have the same $\hat{\alpha}$ eigenvalue, say α_1. Then

$$\hat{\alpha}\,|\psi_2\rangle = \hat{\alpha}(\langle\psi_1|\,\psi_2\rangle\,|\psi_1\rangle + |\phi\rangle)$$

$$\overset{D9}{=} \langle\psi_1|\,\psi_2\rangle\,\hat{\alpha}\,|\psi_1\rangle + \hat{\alpha}\,|\phi\rangle$$

$$= \langle\psi_1|\,\psi_2\rangle\,\alpha_1\,|\psi_1\rangle + \alpha_1\,|\phi\rangle = \alpha_1\,|\psi_2\rangle$$

Therefore $|\psi_1\rangle$ and $|\psi_2\rangle$ have the same $\hat{\alpha}$ eigenvalue. Since $|\psi_1\rangle$ and $|\psi_2\rangle$ were arbitrary, this means that all states have the same eigenvalue, say c. Then

$$\hat{\alpha}\,|\psi\rangle = c\,|\psi\rangle$$

Therefore by Definition 7, $\hat{\alpha} = c\hat{I}$, which means $\hat{\alpha} \propto \hat{I}$. □

Lemma 18. *All eigenvalues of Hermitian operators are real.*

Proof. Let $\hat{a}$ be a Hermitian operator and let $|a\rangle$ be an eigenvector of $\hat{a}$ with eigenvalue a so that $\hat{a}\,|a\rangle = a\,|a\rangle$. Taking the Hermitian conjugate and using Definitions 10 and 2, $\langle a|\,\hat{a}^\dagger = \langle a|\,a^*$.

$$a = \langle a|\,\hat{a}\,|a\rangle \overset{D11}{=} \left\langle a\middle|\,\hat{a}^\dagger\,\middle|a\right\rangle = a^*$$

which shows that a is real. □

Lemma 19. *A transformation $\hat{U}$ on states preserves the length of all vectors if an only if the transformation is unitary, i.e. $\hat{U}^\dagger = \hat{U}^{-1}$.*

Proof. Assume $\hat{U}$ preserves the length of all vectors, so

$$|\hat{U}\,|\chi\rangle\,|^2 = |\,|\chi\rangle\,|^2$$

By Definition 5,

$$\left\langle\chi\middle|\,\hat{U}^\dagger\hat{U}\,\middle|\chi\right\rangle = \left\langle\chi\middle|\,\hat{I}\,\middle|\chi\right\rangle$$

The operator $(\hat{U}^\dagger\hat{U})$ is Hermitian because $(\hat{U}^\dagger\hat{U})^\dagger \overset{L7}{=} \hat{U}^\dagger\hat{U}^{\dagger\dagger} \overset{L6}{=} \hat{U}^\dagger\hat{U}$. So by Lemma 9,

$$\hat{U}^\dagger\hat{U} = \hat{I}$$

which proves the forward implication. Now assume $\hat{U}$ is unitary. Then for any state $|\chi\rangle$,

$$|\hat{U}\,|\chi\rangle\,|^2 \overset{D5}{=} \left\langle\chi\middle|\,\hat{U}^\dagger\hat{U}\,\middle|\chi\right\rangle \overset{D12}{=} \left\langle\chi\middle|\,\hat{I}\,\middle|\chi\right\rangle \overset{D5}{=} |\,|\chi\rangle\,|^2$$

Since norms are always positive, taking the square root of both ends proves the reverse implication. □