1.5. The Schmidt decomposition

We have seen that for any given orthonormal basis $\{\eta_\alpha\}$ of $\mathcal{H}_2$, any vector Ψ in $\mathcal{H}_1 \otimes \mathcal{H}_2$ can be written in the form

$$\Psi = \sum_\alpha |\zeta_\alpha\rangle|\epsilon_\alpha\rangle,$$

and we shall now show that by being more flexible about the choice of basis for $\mathcal{H}_2$ it is possible to get a more symmetric expansion.

Schmidt Decomposition Theorem 1.5.1. Any vector $\Psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ can be expressed in the form

$$\Psi = \sum_j c_j |\xi_j\rangle|\eta_j\rangle,$$

for non-negative real c_j, and orthonormal sets $\xi_j \in \mathcal{H}_1$ and $\eta_j \in \mathcal{H}_2$ $(j = 1, 2, \ldots)$. There are density operators ρ_1 on $\mathcal{H}_1$ and ρ_2 on $\mathcal{H}_2$ such that

$$\langle\Psi|(A \otimes 1)\Psi\rangle = \mathrm{tr}[A\rho_1], \qquad \langle\Psi|(1 \otimes B)\Psi\rangle = \mathrm{tr}[B\rho_2],$$

for all observables A and B on $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, and the $\{\xi_j\}$ may be chosen to be the eigenvectors of ρ_1 corresponding to non-zero eigenvalues $\{p_j\}$, the vectors $\{\eta_j\}$, the corresponding eigenvectors for ρ_2, and the positive scalars $c_j = \sqrt{p_j}$.

Proof. We already know that

$$\langle\Psi|(A \otimes 1)\Psi\rangle = \mathrm{tr}[A\rho_1]$$

for a suitable density operator ρ_1. The operator ρ_1 on $\mathcal{H}_1$ is self-adjoint and so there is an an orthonormal basis of eigenvectors $|\xi_j\rangle$ such that

$$\rho_1\xi_j = p_j|\xi_j\rangle,$$

with non-negative p_j. We now interchange the roles of $\mathcal{H}_1$ and $\mathcal{H}_2$ in the argument leading to the earlier expansion. Since the $|\xi_j\rangle$ form a basis, we can pick an arbitrary basis $|\phi_j\rangle$ of $\mathcal{H}_2$ and expand Ψ as

$$\Psi = \sum_{j,k} c_{jk}|\xi_j\rangle|\phi_j\rangle = \sum_j |\xi_j\rangle(\sum_k c_{jk}|\phi_j\rangle),$$

so that setting $\sum_k c_{jk}|\phi_j\rangle = c_j|\widetilde{\eta}_j\rangle$ for some unit vector $|\eta_j\rangle$ and positive c_j we have

$$\Psi = \sum_j c_j|\xi_j\rangle|\widetilde{\eta}_j\rangle.$$

Now, we know that

$$\mathrm{tr}[A\rho_1] = \sum_j \langle\xi_j|A\rho_1\xi_j\rangle = \sum_j p_j\langle\xi_j|A\xi_j\rangle,$$

and calculate that

$$\langle\Psi|(A \otimes 1)\Psi\rangle = \sum_j \bar{c}_j c_k \langle\xi_j|A\xi_k\rangle\langle\widetilde{\eta}_j|\widetilde{\eta}_k\rangle,$$

so, by comparison, we deduce that $\langle\widetilde{\eta}_j|\widetilde{\eta}_k\rangle$ vanishes unless $j = k$, and since the $\widetilde{\eta}_j$ are unit vectors when $j = k$ we get $|c_j|^2 = p_j$, so that we may take $c_j\sqrt{p_j}$. Clearly we can drop the terms where $p_j = 0$ from the sum, to ensure that we have only positive c_j.

The formulae are now symmetric in the two spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ so that there must also exist a density operator ρ_2 such that

$$\langle\Psi|(1\otimes B)\Psi\rangle = \mathrm{tr}[B\rho_2],$$

and we may check that the $\widetilde{\eta}_j$ are its eigenvectors, with the non-zero eigenvalues $p_j = c_j^2$. ◇

One particularly striking consequenc eof this proof is the fact that the two partial traces ρ_1 and ρ_2 have the same non-zero eigenvalues. This is all the more surprising when one recalls that $\mathcal{H}_1$ and $\mathcal{H}_2$ need not have the same dimension. Nonetheless they must have the same non-zero eigenvalues. This means that if $\dim(\mathcal{H}_1) < \dim(\mathcal{H}_2)$ then, because ρ_1 cannot have more than $\dim(\mathcal{H}_1)$ positive eigenvalues, neither can ρ_2.

The Schmidt decomposition, gives an elegant form for the entanglement of vectors describing states of a system formed from two subsystems. (If there is only a single term in the Schmidt decomposition then Ψ is not entangled, but otherwise it is entangled and the number of terms in the sum the it Schmidt rank gives a measure of how entangled it is. Unfortunately there is no such expansion of general vectors when three or more subsystems are combined, and good measures of entanglement are harder to find.

We conclude this part by noting that one does not need to start with a pure state $\Psi \in \mathcal{H}_1\otimes\mathcal{H}_2$ in order to get density operators ρ_1 and ρ_2 on the two subsystems $\mathcal{H}_1$ and $\mathcal{H}_2$¿

Theorem 1.5.2. Let ρ be a density operator on $\mathcal{H}_1\otimes\mathcal{H}_2$ then there are density operators ρ_1 and ρ_2 on $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, such that

$$\mathrm{tr}[(A\otimes 1)\rho] = \mathrm{tr}[A\rho_1], \qquad \mathrm{tr}[(1\otimes B)\rho] = \mathrm{tr}[B\rho_2],$$

for all observables A and B on $\mathcal{H}_1$ and $\mathcal{H}_2$.

Proof. The density operator ρ is self-adjoint and so has eigenvectors Ψ_α with nonnegative eigenvalues q_α. We know from earlier calculations that

$$\mathrm{tr}[(A\otimes 1)\rho] = \sum_\alpha q_\alpha\langle\Psi_\alpha|(A\otimes 1)\Psi_\alpha\rangle,$$

and for Ψ_α there is a density operator $\rho_{1\alpha}$ such that

$$\langle\Psi_\alpha|(A\otimes 1)\Psi_\alpha\rangle = \mathrm{tr}[A\rho_{1\alpha}],$$

which combine to give

$$\mathrm{tr}[(A\otimes 1)\rho] = \sum_\alpha q_\alpha \mathrm{tr}[A\rho_{1\alpha}].$$

It is therefore clear that by taking $\rho_1 = \sum_\alpha q_\alpha\rho_{1\alpha}$ we obtain $\mathrm{tr}[(A\otimes 1)\rho] = \mathrm{tr}[A\rho_1]$. The operator ρ_1 defined by this formula is positive since for any $\phi\in\mathcal{H}_1$ we have

$$\langle\phi|\rho_1\phi\rangle = \sum_\alpha q_\alpha\langle\phi|\rho_{1\alpha}\phi\rangle \geq 0,$$

since it is a non-negative linear combination of non-negative terms. Moreover, by taking $A = 1$ we get $\mathrm{tr}[\rho] = \mathrm{tr}[\rho_1]$, from which it follows that $\mathrm{tr}[\rho_1] = 1$. ◇

We should remark that the traces involved in the various different expressions are actually traces over different spaces. For example, in $\mathrm{tr}[(A\otimes 1)\rho]$ the trace is taken over $\mathcal{H}_1\otimes\mathcal{H}_2$, whilst in $\mathrm{tr}[A\rho_1]$ it goes only over $\mathcal{H}_1$. It is for this reason that ρ_1 and ρ_2 are called *partial traces* of ρ, because, in passing from $\mathrm{tr}[(A\otimes 1)\rho]$ to $\mathrm{tr}[A\rho_1]$ the trace over $\mathcal{H}_2$ has already been carried out.

1.6. POVM measurements

When doing a measurement on a system described by $\mathcal{H}_1$ we typically use some kind of measuring instrument with its own space $\mathcal{H}_2$, and states $\{\epsilon_j \in \mathcal{H}_2\}$ labelled by the possible outcomes. In order to distinguish these states we choose the vectors to be orthonormal.

Suppose now that we start with the system desrcibed by a vector $\Psi \in \mathcal{H}_1$ and the measuring instrument in some rest state Ω. In the course of the measurement the system and instrument interact, and for a good measurement we want them to evolve to a state

$$\widetilde{\Psi} = \sum_j |\psi_j\rangle|\epsilon_j\rangle$$

for suitable vectors $|\psi_j\rangle \in \mathcal{H}_1$. In other words, the measurement should send the system and measuring instrument into an entangled state. Ideally, if we are measuring some observable A the ψ_j would be its eigenvectors corresponding to eigenvalues which are labelled by the measuring instrument state ϵ_j.

The interaction which occurs should be linear, either the unitary evolution described by solving Schrödinger's equation, or a projection coming from a von Neumann measurement. The instrument labels are independent of the initial system state Ψ, so each $|\psi_j\rangle$ that depends linearly on Ψ, and can be written as $M_j|\Psi\rangle$ for a suitable linear transformation M_j on $\mathcal{H}_1$. We therefore have

$$|\Psi\rangle|\Omega\rangle \mapsto \sum_j |M_j\Psi\rangle \otimes |\epsilon_j\rangle.$$

The partial trace density operator ρ_1 for the original state $|\Psi\rangle|\Omega\rangle$ is given by $\langle\Psi|A\Psi\rangle = \mathrm{tr}[A\rho_1]$ for any observable A on $\mathcal{H}_1$, but the expectation value after the measurement is, by the orthonormality of the $|\epsilon_j\rangle$,

$$\sum_{j,k}\langle M_j\Psi|AM_k\Psi\rangle\langle\epsilon_j|\epsilon_k\rangle = \sum_j\langle M_j\Psi|AM_j\Psi\rangle = \sum_j\langle\Psi|M_j^*AM_j\Psi\rangle = \sum_j \mathrm{tr}[M_j^*AM_j\rho_1].$$

By the usual change of order of terms in the trace this can also be written as $\sum_j \mathrm{tr}[AM_j\rho_1M_j^*]$, showing that the measurement has the effect of sending ρ_1 to $\sum_j M_j\rho_1M_j^*$. Taking $A = 1$, the final expectation value is $\sum_j \mathrm{tr}[M_j^*M_j\rho_1]$, and we might expect this to be 1. However, our measuring apparatus may not be perfect, indeed its states ϵ_j may not cover all the possible outcomes, and in general we can only expect $\sum_j \mathrm{tr}[M_j^*M_j\rho_1] \leq 1 = \mathrm{tr}[\rho_1]$. (Another way of thinking about this is that so long as we have normal unitary evolutions described by solving Schrödinger's equation then probabilities are conserved and we should end up with 1, but projections, representing von Neumann measurements, do not preserve norms and that permits us to get answers which my be smaller.) This inequality is achieved for all ρ_1 if $\sum_j M_j^*M_j \leq 1$. in the sense that $1 - \sum_j M_j^*M_j$ is a positive operator.

Definition 1.6.1. A *generalised positive operator-valued measurement* takes a density operator ρ_1 on the system to $\sum_j M_j\rho_1M_j^*$, with $\sum_j M_j^*M_j \leq 1$. If there is equality then it is said to be a *positive operator-valued measurement (POVM)*. The operators M_j are called *Kraus operators.*

A special case occurs for von Neumann measurements with $\Psi \mapsto P\Psi$, where P is a projection (usually onto some eigenstate of the observable A which is being measured). Then $\rho_1 = |\Psi\rangle\langle\Psi| \mapsto P|\Psi\rangle\langle\Psi|P^* = P\rho_1P^*$, showing that we have a measurement with a single Kraus operator $M_1 = P$.

Definition 1.6.2. If the Kraus operators are projections then on ereferes to a *projection-valued* (or *projective*) measurement.

Originally only projective measurments were used, but it turned out that positive operator-valued measurements could also be implemented and, in some situations, offered real advantages. (One early suggested application was for detecting and decoding signals from deep space probes, where only a few photons a second might arrive back.)

1.7. Reconstructing a measurement from the Kraus operators*

This section is not strictly part of the course, but is included for anybody who wants more detail. It is possible to reconstruct a lot of details of measurements from the Kraus operators alone.

We first introduce an auxiliary (or ancilliary) set of vectors ϵ_j indexed by the same set as the M_j, and consider the map

$$M : \psi \otimes \epsilon_0 \mapsto \sum_j |M_j\psi\rangle|\epsilon\rangle.$$

The condition that $\sum_j M_j^* M_j \leq 1$ tells us that

$$\|M\psi\|^2 = \sum_{j,k}\langle M_j\psi|M_k\psi\rangle\langle\epsilon_j|\epsilon_k\rangle = \sum_j\langle\psi|M_j^*M_j\psi\rangle \leq \|\psi\|^2,$$

so that M is a contraction, that is it diminishes norms. We can now show that every generalised POVM is obtained form a unitary oeprator and a projection, just as our earlier discussion of normal evolution and von Neumann measurements might suggest.

Naimark's Dilation Theorem 1.7.1. *Given a contraction map M on a space $\mathcal{H}$, there exist a space $\mathcal{K}$ containing $\mathcal{H}$ as a subspace and a unitary mapping U on $\mathcal{K}$, such that $M = PU$, where P is the projection onto $\mathcal{H}$.*

Proof. We take $\mathcal{K} = \mathcal{H} \oplus \mathcal{H}$ and identify $\mathcal{H}$ with the first summand, so that the projection P maps $(\psi_1 \oplus \psi_2$ to ψ_1.

Since $1 - M^*M$ is positive we can form its square root $\sqrt{1 - M^*M}$, and similarly for $\sqrt{1 - MM^*}$. We can then write U in block matrix form as

$$U = \begin{pmatrix} M & \sqrt{1 - MM^*} \\ \sqrt{1 - M^*M} & -M^* \end{pmatrix}.$$

We check that

$$\begin{aligned} U^*U &= \begin{pmatrix} M^* & \sqrt{1 - M^*M} \\ \sqrt{1 - MM^*} & -M \end{pmatrix}\begin{pmatrix} M & \sqrt{1 - MM^*} \\ \sqrt{1 - M^*M} & -M^* \end{pmatrix} \\ &= \begin{pmatrix} M^*M + (1 - M^*M) & M^*\sqrt{1 - MM^*} - \sqrt{1 - M^*M}M^* \\ \sqrt{1 - MM^*}M - M\sqrt{1 - M^*M} & MM^* + (1 - MM^*) \end{pmatrix} \\ &= \begin{pmatrix} 1 & M^*\sqrt{1 - MM^*} - \sqrt{1 - M^*M}M^* \\ \sqrt{1 - MM^*}M - M\sqrt{1 - M^*M} & 1 \end{pmatrix}. \end{aligned}$$

Now it is clear that $(1 - MM^*)M = M(1 - M^*M)$, but then checking the action on the eigenvectors of $(1 - M^*M)$ which are also the eigenvectors of $\sqrt{1 - M^*M}$, we see that $\sqrt{1 - MM^*}M = M\sqrt{1 - M^*M}$ and, similarly, the other off-diagonal entry vanishes, showing that $U^*U = 1$, Similarly, we have $UU^* = 1$, so that U is unitary. Finally

$$PU(\psi \oplus 0) = P(M\psi \oplus \sqrt{1 - M^*M}\psi) = M\psi,$$

so that $PU = M$. ◇

We can now use J to identify $\mathcal{H}$ and its image in $\mathcal{K}$, and $J^* = P$ with the projection onto that first component. We then have an identification $A = PU$, with a unitary followed by a projection.

It is possible to generalise this still further and look at the individual terms $M_j\rho M_j^*$, but the discussion is very similar and we shall leave it at this point.

2 Quantum Information Processing

2.1. Quantum Gates

In normal computers and normal digital communications information is encoded as a string of bits, that is 0s and 1s. Quantum computers replace each bit by a vector in the two-dimensional space spanned by two vectors $e_0 = |0\rangle$ and $e_1 = |1\rangle$.

Definition 2.1.1. A quantum bit or *qubit* is encoded by a two-dimensional quantum state, that is a vector (up to multiples) in a two dimensional space $V \cong {}^2$.

One could of course use a three state quantum system and model it by ^{3}, a so-called qutrit, or even higher dimensional spaces, but we shall limit the discussion to qubits, which are the most widely discussed and used. Higher dimensional spaces appear naturally when one prcesses several qubits. For examples two qubits require two systems described by V, and so overall one has the four-dimensional space $V \otimes V$. Similarly for three qubits one uses the eight dimensional space $V \otimes V \otimes V$.

Definition 2.1.2. A k qubit system can be described using the space $\otimes^k V$, defined inductively by $\otimes^1 V = V$, and $\otimes^k V = (\otimes^{k-1} V) \otimes V$.

In a classical computer the bits are processed using a sequence of digital logic gates, but in a quantum computer one carries out quantum transformations, that is unitary transformations of the spaces.

The simplest classical gates work on a single bit of information. There are only four possible such gates, since each must take 0 to either 0 or 1, and likewise take 1 to 0 or 1, two choices of two possible outcomes. Two of these four are rather boring since they either take both 0 and 1 to 0, or both 0 and 1 to 1. Another is just the identity which takes 0 to 0 and 1 to 1. The only really interesting classical gate is the NOT gate which takes 0 to 1 and 1 to 0. (In logical applications we could think of 0 as false and 1 as true, and these are reversed by the NOT gate, whence its name.)

In quantum information processing even one qubit allows many more possibilities, since, as already remarked, any linear transformation T on ^{2} can operate on a qubit, sending $|\psi\rangle \mapsto T|\psi\rangle$. Since vectors differing only by scalar multiples define the same quantum state, we may as well restrict our attention to normalised vectors and unitary transformations T, but there are still a lot of those.

The NOT gate also works for quantum systems taking $|0\rangle$ to $|1\rangle$, and $|1\rangle$ to $|0\rangle$. This linear transformation defined by its effect on the basis has matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_1$$

where σ_1 is the first Pauli spin matrix. As an indication of the greater flexibility available in quantum computing we note that since

$$[\frac{1}{2}(1+i)(1-i\sigma_1)]^2 = \frac{1}{4}(1-1+2i)(1-\sigma_1^2-2i\sigma_1) = \frac{1}{4}(2i)(-2i\sigma_1) = \sigma_1,$$

the NOT gate in quantum computers has a square root, with matrix

$$\frac{1}{2}(1+i)(1-i\sigma_1) = \frac{1}{2}(1+i)\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}.$$

This gate is important because it can be implemented using a beam splitter.

However, there are other single qubit gates. Amongst the most important are the unitary phase shift gates which fix $|0\rangle$ but shift the phase of $|1\rangle$ by $\exp(i\phi)$, that is they have matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} = e^{i\frac{1}{2}\phi} \begin{pmatrix} e^{-i\frac{1}{2}\phi} & 0 \\ 0 & e^{i\frac{1}{2}\phi} \end{pmatrix}.$$

These can also be implemented by simple physical means.

Another useful and easily implemented gate is the Hadamard gate with matrix

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}}(\sigma_1 + \sigma_3).$$

Of course, one expects computers to process more than just one bit or qubit. In classical computers the controlled not or CNOT gate provides an example. Here the first bit controls what happens to the second: if the first bit is 0 then the second bit is left alone, if the first bit is one then NOT is applied to the second. The quantum analogue fixes $e_0 \otimes \psi$ but sends $e_1 \otimes e_j$ to $e_1 \otimes e_{1-j}$. In termsof the basis $\{e_{00}, e_{01}, e_{10}, e_{11}\}$, where $e_{jk} = e_j \otimes e_k$, the matrix is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

It can also be written in terms of the Pauli spin matrices as

$$\frac{1}{2}(1 \otimes 1 + 1 \otimes \sigma_1 + \sigma_3 \otimes 1 - \sigma_3 \otimes \sigma_1),$$

Definition 2.1.3. A set of quantum gates is called *universal* for a system if they generate the whole group of unitary transformations of the state space of the system, in the sense that given any unitary transformation U and any $\epsilon > 0$, there is a a finite product of the unitary gates $U_1 U_2 \ldots U_n$ which is within a distance ϵ of U, that is

$$\|(U - U_1 U_2 \ldots U_n\|_{\mathrm{tr}} < \epsilon$$

where the norm is the trace norm introduced earlier.

It can be shown that the phase shift gate with $\phi = \pi/4$ and the Hadamard gate from a universal set for processing a single qubit, that is for the unitary operators on ^{2}, and those two operators on eahc of two qubits, together with the controlled NOT, form a universal set for two qubits.

2.2. A sketch of the universality theorems*

The proof of this falls outside the scope of this course but for those who are interested we outline how the proof works.

Theorem 2.2.1. In the two dimensional space V any two noncommuting unitary transformations neither of which has finite order form a universal set of gates.

Proof. Any unitary operator on V can be diagonalised, by the same procedure used for self-adjoint operators. (Alternatively, we can derive the unitary result from the self-adjoint result. We may as well assume that 1 is not an eigenvalue of U, since we can always multiply U by a suitable number of unit modulus to ensure this. Then we can form $A = i(U+1)(U-1)^{-1}$, which is self-adjoint, since

$$A^* = -i(U^*-1)^{-1}(U^*+1) = -i(1-U)^{-1}(1+U) = i(1+U)(U-1)^{-1} = A.$$

This self-adjoint operator admits a basis of eigenvectors $\{e_j\}$, and since $U = (iA-1)(iA+1)^{-1}$ these are also eigenvectors of U. If U is a unitary operator which does not have finite order then its eigenvalues have the form $\exp(i\theta_j)$ with $\theta_j/2\pi$ irrational. The powers of these generate all complex numbers of unit modulus. We can therefore assume that we have unitary operators of the form

$$U_\lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix}$$

for all λ such that $\overline{\lambda}\lambda = 1$. We shall call these U-type operators. The second unitary operator V, generates a similar group but with respect to a different basis (since they do not commute). Now any unitary operator can be expressed as a product of a U-type operator then a V-type then another U-type, completing the proof. ◇

The phase shift $T = T_{\frac{1}{4}\pi}$ and Hadamard gate H actually do have finite order, but their products TH and HT do not, so that this theorem applies, and shows that they give a universal set.

Theorem 2.2.2. In a two qubit space a universal set is provided by the single bit unitary transformations and the CNOT gates linking any two.

Proof. By the previous result we know that we can get all unitary operators on each single qubit space. and one then shows that with CNOT these generate the unitary group on the four-dimensional two qubit space.

Together these show that a very small number of gates is sufficient to give a universal set, for example, a two qubit system needs only two gates for each bit plus a CNOT gate.

2.3. No-cloning

In classical information processing we are always making copies, whether on paper using the photocopier, or electronically in our computer. The No-cloning Theorem says that it is not possible to make an exact quantum copies of arbitrary quantum states, that is there is no good quantum copier which can start with an arbitrary quantum state $|\psi\rangle \in {}^2$ and reliably turn out a copy so that we end up with $|\psi\rangle|\psi\rangle$.

The No-cloning Theorem 2.3.1. There is no linear map $C : {}^2 \mapsto \otimes^{22}$ such that $C|\psi\rangle = |\psi\rangle|\psi\rangle$ for all $\psi \in {}^2$.

Proof. A naive proof would simply not that the copying map C should work for $\lambda|\psi\rangle$ for any non-zero scalar λ, so that

$$C(\lambda|\psi\rangle) = (\lambda|\psi\rangle)(\lambda|\psi\rangle) = \lambda^2|\psi\rangle|\psi\rangle.$$

However, the linearity of C also gives

$$C(\lambda|\psi\rangle) = \lambda C|\psi\rangle = \lambda|\psi\rangle\lambda|\psi\rangle,$$

and this cannot be reconciled with the earlier formula unless $\lambda^2 = \lambda$ for all λ, and this is clearly false.

Now, this proof is not entirely convincing, since at least $\lambda|\psi\rangle$ and $\lambda^2|\psi\rangle|\psi\rangle$ are multiples, and so define the same quantum state. We therefore give a second more convincing argument. By linearity

$$C(c_0|0\rangle + c_1|1\rangle) = c_0^2|0\rangle|0\rangle + c_0c_1(|1\rangle|0\rangle + |0\rangle|1\rangle) + c_1^2|1\rangle|1\rangle.$$

On the other hand, this should be the same as

$$c_0C(|0\rangle) + c_0C(|0\rangle) = c_0|0\rangle|0\rangle + c_1|1\rangle|1\rangle.$$

Comparing coefficients of the different basis vectors shows that

$$c_0^2 = c_0, \qquad c_1^2 = c_1, \qquad c_0c_1 = 0.$$

The last equality shows that one of the coefficients c_0, c_1 must vanish, so that $|\psi\rangle$ is actually a multiple of either $|0\rangle$ or $|1\rangle$. ◇

The compensation for this is that, although quantum theory forbids making copies of a state, it allows us to reproduce a state wherever we like, at the expense of losing the original. This is the basis of quantum teleportation.