All this to say that there is a strong connection rotations in three dimension and the Pauli matrices we introduced. We will not need the details or proofs beyond this for cryptographic purposes.

We end this discussion by introducing the Bloch sphere, which gives a nice geometric interpretation (in three dimensions) to the vectors we introduced to describe the Silver atom.

Notation 24 (Bloch Sphere). Consider the following parametrisation: $|\theta, \phi\rangle := \cos\theta/2\,|0\rangle + e^{i\phi}\sin\theta/2\,|1\rangle$ where $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$. This allows one to interpret the vector $|\theta, \phi\rangle$ as a vector on the surface of a three dimensional sphere by taking θ as the polar angle and ϕ as the azimuthal, as depicted in Figure 2.4.

Note 25. Observe that the eigenvectors of σ_z, σ_x, and σ_y on the three dimensional sphere, correspond to vectors along the z, x and y axis respectively, as detailed below

$$
\begin{aligned}
|0\rangle &= |0, \phi\rangle && \text{for any } \phi \\
|1\rangle &= |\pi, \phi\rangle && \text{for any } \phi \\
|+\rangle &= |\pi/2, 0\rangle \\
|-\rangle &= |\pi/2, \pi\rangle \\
|\tilde{+}\rangle &= |\pi/2, \pi/2\rangle \\
|\tilde{+}\rangle &= |\pi/2, 3\pi/2\rangle
\end{aligned}
$$

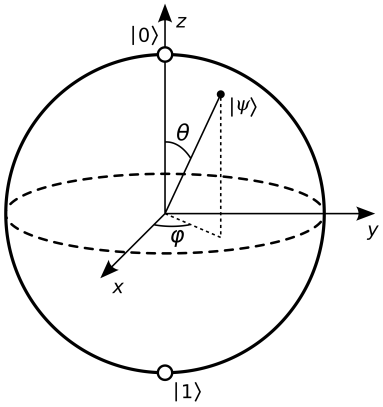

Figure 2.4: Bloch Sphere (taken from Wikipedia)

2.3.1.2 The uncertainty principle

Note that if two observables O_1 and O_2 commute, i.e. $[O_1, O_2] = 0$, then the order in which they are measured, does not matter. This is a direct consequence of the following fact.

Fact 26. *Let O_1, O_2 be Hermitian matrices in $\mathbb{C}^d$. If $[O_1, O_2] = 0$, then there is a basis $\{|v_1\rangle, \ldots |v_d\rangle\}$ in which O_1, O_2 are simultaneously diagonal, i.e. $O_1 = \sum_i \lambda_i^{(1)} |v_i\rangle\langle v_i|$ and $O_2 = \sum_i \lambda_i^{(2)} |v_i\rangle\langle v_i|$.*

Proof sketch. The key observation to understand this fact is the following: Consider an eigenvector $|v\rangle$ of O_1 with eigenvalue λ. Then, $O_2|v\rangle$ is also an eigenvector of O_1 with eigenvalue λ since

$$O_1(O_2|v\rangle) = O_2 O_1 |v\rangle = \lambda(O_2|v\rangle) \tag{2.6}$$

. How does this help? To illustrate the point, assume that O_1 has a two eigenvectors with eigenvalue λ (if it had 1, then there is nothing to do; and the more than two case follows analogously).

Now, given this assumption, Equation (2.6) shows that O_2 leaves O_1's λ-valued eigenspace invariant. Thus, one can restrict O_2 to this eigenspace, i.e. $\mathsf{span}\{|v_1\rangle, O_2|v_1\rangle\}$, and diagonalise O_2. The resulting basis will be an eigenbasis for both O_1 and O_2, restricted to this subspace. The argument can be repeated for all λ in the spectrum of O_1. □

Using Fact 26, one can show that indeed the measurement order does not matter when the observables commute.

Exercise 27. Let O_1, O_2 be as above and let $|\psi\rangle \in \mathbb{C}^d$ denote any quantum state. Consider the following experiments:

1. Measure O_1 first, to obtain o_1 and then measure O_2 to obtain o_2
2. Measure O_2 first, to obtain o_2 and then measure O_1 to obtain o_1

Prove that the probability one obtains $o_1 = a$ and $o_2 = b$ is the same in both cases, i.e.

$$\Pr\left[(o_1, o_2) = (a, b) : \text{Experiment 1}\right] = \Pr[(o_1, o_2) = (a, b) : \text{Experiment 2}]$$

for all $a \in \text{spectrum}(O_1)$ and all $b \in \text{spectrum}(O_2)$.

The discussion around commuting observables is interesting because it shows that in quantum systems, some properties cannot be simultaneously measured with arbitrary precision. For instance, measuring the x component of the spin of an electron, i.e. measuring σ_x, will 'disturb' the spin of the electron along the other components. This is not a limitation of the measurement apparatus but a fundamental limitation imposed by quantum theory and the fact that $[\sigma_x, \sigma_y] \neq 0$ and $[\sigma_x, \sigma_z] \neq 0$. While here we are not dealing with continuous variables, it is worth mentioning that with position x and momentum p of a particle, essentially the same issue arises. The limit on precisely measuring both position and momentum was famously quantified by Heisenberg as follows.

For any observable O, and quantum state $|\psi\rangle$, let $\Delta O := O - \langle O\rangle$ where, recall, $\langle O\rangle = \langle\psi| O |\psi\rangle$. Then, the expectation value of $(\Delta O)^2$ captures the 'dispersion' in the observed value of O from the expected value of O. To see this, note that if $|\psi\rangle$ is an eigenvector of O, then $\left\langle (\Delta O)^2 \right\rangle = 0$.

Theorem 28 (Uncertainty principle). *Let O_1, O_2 be two observables. Then for every state $|\psi\rangle$, the uncertainty as quantified by the dispersion in O_1 and O_2 together, is lower bounded by the commutator of O_1 and O_2, i.e.*

$$\left\langle (\Delta O_1)^2 \right\rangle \left\langle (\Delta O_2)^2 \right\rangle \geq \frac{1}{4} \left|\langle [O_1, O_2]\rangle\right|^2 .$$

We will prove variants of this statement as needed later. For now, we omit the proof and note that the key ingredient in the proof is the Cauchy-Schwarz inequality.

Lemma 29 (Cauchy-Schwarz inequality). *For two vectors $|u\rangle, |v\rangle \in \mathbb{C}^d$, it holds that $\langle u|u\rangle \langle v|v\rangle \geq |\langle u|v\rangle|^2$.*

2.3.1.3 Applying the axioms to the double slit experiment

I wanted to include an explanation of the double slit experiment using Axiom 19 but in the interest of time, I leave this as an exercise.

Exercise 30. Understand what the Mach-Zhander interferometer is. Can one view it as a simplification of the Double-Slit experiment (see Definition 10 and (13))? Explain the observations made using a Mach-Zhander interferometer starting from Axiom 19.

2.3.1.4 Solving the Schrödinger Equation

So far, we took $H = 0$. We now see what happens when $H \neq 0$. Recall that H only appears in the differential equation, Equation (2.1), Postulate 2. For our purposes, this can be simplified to the point that we do not need to worry about differential equations. To see this, we first recall the following fact.

Fact 31 (Hermitian Exponentiation). *Consider a vector space $\mathbb{C}^d$ and denote by $\mathrm{U}(d)$ the set of all unitary matrices, i.e.*

$$\mathrm{U}(d) := \{U : U^\dagger U = \mathbb{I}\}.$$

Then it holds that

$$\mathrm{U}(d) = \{e^{iH} : H = H^\dagger\}.$$

When H is restricted to also have zero trace, then the group becomes the special unitary group, i.e.

$$\mathrm{SU}(d) = \{e^{iH} : H = H^\dagger, \mathsf{tr}(H) = 0\}$$

where $\mathrm{SU}(d) := \{U : U^\dagger U = 1, \det(U) = 1\}$.

The second relevant observation is the following.

Claim 32 (Solution to Schrödinger's equation for time independent Hamiltonians). Using the notation in Axiom 19, suppose $H(t)$ is independent of time. Then the solution to Equation (2.1) is

$$|\psi(t')\rangle = e^{-i(t'-t)H} |\psi(t)\rangle$$

for all t', t.

Using the claim above and Fact 31, it is clear that by choosing H appropriately, one can apply any unitary U to $|\psi\rangle$.

However, not all unitaries are 'easy' to implement so there is more to be said. We cover this later when we discuss the circuit model in Subsection 2.4.1.

2.3.2 Multiple systems

These axioms were for a single system. Suppose two systems are involved, such as two different Silver atoms. How does quantum mechanics apply to these? To this end, we introduce some notation anticipating its use in computing and cryptographic contexts. We are following the convention in [3].

Notation 33. We use the word *register* to refer to a physical system (such as a Silver atom). We typically use capital letters in sans serif font such as $\mathsf{X}, \mathsf{Y}, \mathsf{Z}$ to refer to registers and calligraphic letters $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ to refer to their corresponding Hilbert spaces.

Definition 34 (Quantum state of multiple registers/systems)**.** Given two registers X and Y, the combined state of the two registers is given by vector in $\mathcal{X} \otimes \mathcal{Y}$ where the vector space $\mathcal{X} \otimes \mathcal{Y}$ is spanned by $\{|u\rangle \otimes |v\rangle : |u\rangle \in \mathcal{X}, |v\rangle \in \mathcal{Y}\}$. We also use $|uv\rangle$ or $|u\rangle |v\rangle$ to write $|u\rangle \otimes |v\rangle$ briefly.

Informally, for those not familiar with the tensor product notation, when we write $|u\rangle \otimes |v\rangle$, we basically need to keep two properties of the tensor product in mind

1. $(|u\rangle \otimes |v\rangle)^\dagger = \langle u| \otimes \langle v|$ and
2. $(\langle u'| \otimes \langle v'|)\,(|u\rangle \otimes |v\rangle) = \langle u'|u\rangle \langle v'|v\rangle$.

With these, it is easy to check the following.

Exercise 35. If $\{|x_i\rangle\}_{i\in I}$ and $\{|y_j\rangle\}_{j\in J}$ are orthonormal bases for vector spaces $\mathcal{X}$ and $\mathcal{Y}$ respectively, then $\{|x_i\rangle \otimes |y_j\rangle\}_{i\in I, j\in J}$ spans the vector space $\mathcal{X} \otimes \mathcal{Y}$.

It may be worth working out the following to ensure everything is clear.

Exercise 36. Explicitly write down the tensor products assuming systems A and B are both two-dimensional:

1. $\mathbb{I}_A \otimes \mathbb{I}_B$
2. $\mathbb{I}_A \otimes (\sigma_x)_B$
3. $(\sigma_x)_A \otimes (\sigma_x)_B$
4. $(\sigma_x)_A \otimes (\sigma_z)_B$

To see why the tensor product appears, perhaps it is worth looking at how similar situations are handled in classical probability theory. Consider a process that involves rolling two die. Now each dice can produce 6 outcomes. However, when two die are concerned, the sample space of outcomes becomes 6×6. Intuitively, quantum theory is generalising this idea to vector spaces, using tensor products.

Now that we understand how to describe the quantum state of multiple registers/systems, it is easy to extend Axiom 19—they remain unchanged, except that the initial state must be taken to be a vector in the tensor product of the Hilbert spaces corresponding to the various systems.

Axiom 37 (Continuation of Axiom 19)**.** *In addition to the three axioms stated in Axiom 19 for a single system, the fourth axiom which accounts for multiple systems, is as follows.*

4. The joint state of multiple quantum systems corresponding to Hilbert spaces $\mathcal{H}_1 \dots \mathcal{H}_k$ is specified by a vector in $\mathcal{H} := \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k$. Furthermore, if the state of the individual systems is known to be $|\psi_1\rangle \in \mathcal{H}_1, \dots, |\psi_k\rangle \in \mathcal{H}_k$, then the joint state of all the systems is given by $|\psi_1\rangle \otimes \cdots \otimes |\psi_k\rangle \in \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k$.

What is more interesting now, is that one can ask the following question: Suppose one has two registers S and E, denoting the system of interest and the environment, in the state $|\psi\rangle_{\mathsf{S}} |\phi\rangle_{\mathsf{E}}$. Suppose S and E are made to interact by applying some Hamiltonian H_{SE} acting on both $\mathcal{S}$ and $\mathcal{E}$ (the corresponding Hilbert spaces of registers S and E). And finally, system E is removed. Given H_{SE} and $|\phi\rangle$, clearly, the system S goes from the state $|\psi\rangle$ to some different $\mathcal{C}(|\psi\rangle)$. Clearly, this is a physical process and yet, one can show that the transformation from $|\psi\rangle$ to $\mathcal{C}(|\psi\rangle)$ is not unitary in general. To see this, we first introduce the notion of mixed and pure states. And then, we state an equivalent version of the axioms of quantum mechanics to allow for such transformations, without having to explicitly invoke the existence of an environment.

2.3.2.1 Pure and Mixed States (tracing out systems)

So far, we have only discussed what are called *pure states*—the state of a system S is a vector $|\psi\rangle \in \mathcal{S}$ in the Hilbert space corresponding to S. However, it is easy to see that the formalism we have introduced does not handle classical probabilities conveniently. Suppose that Aman prepares system S in the state $|\psi_0\rangle$ and $|\psi_1\rangle$ with probability p_0 and p_1 respectively[4] and gives this state to Basanti, without telling her which state he prepared. How should Basanti write the state of system S, from her point of view?

First, note that the axioms of quantum mechanics (as in Axiom 19) allow Basanti to compute what happens in any given experiment. For instance, suppose Basanti evolves the system using the unitary U and then measures the observable O. Then, the expected value of the observable, can be computed by computing what happens in each case—when the state is $|\psi_0\rangle$ and when the state is $|\psi_1\rangle$—and then weighting these outcomes with p_0 and p_1. More precisely, the expected value O will be

$$p_0 \langle\psi_0| U^\dagger O U |\psi_0\rangle + p_1 \langle\psi_1| U^\dagger O U |\psi_1\rangle . \tag{2.7}$$

Such calculations can be done for simple situations but become quite cumbersome in general.

To remedy this, the *density matrix formalism* was introduced, which allows one to handle both pure states and mixed states—classical mixtures of pure states states (like the one discussed above).

Definition 38. Let S be a register and $\mathcal{S} = \mathbb{C}^d$ be its associated vector space. Then the state of S is given by a *density matrix* ρ in $\mathbb{C}^{d\times d}$ which satisfies the following conditions:

1. ρ is Hermitian,
2. $\rho \geq 0$, i.e. all its eigenvalues are non-negative
3. $\mathsf{tr}(\rho) = 1$.

The set of all density matrices is denoted by $\mathsf{D}(\mathcal{S})$.

Let us write down Basanti's state using this density matrix formalism:

$$\rho = p_0 |\psi_0\rangle \langle\psi_0| + p_1 |\psi_1\rangle \langle\psi_1| .$$

Exercise 39 (below) claims that one can write $\langle\psi_0| O |\psi_0\rangle = \mathsf{tr}(O |\psi_0\rangle \langle\psi_0|)$ and therefore the expected value of O can be computed as $\mathsf{tr}(O\rho)$ which matches the expression in Equation (2.7). It is straightforward to observe how Axiom 19 applies to density matrices. Since ρ is a Hermitian matrix, it can always be diagonalised (see Theorem 17). Writing $\rho = \sum_{i=1}^d p_i |\phi_i\rangle \langle\phi_i|$ where $p_i \geq 0$ and $\{|\phi_i\rangle\}_{i=1}^d$ are orthonormal, we note that if the states $|\phi_i\rangle$ are mapped to $|\phi_i'\rangle$ using Axiom 19, then the corresponding density matrix $\rho' = \sum_{i=1}^i p_i |\phi_i'\rangle \langle\phi_i'|$. For instance, if $|\phi'\rangle = |\phi(t)\rangle = e^{-itH} |\phi\rangle = U |\phi\rangle$, ρ maps to $U\rho U^\dagger$.

Exercise 39. Prove that for any state $|\psi\rangle \in \mathbb{C}^d$ and any matrix $O \in \mathbb{C}^{d\times d}$, $\langle\psi| O |\psi\rangle = \mathsf{tr}(O |\psi\rangle \langle\psi|)$.

Before we return to the question of interacting a system with an environment and then removing the environment, we briefly mention a few noteworthy properties of density matrices. First, even though we motivated the density matrix by saying that a classical mixture of states $|\psi_0\rangle$ and $|\psi_1\rangle$ can be viewed as a density matrix, the inverse is not true—given a density matrix, (in general) there is no unique set of pure states and probabilities that constitute the density matrix.

[4] We impose $p_0 < 1$ and $p_1 = 1 - p_0$ to ensure these really are probabilities.

Exercise 40. Let $\rho := \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|$ and $\sigma := \frac{1}{2}|u\rangle\langle u| + \frac{1}{2}|v\rangle\langle v|$ where $|u\rangle := \sqrt{3/4}|0\rangle + \sqrt{1/4}|1\rangle$ and $|v\rangle := \sqrt{3/4}|0\rangle - \sqrt{1/4}|1\rangle$. Is $\rho = \sigma$?

The other property is the following observation which follows immediately from Theorem 17.

Exercise 41. A density matrix ρ has rank 1 if and only if it corresponds to a pure state.

We collect some linear algebra notation, as we will now start relying on matrices (or linear operators) acting on vector spaces more heavily.[5]

Notation 42. Let $\mathcal{X}, \mathcal{Y}$ be two vector spaces. We define the following sets, relative to $\mathcal{X}, \mathcal{Y}$.

1. Linear operators.
   a) $\mathrm{L}(\mathcal{X}, \mathcal{Y})$: The set of liner operators from $\mathcal{X}$ to $\mathcal{Y}$.
   b) $\mathrm{L}(\mathcal{X})$: The set of linear operators from $\mathcal{X}$ to $\mathcal{X}$.
2. Unitary operators. $\mathrm{U}(\mathcal{X}) := \{U \in \mathrm{L}(\mathcal{X}) : U^\dagger U = \mathbb{I}\}$
3. Hermitian operators. $\mathrm{Herm}(\mathcal{X}) := \{M \in \mathrm{L}(\mathcal{X}) : M^\dagger = M\}$
4. Positive semidefinite operators. $\mathrm{Pos}(\mathcal{X}) := \{M \in \mathrm{Herm}(\mathcal{X}) : M \geq 0\}$ where $M \geq 0$ means that all its eigenvalues are non-negative.

2.3.2.2 Reduced states and Purifications

Consider a system register S and an environment register E. Suppose the state of the two registers initially is given by ρ_{SE}. How should one describe the state of register S by itself?

Stated in a more operational sense, suppose Aman prepares/initialises registers SE into some initial state ρ_{SE} and gives register S to Basanti. How should Basanti describe the state of register S?

It turns out that the correct operation that captures this situation, is the partial trace operation defined as follows.

Definition 43 (Partial Trace)**.** Let $M_{\mathsf{SE}} \in \mathrm{L}(\mathcal{S} \otimes \mathcal{E})$ be a matrix. Then,

$$\mathrm{tr}_{\mathsf{E}}(M_{\mathsf{SE}}) := \sum_{i \in I} \langle e_i|_{\mathsf{E}} M_{\mathsf{SE}} |e_i\rangle_{\mathsf{E}}$$

where $\{|e_i\rangle\}_{i \in I}$ is any orthonormal basis for $\mathcal{E}$. Equivalently,

$$\mathrm{tr}_{\mathsf{E}}\left(|s\rangle\langle s'|_{\mathsf{S}} \otimes |e\rangle\langle e'|_{\mathsf{E}}\right) = |s\rangle\langle s'|_{\mathsf{S}} \, \mathrm{tr}(|e\rangle\langle e'|)$$

for any set of vectors $|s\rangle, |s'\rangle \in \mathcal{S}$ and $|e\rangle, |e'\rangle \in \mathcal{E}$.

Notation 44 (Reduced Density matrix). When $\rho_{\mathsf{SE}} \in \mathrm{Pos}(\mathcal{S} \otimes \mathcal{E})$ is a density matrix, $\rho_{\mathsf{S}} := \mathrm{tr}_{\mathsf{E}}(\rho_{\mathsf{SE}})$ is termed the *reduced density matrix*.

Exercise 45. Prove that ρ_{S} is indeed a density matrix, if ρ_{SE} is a density matrix.

Why is partial trace the correct operation here? It is not hard to see that we want the description ρ_{S} to be such that any observable acting non-trivially on register S should produce the same statistic, whether ρ_{S} is used or ρ_{SE} is. Partial trace satisfies this requirement.[6]

Exercise 46. Let AB be registers with $\mathcal{A}$ and $\mathcal{B}$ both $\mathbb{C}^2$. Write down the density matrix ρ_{AB} and the reduced density matrix ρ_{B} corresponding to the following pure states.

1. $|\psi\rangle_{\mathsf{A}} \otimes |\phi\rangle_{\mathsf{B}}$
2. $\sqrt{p_0}|0\rangle_{\mathsf{A}} \otimes |\psi_0\rangle_{\mathsf{B}} + \sqrt{p_1}|1\rangle_{\mathsf{A}} \otimes |\psi_1\rangle_{\mathsf{B}}$ where $p_0 \leq 1$ and $p_1 = 1 - p_0$.
3. $(|00\rangle + |11\rangle)_{\mathsf{AB}}/\sqrt{2}$

[5] We use the terms linear operator and matrices interchangeably as they are equivalent in finite dimensions.
[6] For details, see page Box 2.6 of Nielsen and Chuang (10th edition).

From Exercise 46, it is clear that pure states can give rise to mixed states when a part of the system is traced out. Can one do the opposite? Given a mixed state, can one always 'purify' it, i.e. given ρ_{A}, can one find $|\psi\rangle_{\mathsf{AB}}$ such that $\mathsf{tr}_{\mathsf{B}}(|\psi\rangle\langle\psi|_{\mathsf{AB}}) = \rho_{\mathsf{A}}$?

Exercise 47 (Purification). Let $\rho_{\mathsf{A}} = \sum_{i\in I} \lambda_i |u_i\rangle\langle u_i|_{\mathsf{A}}$ be the spectral decomposition of a density matrix $\rho_{\mathsf{A}} \in \mathcal{A}$. Then, verify that $|\psi\rangle_{\mathsf{AB}} = \sum_{i\in I} \sqrt{\lambda_i} |u_i\rangle_{\mathsf{A}} |i\rangle_{\mathsf{B}}$ satisfies $\mathsf{tr}_{\mathsf{B}}(|\psi\rangle\langle\psi|_{\mathsf{AB}}) = \rho_{\mathsf{A}}$. Is there any other state satisfying this property? Give an example.

As you must have noticed, there is in fact an entire family of purifications for any given density matrix ρ_{A},

$$\{(U_{\mathsf{A}} \otimes \mathbb{I}_{\mathsf{B}}) |\psi\rangle_{\mathsf{AB}}\}$$

where $|\psi\rangle_{\mathsf{AB}}$ is as above. It turns out that this is essentially a complete characterisation of purifications, sometimes also called Uhlman's theorem.

Theorem 48 (Unitary equivalence of purifications.). *Let $\rho_{\mathsf{A}} \in \mathcal{A}$ be a density matrix and suppose $|\psi\rangle_{\mathsf{AB}}, |\phi\rangle_{\mathsf{AB}} \in \mathcal{A}\otimes\mathcal{B}$ are purifications of ρ_{A}, i.e.*

$$\mathsf{tr}_{\mathsf{B}}(|\psi\rangle\langle\psi|) = \rho_{\mathsf{A}} = \mathsf{tr}_{\mathsf{B}}(|\phi\rangle\langle\phi|).$$

Then, there is a unitary $U \in \mathsf{U}(\mathcal{B})$ such that $|\psi\rangle_{\mathsf{AB}} = \mathbb{I}_{\mathsf{A}} \otimes U_{\mathsf{B}} |\phi\rangle$.

Proof sketch. Assume for now (we prove it shortly, in Lemma 51), that every pure bipartite state can be written as

$$|\psi\rangle_{AB} = \sum_{i\in I} a_i |u_i\rangle |v_i\rangle \tag{2.8}$$

where $a_i > 0$, $\{|u_i\rangle\}_i$ and $\{|v_i\rangle\}$ are both orthonormal basis for $\mathcal{A}$ and $\mathcal{B}$ respectively, and $I \subseteq \mathbb{N}$. Similarly, we can write

$$|\phi\rangle_{AB} = \sum_{i\in I'} a'_i |u'_i\rangle |v'_i\rangle \tag{2.9}$$

where $a'_i > 0$, $\{|u'_i\rangle\}$ and $\{|v_i\rangle\}$ are orthonormal basis for $\mathcal{A}$ and $\mathcal{B}$ respectively.

Since $\mathsf{tr}_B(|\psi\rangle\langle\psi|) = \mathsf{tr}_B(|\phi\rangle\langle\phi|)$, from Equation (2.8) and (2.9) it must be the case that $a_i = a'_i$, $|v_i\rangle = |v'_i\rangle$, and $I = I'$ (we make the simplifying assumption that all a_is are distinct). Evidently, $|\psi\rangle_{AB}$ can be mapped to $|\phi\rangle$ by the unitary $\mathbb{I}_A \otimes U_B$ where $U_B := \sum_i |u'_i\rangle\langle u_i|$ maps $|u_i\rangle$ to $|u'_i\rangle$ (see Exercise 16). □

The proof of this fact relied on Equation (2.8) which is known as the Schmidt decomposition. However, before we do that, it is instructive to introduce entanglement.

2.3.2.3 Entanglement

Perhaps one of the most surprising features of quantum mechanics, is *entanglement*. The definition itself is unassuming. We give the simplest definition—for pure bipartite states (i.e. involving two systems). Let $|\psi\rangle_{\mathsf{AB}} \in \mathcal{A}\otimes\mathcal{B}$ be a state on two registers AB. Then, $|\psi\rangle_{\mathsf{AB}}$ is entangled if it cannot be written as $|\phi\rangle_{\mathsf{A}} \otimes |\phi'\rangle_{\mathsf{B}}$ for any choice of $|\phi\rangle_{\mathsf{A}} \in \mathcal{A}$ and $|\phi'\rangle_{\mathsf{B}} \in \mathcal{B}$. Let us work through some examples.

Example 49. Write the following states as tensor products of two states, when possible:

1. $|\psi\rangle = (|0\rangle\otimes|0\rangle + |1\rangle\otimes|0\rangle)/\sqrt{2}$
2. $|\psi'_{\pm}\rangle = (|0\rangle|0\rangle \pm |1\rangle|0\rangle \pm |0\rangle|1\rangle + |1\rangle|1\rangle)/\sqrt{2}$
3. $|\psi_{\pm}\rangle = (|0\rangle\otimes|0\rangle \pm |1\rangle\otimes|1\rangle)/\sqrt{2}$
4. $|\phi_{\pm}\rangle = (|0\rangle\otimes|1\rangle \pm |1\rangle\otimes|0\rangle)/\sqrt{2}$

Also, compute partial traces $\mathsf{tr}_A(\cdot)$ and $\mathsf{tr}_b(\cdot)$ for each. (The states in (3) and (4) are known as *Bell states*.)

At this point, it is unclear why this fact carries any significance. However, it turns out that states that are entangled, are a very powerful resource for various communication related tasks because parties in possession of such states (like the Bell states in Example 49), can produce correlations that are classically impossible. We will look at these in detail when we study self-testing/non-local games in later chapters. Despite being powerful, entanglement is not 'all powerful'—it does not allow the entangled parties to communicate instantaneously even though it does allow one to 'teleport' quantum states (discussed in Subsection 2.4.1). The formal definition of entanglement is as follows.

Definition 50 (Separable and entangled states)**.** Let AB denote two registers and $\mathcal{A}, \mathcal{B}$ their corresponding Hilbert spaces. A density matrix $\rho_{\mathsf{AB}} \in \mathcal{A} \otimes \mathcal{B}$ is *separable* if there exist density matrices $\sigma_{\mathsf{A}}^{(i)} \in \mathcal{A}$ and $\sigma_{\mathsf{B}}^{(i)} \in \mathcal{B}$ indexed by i, such that

$$\rho_{\mathsf{AB}} = \sum_{i \in I} p_i \sigma_{\mathsf{A}}^{(i)} \otimes \sigma_{\mathsf{B}}^{(i)}$$

where $p_i > 0$ and $I \subseteq \mathbb{N}$. If a state is not separable, then ρ_{AB} is *entangled*.

The definition above specifies when a state is entangled but for pure bipartite states, one can also give a nice quantitative characterisation of entanglement—how entangled is a state?—in terms of what is known as the Schmidt rank. Indeed, it is related to the Schmidt decomposition we mentioned above.

Lemma 51 (Schmidt decomposition and rank)**.** *Given pure bipartite state $|\psi\rangle_{\mathsf{AB}} \in \mathcal{A} \otimes \mathcal{B}$, there are orthonormal vectors $\{|u_i\rangle\}_i$ for $\mathcal{A}$ and $\{|v_i\rangle\}_i$ for $\mathcal{B}$ such that $|\psi\rangle_{\mathsf{AB}} = \sum_i c_i |u_i\rangle_{\mathsf{A}} \otimes |v_i\rangle_{\mathsf{B}}$ where $c_i \geq 0$. Given $|\psi\rangle$, the Schmidt rank is the number of positive coefficients c_i in its Schmidt decomposition (this number is independent of which Schmidt decomposition is used to compute it).*

Proof sketch. This admits a surprisingly simple proof. We focus on the case where $\mathcal{A}$ and $\mathcal{B}$ have the same dimension. Recall, that singular value decomposition which states that any matrix M can be expressed as $U\Sigma V$ where U and V are unitary and Σ is a diagonal matrix with non-negative entries c_i. Let us write $|\psi\rangle = \sum_{ij} \alpha_{ij} |i\rangle |j\rangle$ where $\alpha_{ij} \in \mathbb{C}$ and $\{|i\rangle\}_i$ and $\{|j\rangle\}_j$ are orthonormal basis for $\mathcal{A}$ and $\mathcal{B}$. Now treating α_{ij} as entries of the matrix M, we can write

$$\begin{aligned}
|\psi\rangle &= \sum_{ij} \sum_{\ell m} U_{i\ell} \Sigma_{\ell m} V_{mj} |i\rangle |j\rangle \\
&= \sum_{\ell m} \Sigma_{\ell m} \underbrace{\left(\sum_i U_{i\ell} |i\rangle \right)}_{:=|u_\ell\rangle} \underbrace{\left(\sum_j V_{mj} |j\rangle \right)}_{:=|v_m\rangle} \\
&= \sum_{\ell m} \Sigma_{\ell m} |u_\ell\rangle |v_m\rangle \\
&= \sum_{\ell} c_\ell |u_\ell\rangle |v_\ell\rangle
\end{aligned}$$

as asserted. □

While the pure bipartite case admits a clean characterisation, this is rare. Characterisation of multipartite entanglement and that of mixed states are both active research areas. We end this brief discussion with an example of a tripartite entangled state.

Exercise 52. Extend the definition of entanglement to three parties (as in, extend Definition 50). Is the following state (known as the *GHZ* state) entangled according to your definition?

$$|\psi\rangle = \frac{|000\rangle_{\mathsf{ABC}} + |111\rangle_{\mathsf{ABC}}}{\sqrt{2}}$$

What is the reduced state on BC when system A is traced out?

2.3.2.4 POVMs

In quantum mechanics, the most general measurement is given by what are known as Positive Operator Valued Measures.

Definition 53 (POVM)**.** Let S be a register and let $\mathcal{S}$ be the associated vector space. A Positive Operator Valued Measure (POVM) is a set of positive semi-definite operators $\mathcal{M} := \{M_1, \dots M_k\}$ where $M_1, \dots, M_k \in \mathrm{Pos}(\mathcal{S})$ such that $\sum_{i=1}^{k} M_i = \mathbb{I}$ and elements M_i of $\mathcal{M}$ are called POVM elements.

Axiom 54. *Given a POVM $\mathcal{M}$ as stated above, and a quantum state $\rho \in \mathcal{S}$, the probability that outcome i is obtained when register S is measured using $\mathcal{M}$, is given by $\mathsf{tr}(M_i\rho)$.*

One can show (by linearity and probability conservation arguments) that this is the most general measurement possible in quantum mechanics—including the case where a standard measurement is made on a system and an environment register. This characterisation proves to be very helpful in various operational settings where one wants to bound the power of an adversarial party.

Note that Axiom 54 does not say what the post-measurement state is. This may not be unique.

Remark 55 (Kraus Operators and 'measurement operators')*.* One often considers what are called *Kraus operators* associated with $M_i := N_i^\dagger N_i$ to describe the post measurement state. These are not unique but for every POVM element M_i, such a decomposition can be constructed. Suppose the 'measurement operators' $\{N_i^\dagger N_i\}_i$ are measured. Then, the post measurement state, upon seeing outcome i, is given by $N_i\rho N_i^\dagger / \mathsf{tr}(N_i\rho N_i^\dagger)$.

So far, we saw how measurements can be generalised as POVM and 'measurement operators'. How is unitary evolution generalised in the system-environment way of thinking? The most general transformation a quantum state can undergo, is termed a quantum channel.

2.3.2.5 Quantum Channels—Axioms, restated

Let us directly start with the definition of a quantum channel. We discuss its connection to the 'original' postulates later.

Axiom 56. *We have the following.*

State. *Each quantum system has an associated Hilbert space, $\mathcal{H}$ and its state is given by a density matrix, i.e. $\rho \in \mathrm{Pos}(\mathcal{H})$ and $\mathsf{tr}(\rho) = 1$.*

Channels. *Every physical process that acts on the state ρ, can be represented by a sequence of maps $\{\mathcal{E}_i\}$, known as* channels*, where $\mathcal{E}_i : \mathrm{Pos}(\mathcal{H}) \to \mathrm{Pos}(\mathcal{H}_i)$ (here $\mathcal{H}_i$ is also a Hilbert space) and i denotes the outcome index, associated with applying the physical process. These maps, together, satisfy property that $\sum_i \mathsf{tr}\left(\mathcal{E}_i(\sigma)\right) = 1$. $\{\mathcal{E}_i\}_i$. Each map $\mathcal{E} \in \{\mathcal{E}_i\}_i$ satisfies the following:*

1. *$\mathsf{tr}[\mathcal{E}(\rho)]$ is the probability that $\mathcal{E}$ occurs, i.e. $0 \leq \mathsf{tr}[\mathcal{E}(\rho)] \leq 1$ for any state ρ*
2. *$\mathcal{E}$ is a convex-linear map, i.e. for any probabilities $\{p_i\}$, it holds that*

$$\mathcal{E}\left(\sum_i p_i\rho_i\right) = \sum_i p_i\mathcal{E}(\rho_i)$$

3. *Finally, $\mathcal{E}$ is a CPTP map, i.e. $\mathbb{I} \otimes \mathcal{E}(A)$ is positive for any positive operator A on the joint system.*

Multiple systems. *The joint state of multiple quantum systems is given by $\rho \in \mathrm{Pos}(\mathcal{H})$ satisfying $\mathsf{tr}(\rho) = 1$ where $\mathcal{H}$ is a tensor product of the Hilbert space of the constituent systems and if the state of the constituent systems is known to be $\rho_1 \dots \rho_k$ then the state of the joint system is given by $\rho_1 \otimes \dots \otimes \rho_k \in \mathrm{Pos}(\mathcal{H})$.*

While it is outside the scope of these notes to prove the equivalence of Axiom 19 and Axiom 56, the equivalence can nonetheless be made plausible. To this end, we first note that a channel admits what is sometimes known as the operator sum representation.

Theorem 57. *The map $\mathcal{E}$ satisfies Axiom 56 if and only if*

$$\mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger$$

for some operators $\{E_i\}$ satisfying $\sum_i E_i^\dagger E_i \leq \mathbb{I}$.

First, let us quickly check that all the operations we have discussed so far, can in fact, be expressed in this formalism.

- Unitary evolution. Suppose the system evolves under the unitary U. The corresponding sequence of channels, is a single channel $\mathcal{E}$ (i.e. $k = 1$). It acts as $U\rho U^\dagger$. Clearly, this is a channel because $U^\dagger U = \mathbb{I}$.
- Observables. Suppose observable $O = \sum_i o_i \Pi_i$ is being measured. Then the sequence of channels is $\{\mathcal{E}_i\}_i$ where $\mathcal{E}_i$ acts as $\Pi_i \rho \Pi_i$. Again, this is clearly a channel because $\Pi_i \leq \mathbb{I}$ (recall Exercise 16 says Π_i has eigenvalues 0 or 1).
- Measurement operators. Suppose the measurement operators are given by $\{N_i^\dagger N_i\}_i$. Then the sequence of channels is $\{\mathcal{E}_i\}_i$ where $\mathcal{E}_i$ acts as $N_i \rho N_i^\dagger$. This too is clearly a channel because $N_i^\dagger N_i \geq 0$ and $\sum_i N_i^\dagger N_i = \mathbb{I}$ which means $N_i^\dagger N_i \leq \mathbb{I}$.

Now, suppose the initial state of our system was ρ and we measured $\{N_i^\dagger N_i\}_i$ (as stated in Remark 55) and obtained outcome i. The state of the system would then become proportional to $N_i \rho N_i^\dagger$ (where the proportionality constant is $1/\mathsf{tr}(N_i \rho N_i^\dagger)$) and the probability of obtaining outcome i, recall, was $\mathsf{tr}(N_i \rho N_i^\dagger)$. Suppose the measurement device was faulty and it is unable to distinguish between outcomes i and j. Then, the state of the system would become proportional to

$$\mathsf{tr}(N_i \rho N_i^\dagger) \cdot \frac{N_i \rho N_i^\dagger}{\mathsf{tr}(N_i \rho N_i^\dagger)} + \mathsf{tr}(N_j \rho N_j^\dagger) \cdot \frac{N_j \rho N_j^\dagger}{\mathsf{tr}(N_j \rho N_j^\dagger)} = N_j \rho N_j^\dagger + N_j \rho N_j^\dagger.$$

This makes it plausible that the most general quantum operation can be understood in terms of a coarse-grained measurement process. One can prove that starting with a quantum system, interacting with an environment and making measurements (as in Axiom 19, Axiom 37 and Remark 55), the most general operation possible, can be expressed as a quantum channel as stated in Axiom 56.

We end this section by stating another very useful characterisation of quantum channels—which is a consequence of the so-called Steinspring dialation theorem.

Theorem 58 (Steinspring Representation of Channels)**.** *For any channel $\mathcal{E}$ from $\mathcal{X}$ to $\mathcal{Y}$ there is always (i) a Hilbert space $\mathcal{Z}$ (with dimension at most product of $\mathcal{X}$ and $\mathcal{Y}$) and (ii) an isometry A (i.e. $A^\dagger A = \mathbb{I}$) from $\mathcal{X}$ to $\mathcal{Y} \otimes \mathcal{Z}$ such that*

$\mathcal{E}(X) = \mathsf{tr}_{\mathcal{Z}}(AXA^\dagger)$ for every X.

The following statement, together with the theorem above, shows how one can 'purify' a channel.

Claim 59. One can express an isometry $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ as $A = U(\mathbb{I}_{\mathcal{X}} \otimes |0\rangle_{\mathcal{X}'})$ where $U \in \mathrm{U}(\mathcal{Y} \otimes \mathcal{Z})$ and $\mathcal{X}' = \mathcal{Y} \otimes \mathcal{Z} \backslash \mathcal{X}$. To be concrete, suppose $\mathcal{X}$ is an n qubit register, $\mathcal{Y}$ is an m qubit register and $\mathcal{Z}$ is an $n + m$ qubit register. Thus, U acts on an $n + 2m$ qubit register.

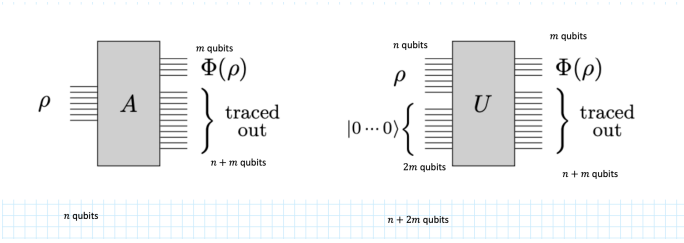

Figure 2.5: Purifying a quantum channel Φ. Left: uses Theorem 58 to express Φ as an isometry with some qubits traced out. Right: expresses the channel as a unitary acting on auxiliary qubits initialised to zero and tracing out some of the resulting qubits (using Claim 59). Figure stolen and slightly modified from [3].

«In progress beyond this point»

Remark 60. TODO: One can view a sequence of channels in Axiom ... as a single trace preserving channel using the following.

§ 2.4 Quantum information/computation

Henceforth, we almost exclusively work with qubits. We also assume that all channels are trace preserving.

Definition 61. A *qubit* is a two dimensional register (as in Notation 33), i.e. a quantum system whose state is described by a two dimensional Hilbert space. The standard basis of this qubit, $|0\rangle, |1\rangle$, will also be referred to as the computational basis.

Registers comprising qubits will typically be written as $\mathsf{X}, \mathsf{Y}, \mathsf{Z}$ and the binary representation of integers x, y, z will be used to describe the computational basis of the underlying qubits, as detailed below.

Notation 62. For an n qubit register X, and an integer $x \in \{0, \ldots 2^n - 1\}$ we use $|x\rangle \in \mathcal{X}$ to denote the state $|x_1\rangle \otimes \ldots |x_n\rangle$ where x_i denotes the ith bit of x in binary representation. E.g. $x = 2, n = 2$, $|2\rangle \in \mathcal{X}$ is used to denote the state $|1\rangle \otimes |1\rangle \in \mathcal{X}$.

2.4.1 The Circuit Model and Teleportation**

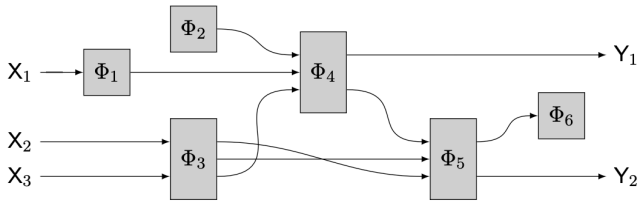

Figure 2.2: An example of a quantum circuit. The input qubits are labelled $\mathsf{X}_1, \mathsf{X}_2, \mathsf{X}_3$, the output qubits are labelled Y_1 and Y_2, and the gates are labelled by (hypothetical) quantum channels $\Phi_1, \ldots, \Phi_6$.

Figure 2.6: Example of quantum circuit. The input qubits are labelled X_1, X_2, X_3 the output qubits are Y_1, Y_2 and gates are given by the channels Φ_i. Taken from [3].

Definition 63 (Quantum Circuit)**.** Let GateSet be a set of quantum channels. A *quantum circuit* relative to a GateSet is an acyclic network of *quantum gates* connected by *wires* where a quantum gate is a quantum channel in GateSet acting on the qubits which in turn are represented by wires.

Here, we take the GateSet to be the following.

1. Toffoli gate. The channel corresponding to the unitary $T : |a\rangle |b\rangle |c\rangle \mapsto |a\rangle |b\rangle |c \oplus ab\rangle$.
2. Hadamard gate. The channel corresponding to the unitary $H : |a\rangle \mapsto \frac{1}{\sqrt{2}} |0\rangle + \frac{(-1)^a}{\sqrt{2}} |1\rangle$.
3. Phase-shift gate. The channel corresponding to the unitary $P : |a\rangle \mapsto i^a |a\rangle$
4. Auxiliary gate. The non-unitary gate that takes no input but produces a single output qubit in state $|0\rangle$.
5. Erasure gate. The non-unitary gate that takes a qubit as input and produces no output. The effect is represented by partial trace.

   The *size* of a quantum circuit Q is denoted by $\mathrm{size}(Q)$ and defined to be the number of quantum gates comprising it.

One can take other gate sets as well but we choose this for concreteness (it is also quite standard). Our choice ofGateSet is universal in the following strong sense.

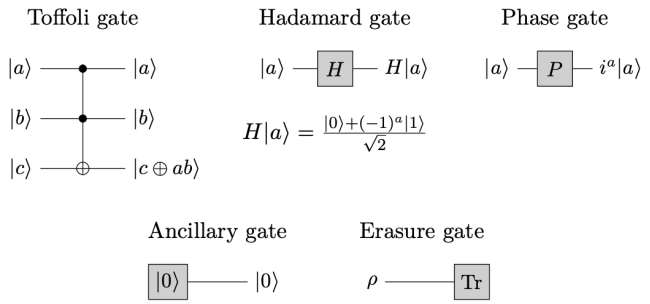

Figure 2.3: Commonly used notation for denoting gates from the universal gate set.

Figure 2.7: Commonly used notation for denoting gates from GateSet. From [3].

Theorem 64. *Let $\mathcal{E}$ be any quantum channel from n qubits to m qubits. Then for every $\epsilon > 0$ there is a quantum circuit Q with n input qubits and m output qubits such that Q ϵ-approximates $\mathcal{E}$. Furthermore, for a fixed choice of n and m, $\text{size}(Q) \leq \text{poly}(\log(1/\epsilon))$ for some fixed polynomial* poly.

We formalise what we mean by ϵ-approximates in Subsection 2.4.2 below.

Circuit, as we have described them, can be purified in the following sense.

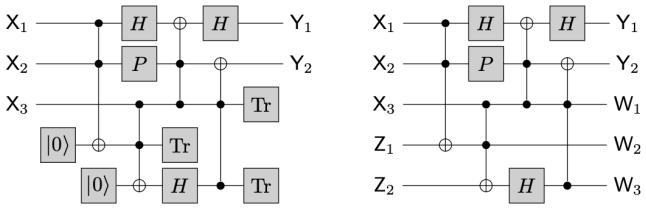

Figure 2.4: A general quantum circuit (left) and its unitary purification (right).

Figure 2.8:

Example 65 (Teleportation circuit.)**.** Temp

$$|\psi\rangle_C = \alpha|0\rangle_C + \beta|1\rangle_C.$$

ne subscript C above is used onl

$$|\Phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B),$$

$$|\psi\rangle_C \otimes |\Phi^+\rangle_{AB} = (\alpha|0\rangle_C + \beta|1\rangle_C) \otimes \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B).$$

Alice will then make a local measurement in the Bell basis (i.e. the four Bell states) on the two particles in her possession. To make the result of her measurement clear, it is best to write the state of Alice's two qubits as superpositions of the Bell basis. This is done by using the following general identities, which are easily verified:

$$|0\rangle \otimes |0\rangle = \frac{1}{\sqrt{2}}(|\Phi^+\rangle + |\Phi^-\rangle),$$

$$|0\rangle \otimes |1\rangle = \frac{1}{\sqrt{2}}(|\Psi^+\rangle + |\Psi^-\rangle),$$

$$|1\rangle \otimes |0\rangle = \frac{1}{\sqrt{2}}(|\Psi^+\rangle - |\Psi^-\rangle),$$

and

$$|1\rangle \otimes |1\rangle = \frac{1}{\sqrt{2}}(|\Phi^+\rangle - |\Phi^-\rangle).$$

$$\begin{aligned} &|\psi\rangle_C \otimes |\Phi^+\rangle_{AB} = \\ &\frac{1}{2}\Big[\, |\Phi^+\rangle_{CA} \otimes (\alpha|0\rangle_B + \beta|1\rangle_B) \;+\; |\Phi^-\rangle_{CA} \otimes (\alpha|0\rangle_B - \beta|1\rangle_B) \\ &+\; |\Psi^+\rangle_{CA} \otimes (\alpha|1\rangle_B + \beta|0\rangle_B) \;+\; |\Psi^-\rangle_{CA} \otimes (\alpha|1\rangle_B - \beta|0\rangle_B)\Big]. \end{aligned}$$

Definition 66. BQP computations.

2.4.2 Distances between states and channels**

Norms

Definition 67. Consider any operator $O \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$.

- These operators already have the *operator norm* induced by the Euclidean norm on $\mathcal{X}$ and $\mathcal{Y}$ as

$$\|O\| = \max\{\|O|\psi\rangle\| : |\psi\rangle \in \mathcal{X}, \||\psi\rangle\| \leq 1\}.$$

- We also use the *trace norm*, defined as $\|O\|_1 = \mathsf{tr}(\sqrt{OO^\dagger})$.

The following result from linear algebra will be helpful.

Fact 68. *The operator norm and trace norm are dual to each other, i.e. the following holds:*

$$\begin{aligned} \|O\| &= \max\left\{|\langle O', O\rangle| : O' \in \mathrm{L}(\mathcal{X}, \mathcal{Y}), \|O'\|_1 \leq 1\right\}, \\ \|O'\|_1 &= \max\left\{|\langle O', O\rangle| : O \in \mathrm{L}(\mathcal{X}, \mathcal{Y}), \|O\| \leq 1\right\}. \end{aligned}$$

States

Theorem 69 (Operational meaning of trace distance between states)**.** *It holds that*

$$\max\left\{\frac{1}{2}\langle P_0, \rho_0\rangle + \frac{1}{2}\langle P_1, \rho_1\rangle : P_0, P_1 \geq 0, P_0 + P_1 = \mathbb{I}_{\mathcal{X}}\right\} = \frac{1}{2} + \frac{1}{4}\|\rho_0 - \rho_1\|_1.$$

Definition 70. σ is an ϵ-approximation to ρ if $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon$.

Definition 71. Fidelity between ρ and σ is given by

$$\mathrm{F}(\rho,\sigma) = \mathsf{tr}(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}) = \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1 .$$

NB. the second expression makes it clear that $F(\rho,\sigma) = F(\sigma,\rho)$ because $\sqrt{\rho}\sqrt{\sigma}$ and $\sqrt{\sigma}\sqrt{\rho} = (\sqrt{\rho}\sqrt{\sigma})^\dagger$ must have the same singular values.

NB2. When $\sigma = |\psi\rangle\langle\psi|$, the expression simplifies to

$$F(\rho, |\psi\rangle\langle\psi|) = \sqrt{\langle\psi|\rho|\psi\rangle}.$$

Theorem 72 (Uhlmann's theorem, characterising fidelity)**.** *It holds that*

$$\mathrm{F}(\rho,\sigma) = \max_{|\psi\rangle,|\phi\rangle} |\langle\psi|\phi\rangle|$$

where $|\psi\rangle, |\phi\rangle$ are purifications of ρ, σ respectively.

Proposition 73 (Fuchs-van de Graaf inequalities)**.** *Fidelity and trace distance are related as follows:*

$$1 - F(\rho,\sigma) \leq \frac{1}{2}\|\rho-\sigma\|_1 \leq \sqrt{1-F(\rho,\sigma)^2}$$

for all ρ, σ density matrices.

Channels

Notation 74. We use the following. The set of trace preserving channels from $\mathcal{X}$ to $\mathcal{Y}$ are denoted by $C(\mathcal{X},\mathcal{Y})$. The set of linear maps that are neither CPTP nor trace preserving are denoted by $T(\mathcal{X},\mathcal{Y})$.

Diamond norm. Given an operator $\Delta \in T(\mathcal{X},\mathcal{Y})$, the diamond norm is given by

$$\|\Delta\|_\diamond = \max\{\|(\Delta\otimes\mathbb{I}_{\mathcal{W}}(X)\|_1 : X \in \mathrm{L}(\mathcal{X}\otimes\mathcal{W}), \|X\|_1 \leq 1\}$$

where $\mathcal{W}$ is any Hilbert space having dimension at least as large as $\mathcal{X}$ (dimension beyond this does not change the value of the norm).

Proposition 75. *No physical process can distinguish two channels $\mathcal{E}_0, \mathcal{E}_1 \in C(\mathcal{X},\mathcal{Y})$ with bias greater than $\frac{1}{2}\|\mathcal{E}_0 - \mathcal{E}_1\|_\diamond$ when one of them is given with equal probability.*

Definition 76. Channel $\mathcal{E}_1$ is an ϵ-approximation to channel $\mathcal{E}_0$ if $\|\mathcal{E}_0 - \mathcal{E}_1\|_\diamond \leq \epsilon$.

2.4.3 Simon's Problem | Quantum advantage in Query Complexity**

Simon's problem.

The input to an instance of Simon's problem is a function $f : \{0,1\}^n \to \{0,1\}^n$ that has the property that f is 2-to-1 (every value in the range has exactly two preimages) and moreover there is a string $s \in \{0,1\}^n$ such that for every $x, y \in \{0,1\}^n$, $f(x) = f(y)$ if and only if $y = x$ or $y = x + s$, where addition is performed coordinate-wise and modulo 2. The goal is to recover the string s. It is not hard to see that in the worst case any classical algorithm requires at least $\Omega(2^{n/2})$ evaluations of f to determine s. This is because on the one hand for any deterministic algorithm that makes a smaller number of evaluations there is a function f such that all values returned by f are distinct, so no information about s is gained; similarly one can show that for any randomized algorithm if f is chosen at random then it is unlikely that the algorithm will gain any information about s in $\ll 2^{n/2}$ evaluations. On the other hand, by making roughly $\Omega(2^{n/2})$ evaluations at random points then by the birthday paradox one will likely obtain $x \neq y$ such that $f(x) = f(y)$, which immediately reveals $s = x + y$.

Simon showed that there is a quantum algorithm that can solve this problem using only $O(n)$ evaluations, provided that the function f can be evaluated "in superposition". The algorithm first evaluates f on a uniform superposition of inputs, as follows:

$$\begin{aligned} |0^n\rangle|0^n\rangle &\mapsto \frac{1}{\sqrt{2^n}} \sum_x |x\rangle|0^n\rangle \\ &\mapsto \frac{1}{\sqrt{2^n}} \sum_x |x\rangle|f(x)\rangle \ . \end{aligned}$$

It then measures the last register in the computational basis, yielding some $y = f(x_0) = f(x_1)$ where x_0 and $x_1 = x_0 + s$ are the two preimages of y under f. The re-normalized post-measurement state is

$$\frac{1}{\sqrt{2}}(|x_0\rangle + |x_1\rangle)|y\rangle \ . \tag{4.1}$$

Measuring the first register in the Hadamard basis yields a uniformly random $d \in \{0,1\}^n$ such that $d \cdot s = 0$. Repeating the entire procedure $O(n)$ times yields $(n-1)$ linearly independent such d's, which suffices to recover s with high probability.

According to the approximation, the same approach can be applied to any number of "people" and "days". If rather than 365 days there are d, if there are n persons, and if $n \ll d$, then using the same approach as above we achieve the result that if $p(n, d)$ is the probability that at least two out of n people share the same birthday from a set of d available days, then:

$$\begin{aligned} p(n,d) &\approx 1 - e^{-\frac{n(n-1)}{2d}} \\ &\approx 1 - e^{-\frac{n^2}{2d}} . \end{aligned}$$

§ References

[1] Scott Aaronson. Quantum copy-protection and quantum money. In *24th CCC*, pages 229–242, 2009.

[2] S. Carroll. *Something Deeply Hidden: Quantum Worlds and the Emergence of Spacetime.* Penguin Publishing Group, 2020.

[3] Thomas Vidick and John Watrous. Quantum proofs. *Foundations and Trends® in Theoretical Computer Science*, 11(1?2):1?215, 2016.