# Mathematician's introduction to quantum error correction

Will Troiani

August 2020

## Contents



# 1 Quantum computing

The goal of this document is to define a mathematical theory of *qubits* as well as their *measurements*, *time evolutions*, etc, and then to develop a theory of error correction upon this foundation. An optional next step after understanding this mathematics is to find a physical phenomena adhering to

these conditions. This step which we refer to as optional though is not part of this document.

Our standard of *information* will be sequences of binary integers. The form this information can be encoded into during transmission however will be more liberal. A bit of *quantum information*, that is, a *qubit*, will be any norm 1 element of the complex Hilbert space $\mathcal{H} := \mathbb{C}^2$. Actually, we identify elements of $\mathcal{H}$ with linear operators from $\mathbb{C}$ into $\mathcal{H}$ and use Dirac notation. For example, $|0\rangle : \mathbb{C} \longrightarrow \mathcal{H}$ denotes the map defined by linearity and the rule $1 \longmapsto (1, 0)$, whereas $|1\rangle$ denotes the map defined by linearity and the rule $1 \longmapsto (0, 1)$.

**Definition 1.0.1.** A **qubit** is a copy of the $\mathbb{C}$-Hilbert space $\mathbb{C}^2$.

The **state** of a qubit $\mathbb{C}^2$ is a vector $|\psi\rangle \in \mathbb{C}^2$ of norm 1.

A pair $(\mathbb{C}^2, |\psi\rangle)$ consisting of a qubit $\mathbb{C}^2$ and a state $|\psi\rangle \in \mathbb{C}^2$ is a **prepared qubit** and we say $\mathbb{C}^2$ has been **prepared** to $|\psi\rangle$.

If clarity is needed, we will refer to a binary integer as a **classical bit**. This is to help distinguish our standard of information from qubits just defined. If we write a state $|\psi\rangle$ of a qubit $\mathcal{H}$ as a linear combination of the standard basis vectors

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \tag{1}$$

We think of $|\alpha|^2$ and $|\beta|^2$ respectively as *probabilities* of the state $|\psi\rangle$ being in state $|0\rangle$ or state $|1\rangle$ respectively. A qubit where $\alpha \neq 0$ and $\beta \neq 0$ is a **superposition state**.

We can now be more precise; our goal is to develop a theory of communication of classical bits *via* qubits. The manipulation and transmission of classical bits via qubits will loosely be referred to as *quantum computing*. For any physical computation to take place, it is crucial that reliable error correction is possible, simply due to the huge number of components and interactions involved. The following is our central question.

**Question 1.0.2.** What tolerance for error does quantum computing allow for?

This is answered formally by Theorem 4.0.9, which classifies necessary and sufficient conditions for a collection of errors to be correctable.

So far, however, we have only considered systems consisting of a single qubit. Since a system consisting of multiple qubits, any of which may be

in superposition, may *itself* be in a superposition state, the definition of a *composite* system of qubits is not as simple as the *product* of qubits, in the category of Hilbert spaces.

More precisely, say we had two qubits prepared respectively to states $|\psi\rangle, |\varphi\rangle$. Then the pair $(|\psi\rangle, |\varphi\rangle)$ may also be in superposition. That is, the state

$$\alpha(|\psi\rangle, |\varphi\rangle) + \beta(|\psi\rangle, |\varphi\rangle) \tag{2}$$

where $|\alpha|^2 + |\beta|^2 = 1$ is a valid state of the combined system consisting of the two qubits. Thus, states of *pairs* of systems are vectors in a subspace of the Hilbert space freely generated by the pairs $(|\psi\rangle, |\varphi\rangle)$ of states of the qubits. in fact, we will take the composite system to be the tensor product of the two spaces, which means we need to justfiy the bilinearity conditions too. At the time of writing, the author does not know a satisfying way to motivate these conditions mathematically (although $L(X \times Y) \cong L(X) \otimes L(Y)$ is surely relevant).

A qubit as well as any composite of a finite collection of qubits are examples of finite dimensional complex Hilbert spaces. We define a **state space** to be any finite dimensional Hilbert space $\mathcal{H}$.

**Definition 1.0.3.** Let $\mathcal{H}_1, \mathcal{H}_2$ be two state spaces. The **composite state space** is $\mathcal{H}_1 \otimes \mathcal{H}_2$. A **state** of a composite system is a vector $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ which can be written as a linear combination of pure tensors

$$\alpha_1 |\psi_1\rangle + \ldots + \alpha_n |\psi_n\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2 \tag{3}$$

where the coefficients satisfy $|\alpha_1|^2 + \ldots + |\alpha_n|^2 = 1$. The condition that each $|\psi_i\rangle$ is a pure tensor means

$$\forall i = 1, \ldots, n, \exists\, |\psi_i^1\rangle \in \mathcal{H}_1, \exists\, |\psi_i^2\rangle \in \mathcal{H}_2, |\psi_i\rangle = |\psi_i^1\rangle \otimes |\psi_i^2\rangle \tag{4}$$

**Remark 1.0.4.** The tensor product is *not* a product in the category of Hilbert spaces. This is because the states such as (3) are *not* necessarily determined by a choice of state in $\mathbb{H}_1$, and a choice of state in $\mathbb{H}_2$. Thus, it is not a surprise that we observe "bizarre" behavior, as our definition of a coupled system is *not* given by the standard mathematical definition of *product*. A comparison between the monoidal structure of the category of Hilbert spaces and a hypothetical product can be found in [3]).

What degree of access do we have to superposition states? The answer, naturally, is we have access to what we can measure. Rather than one

particular outcome, we define a measurement as a family of possible outcomes with assocaited probabilities; the states of state spaces are probabilistic, and so the measurements will be too. Moreover, we do *not* assume that measurement leaves the state uneffected, and so measurements are operators upon the state space.

**Definition 1.0.5.** A **measurement** on a state space $\mathcal{H}$ is a finite family of linear operators $\{M_m : \mathcal{H} \longrightarrow \mathcal{H}\}_{m\in\mathcal{M}}$ satisfying the **completeness condition**.

$$\sum_{m\in\mathcal{M}} M_m^\dagger M_m = I \tag{5}$$

An element $m \in \mathcal{M}$ is an **outcome** (simply a set of labels).

The **resulting state** after measurement $\{M_m\}_{m\in\mathcal{M}}$ and outcome $m$ is:

$$\frac{M_m |\psi\rangle}{\sqrt{p(m)}} \tag{6}$$

**Remark 1.0.6.** Associated to every measurement and state vector $|\psi\rangle$ there is a value

$$p(m) := \langle\psi| M_m^\dagger M_m |\psi\rangle = \|M_m |\psi\rangle\|^2 \tag{7}$$

It follows from (5) that $p(m) \leq 1$ for all $m, |\psi\rangle$. We understand $p(m)$ as the probability of outcome $m$ on the measurement $\{M_m\}_{m\in\mathcal{M}}$. Under this interpretation, we think of (5) as requiring that the probabilities $p(m)$ sum to 1.

The possibility of superposition states is a liberation, and measurements needing to satisfy the completeness condition is a limitation. For instance, where a classical bit is in one of two states (0 or 1) a qubit has an *infinite* number of possible states (a liberation). On the other hand, only *orthogonal* states can be distinguished, due to the completeness condition (Lemma 1.0.7 below) (a limitation).

**Lemma 1.0.7.** *Let $|\psi_1\rangle, |\psi_2\rangle$ be non-orthogonal states of a qubit $\mathcal{H}$. There is no measurement $\{M_m : \mathcal{H} \longrightarrow \mathcal{H}\}_{m\in\mathcal{M}}$ with $1, 2 \in \mathcal{M}$ satisfying:*

$$p(1) = \langle\psi_1| M_1^\dagger M_1 |\psi_1\rangle = 1 \quad \textit{and} \quad p(2) = \langle\psi_2| M_2^\dagger M_2 |\psi_2\rangle = 1 \tag{8}$$

*Proof.* Assume such a measurement exists. Since $|\psi_1\rangle, |\psi_2\rangle$ are non-orthogonal, there exists $|\varphi\rangle$, orthorgonal to $|\psi_1\rangle$ such that

$$|\psi_2\rangle = \alpha |\psi_1\rangle + \beta |\varphi\rangle \tag{9}$$

for some $\alpha, \beta \in \mathbb{C}$ satisfying $|\alpha|^2 + |\beta|^2 = 1$. Moreover, since $|\psi_1\rangle, |\psi_2\rangle$ are non-orthogonal, we have $\beta \neq 1$. We have

$$\begin{aligned} 1 &= \langle\psi_2|\,|\psi_2\rangle && (10)\\ &= \langle\psi_2|\,M_2^\dagger M_2\,|\psi_2\rangle && (11)\\ &= (\bar{\alpha}\,\langle\psi_1| + \bar{\beta}\,\langle\varphi|)M_2^\dagger M_2(\alpha\,|\psi_1\rangle + \beta\,|\varphi\rangle) && (12)\\ &= |\alpha|^2\,\langle\psi_1|\,M_2^\dagger M_2\,|\psi_1\rangle + \bar{\alpha}\beta\,\langle\psi_1|\,M_2^\dagger M_2\,|\varphi\rangle && (13)\\ &\quad + \bar{\beta}\alpha\,\langle\varphi|\,M_2^\dagger M_2\,|\psi_1\rangle + |\beta|^2\,\langle\varphi|\,M_2^\dagger M_2\,|\varphi\rangle && (14) \end{aligned}$$

Now, by the completeness condition, we have

$$\langle\psi_1| \sum_{m\in\mathcal{M}} M_m^\dagger M_m\,|\psi_1\rangle = \langle\psi_1|\,I\,|\psi_1\rangle = 1 \tag{15}$$

This, combined with $p(1) = 1$ implies $\langle\psi_1|\,M_2^\dagger M_2\,|\psi_1\rangle = 0$. In other words, $\|M_2\,|\psi_1\rangle\|^2 = 0$. This implies $M_2\,|\psi_1\rangle = 0$ (the zero vector). Thus, (13) implies:

$$1 = |\beta|^2\,\langle\varphi|\,M_2^\dagger M_2\,|\varphi\rangle = |\beta|^2\|M_2\,|\varphi\rangle\|^2 \leq |\beta|^2 < 1 \tag{16}$$

This stands in contradiction to (8). □

The liberty of superposition means the following are four valid states of a single qubit.

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \quad \frac{|0\rangle - |1\rangle}{\sqrt{2}} \quad \frac{-|0\rangle + |1\rangle}{\sqrt{2}} \quad \frac{-|0\rangle - |1\rangle}{\sqrt{2}} \tag{17}$$

Since these are non-orthogonal though, Lemma 1.0.7 renders these indistinguishable. Desigining algorithms amongst the push and pull of superposition and measurement is the art of quantum computing. An example of such an art piece is the possibility of transferring two classical bits of information via a single qubit. This possibility is surprising given the previous observation.

**Example 1.0.8.** The following is a state of the composite system $\mathcal{H} := \mathbb{C}^2 \otimes \mathbb{C}^2$.

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}} \in \mathcal{H} \tag{18}$$

Consider the following matrices, we note that these matrices will play an important role later too.

$$I := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad Y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{19}$$

To make the action of $Y$ simpler we will consider $iY$. These act on basis elements as follows.

$$\begin{aligned} I|0\rangle &= |0\rangle & I|1\rangle &= |1\rangle \\ X|0\rangle &= |1\rangle & X|1\rangle &= |0\rangle \\ iY|0\rangle &= -|1\rangle & iY|1\rangle &= |0\rangle \\ Z|0\rangle &= |0\rangle & Z|1\rangle &= -|1\rangle \end{aligned}$$

Applying a choice of these unitary matrices to the first qubit results in the following states of the combined system.

$$\begin{aligned} I: &\quad \frac{|00\rangle + |11\rangle}{\sqrt{2}} \\ X: &\quad \frac{|10\rangle + |01\rangle}{\sqrt{2}} \\ iY: &\quad \frac{|01\rangle - |10\rangle}{\sqrt{2}} \\ Z: &\quad \frac{|00\rangle - |11\rangle}{\sqrt{2}} \end{aligned} \tag{20}$$

A priori, we may have agreed on a correspondence between the operators $I, X, iY, Z$ and respectively the classical bits $00, 10, 01, 11$. Moreover, the states (20) are orthogonal, so Lemma 1.0.7 does not rule out the possibility of distinguishing these states.

Orthogonal states can be distinguished via measurement, this is the content of Lemma 1.0.9 below. Let us emphasise the crucial point of this example: although state (18) is in superposition, the system itself is still thought of as a pair of qubits, one in one hand, one in the other. Thus, if Alice holds the first qubit, and performs one of the transformations (19) then 2 bits of information can be transferred by sending this single qubit to Bob, who holds the second qubit. Bob performs a measurement to distinguish which of the four states (20) the combined system is in, and then extracts the classical bits from the result.

**Lemma 1.0.9.** *If $|\psi_1\rangle, \ldots, |\psi_n\rangle$ are orthogonal states of a state space, then there exists a measurement $\{M_m\}_{m\in\mathcal{M}}$ such that for all $i = 1, \ldots, n$:*

$$p(i) = \langle\psi_i| M_i^\dagger M_i |\psi_i\rangle = 1 \tag{21}$$

*Proof.* The operator

$$E := \sum_{i=1}^{n} |\psi_i\rangle \langle\psi_i| \tag{22}$$

is equal to the identity when restricted to the subspace $\mathrm{Span}\{|\psi_1\rangle, \ldots, |\psi_n\rangle\}$ and when written with respect to the basis $|1\rangle, \ldots, |n\rangle$ of this subspace. A trick to extend this to a measurement of the whole space, and written with respect to the standard basis is to add an operator $M_0$ defined by $I - E$. The set $\{|\psi_i\rangle \langle\psi_i|\}_{i=1,\ldots,n} \cup M_0$ is a measurement distinguishing $|1\rangle, \ldots, |n\rangle$. □

Example 1.0.8 assumed that Alice was able to perform one of the unitary operators (19) to her qubit. For single qubits, *all* the unitary operators constitute the operations we can perform to qubits.

**Definition 1.0.10.** Let $\mathcal{H}$ be a state space. A **single step time evolution** of $\mathcal{H}$ is a unitary operator $U$ on $\mathcal{H}$. A **single step time evolution** of a state vector $|\psi\rangle$ with respect to $U$ is the pair $(|\psi\rangle, U|\psi\rangle)$.

An **evolution** of $\mathcal{H}$ is a sequence of unitary operators $(U_1, ..., U_n)$ on $\mathcal{H}$, an **evolution** of a state vector $|\psi\rangle$ with respect to the evolution $(U_1, ..., U_n)$ is the sequence $(|\psi\rangle, U_1|\psi\rangle, ..., U_n \ldots U_1|\psi\rangle)$.

Example 1.0.8 falls into the special setting where the measurement used to distinguish the states (20) is **projective**, ie, all the measurement operators are projectors.

**Definition 1.0.11.** A linear transformation $P$ is a **projector** if $P^2 = P$.

These simple measurements are sufficient in many situations.

The exact relationship between projective measurements and measurements is given by Proposition 1.0.13 below which says in a precise sense that general measurements are projective measurements augmented by a unitary operator.

**Lemma 1.0.12.** *Let $W \subseteq V$ be a subspace of a Hilbert space $V$, and let $U : W \longrightarrow V$ be a unitary operator. Then $U$ extends to a unitary $U'$ operator on all of $V$.*

*Proof sketch.* Define $U' = U \otimes \mathrm{Id}_{W^\perp}$. □

**Proposition 1.0.13.** *Let $\{M_m\}_{m\in\mathcal{M}}$ be a measurement on $\mathcal{H}$. Then there exists a projective measurement $\{P_m\}_{m\in\mathcal{M}}$, a state space $Q$, and a unitary*

*operator* $U : \mathcal{H} \otimes Q \longrightarrow \mathcal{H} \otimes Q$ *such that for any state* $|\psi\rangle$ *of the composite system* $\mathcal{H} \otimes Q$ *and any* $n \in \mathcal{M}$*:*

$$\langle\psi| U^\dagger P_n^\dagger P_n U |\psi\rangle = \langle\psi| M_n^\dagger M_n |\psi\rangle \tag{23}$$

*Proof.* Let $Q$ be the Hilbert space freely generated by the set $\{|1\rangle, ..., |m\rangle\}$. Define the following linear map.

$$U : \mathcal{H} \longrightarrow \mathcal{H} \otimes Q \tag{24}$$

$$|\psi\rangle = \sum_{m \in \mathcal{M}} M_m |\psi\rangle \otimes |m\rangle \tag{25}$$

We first prove this is unitary, by Corollary A.2.8 it suffices to check that $\langle\psi| U^\dagger U |\psi\rangle = \langle\psi| |\psi\rangle$ for arbitrary $|\psi\rangle \in \mathcal{H}$. We perform the following calculation, note: we have written $\langle\psi| M_m^\dagger \otimes \langle m|$ for the linear functional which sends $a \otimes b$ to the product $\langle\psi| M_m^\dagger a \langle m| b$.

$$\begin{aligned}
\langle\psi| U^\dagger U |\psi\rangle &= \Big( \sum_{m \in \mathcal{M}} \langle\psi| M_m^\dagger \otimes \langle m| \Big) \Big( \sum_{m' \in \mathcal{M}} M_{m'} |\psi\rangle \otimes |m'\rangle \Big) \\
&= \sum_{m \in \mathcal{M}} \sum_{m' \in M} \langle\psi| M_m^\dagger M_{m'} |\psi\rangle \langle m| |m'\rangle \\
&= \sum_{m \in \mathcal{M}} \langle\psi| M_m^\dagger M_{m'} |\psi\rangle \\
&= \langle\psi| |\psi\rangle
\end{aligned}$$

We now want to extend $U$ to a unitary operator on all of $\mathcal{H} \otimes Q$ using Lemma 1.0.12, however we must first identify $\mathcal{H}$ with a subspace of $\mathcal{H} \otimes Q$. There are many ways this can be done, here we choose the basis vector $|1\rangle \in Q$ to be special, and identify $\mathcal{H}$ with $\mathcal{H} \otimes \operatorname{Span} |1\rangle \subseteq \mathcal{H} \otimes Q$.

Now consider the following projective measurement on $\mathcal{H} \otimes Q$:

$$P_m := I_q \otimes |m\rangle \langle m| \tag{26}$$

Then the probability outcome $n$ occurs is:

$$\begin{aligned}
p(n) &= \langle\psi| U^\dagger P_n U |\psi\rangle \\
&= \Big( \sum_{m\in\mathcal{M}} \langle\psi| M_m^\dagger \otimes \langle m| \Big) I_Q \otimes |n\rangle \langle n| \Big( \sum_{m'\in\mathcal{M}} M_{m'} |\psi\rangle \otimes |m\rangle \Big) \\
&= \big( \sum_{m\in\mathcal{M}} \langle\psi| M_m^\dagger \otimes \langle m| \big) \sum_{m'\in\mathcal{M}} M_{m'} |\psi\rangle \otimes |n\rangle \langle n| |m\rangle \\
&= \sum_{m\in\mathcal{M}} \Big( \langle\psi| M_m^\dagger \otimes \langle m| \Big) M_n |\psi\rangle \otimes |n\rangle \\
&= \sum_{m\in\mathcal{M}} \langle\psi| M_m^\dagger M_n |\psi\rangle \langle m| |n\rangle \\
&= \langle\psi| M_n^\dagger M_n |\psi\rangle
\end{aligned}$$

□

**Remark 1.0.14.** The defining equation (25) of the linear map (24) may look opaque. We derive it from a more natural starting point here. See [2, §Partial Trace] for a justification of the natural isomorphisms used in the following calculation.

$$\begin{aligned}
\mathrm{Hom}(Q, \mathrm{Hom}(\mathcal{H}, \mathcal{H})) &\cong \mathrm{Hom}(Q \otimes \mathcal{H}, \mathcal{H}) && (27)\\
&\cong \mathrm{Hom}(\mathcal{H}, \mathcal{H} \otimes Q^*) && (28)
\end{aligned}$$

Then, by identifying $Q$ with $Q^*$ via the anti-linear, isometric bijection given by the Riesz Representation Theorem (see Corollary A.1.4), a linear map $\mathcal{H} \longrightarrow \mathcal{H} \otimes Q$ can be given by a linear map $Q \longrightarrow \mathcal{H} \otimes \mathcal{H}$. We claim that (24) corresponds under this correspondence to the following linear map.

$$\begin{aligned}
Q &\longrightarrow \mathrm{Hom}(\mathcal{H}, \mathcal{H}) && (29)\\
|m\rangle &\longmapsto M_m && (30)
\end{aligned}$$

We now validate this claim. This is a matter of a calculation.

$$\begin{aligned}
\big( |m\rangle \mapsto M_m \big) &\longmapsto \big( |m\rangle \otimes |\psi\rangle \mapsto M_m |\psi\rangle \big) && (31)\\
&\longmapsto \big( \psi \mapsto \sum_{m\in\mathcal{M}} M_m |\psi\rangle \otimes |m\rangle \big) && (32)
\end{aligned}$$

See Corollary [2, 1.2.6] for a justification of the last step.

# 2 The density operator

Let us consider again the Bell state

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{33}$$

thought of as the state of a composite system consisting of two qubits. In Section 1 we thought of this state as having probability $1/2$ that the first and second qubits are in state $|0\rangle$, and a probably $1/2$ that the first and second qubits are in state $|1\rangle$. So what is the probability of the first qubit being in state $|0\rangle$? Presumably $1/2$, but how do we know this?

In short, we have not been precise enough with how the state of a combined system reflects the states of the individual systems.

Obtaining this precision will in fact require reformulating the entirety of what has been done so far, right down to the definition of what a qubit is... Such expositions are excrutiating, so here we provide a justification. In the one qubit case, there is no "combined system", we only have a single qubit. Thus, Section 1 is perfectly valid. In fact, even in situations where composite systems are considered, but scrutinising analysis of the subsystems is not, Section 1 remains valid. For instance, Example 1.0.8 was perfectly precise.

In situations where combined systems are considered and precise analysis of the indivisual subsystems is relevant, such as Example 2.1.10 below, the formalisation of Section 1 is insufficient.

Again, the complication comes from the decision that a composite system is *not* described as a product, but rather a tensor product. Had a composite system been described as a product, then we would have projection morphisms which would be able to relate the multi-qubit case to the single-qubit case. Here though, we need some way of moving from the tensor product of several qubits to a subcollection of qubits. Our tool of choice will be the *partial trace* operator.

## 2.1 Partial trace

For an introduction to the partial trace operator, see [2]

**Example 2.1.1.** We calculate the partial trace of the operator

$$\begin{aligned}\rho &:= \Big(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\Big)\Big(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\Big) \\ &= \frac{|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|}{2}\end{aligned}$$

First consider $\text{Trace}_2(|00\rangle\langle 00|)$. We have, where we write $E_{ij} : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$ for the linear map which maps the $i^{\text{th}}$ basis vector to the $j^{\text{th}}$ basis vector, and similarly for $F_{ij}$ (just applied to the second copy of $\mathbb{C}^2$)

$$|00\rangle\langle 00| = E_{00} \otimes F_{00} \tag{34}$$

and so

$$\text{Trace}_2(|00\rangle\langle 00|) = \text{Trace}(F_{00})E_{00} = |0\rangle\langle 0| \tag{35}$$

Similarly,

$$|11\rangle\langle 00| = E_{01} \otimes F_{01}, \quad |00\rangle\langle 11| = E_{10} \otimes F_{10}, \quad |11\rangle\langle 11| = E_{11} \otimes F_{11} \tag{36}$$

and so

$$\text{Trace}_2(|11\rangle\langle 00|) = \text{Trace}(F_{00})E_{11} = 0 \tag{37}$$

$$\text{Trace}_2(|11\rangle\langle 00|) = \text{Trace}(F_{00})E_{11} = 0 \tag{38}$$

$$\text{Trace}_2(|11\rangle\langle 11|) = \text{Trace}(F_{11})E_{11} = |1\rangle\langle 1| \tag{39}$$

we thus have

$$\text{Trace}_2(\rho) = \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} = I/2 \tag{40}$$

In fact, Example 2.1.1 can be interpreted as deriving the state of a subsystem from a composite system. This involves identifying the state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ with the operator $\rho$ of Example 2.1.1. The following definition generalises Definition 1.0.1 in this way, and also generalises further by allowing for an *ensemble* of states, each weighted by some probability.

Now we can interpret Calculation (40) as extracting the state of the second qubit (as an isolated system) from the composite system. We see the resulting operator corresponds to the state $\frac{1}{2}(|0\rangle + |1\rangle)$ which fits what was said at the start of this section that the state of a single qubit has probability

$1/2$ of being in state $|0\rangle, |1\rangle$ respectively when the combined system is in state (33).

Since positive operators on finite dimension Hilbert spaces are Hermitian (Lemma A.2.12) it follows from the Spectral Decomposition Theorem A.2.2 that positive operators on finite dimensional Hilbert spaces are diagonalisable. It follows that if moreover the trace of a positive operator $H$ is finite, then there exists a finite set of vector $|\psi_1\rangle, \ldots, |\psi_n\rangle$ and probabilities $p_1, \ldots, p_n \in [0, 1]$ so that

$$H = \sum_{i=1}^{n} p_i |\psi_i\rangle \langle\psi_i| \tag{41}$$

Thus we have the following definition of a density operator, which is thought of as an "ensemble of states".

**Definition 2.1.2** (Intrinsic definition of density operator)**.** Let $\mathbb{H}$ be finite dimensional. A **density operator** is a positive operator $\rho : \mathbb{H} \longrightarrow \mathbb{H}$ with trace equal to 1.

**Definition 2.1.3.** Let $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ be a composite system given by the tensor product of two qubits. we have density operator $\rho$ on $\mathcal{H}$, then **tracing over the first copy of** $\mathbb{C}^2$ (respectively, the second copy of $\mathbb{C}^2$) yields the following operators

$$\text{Trace}_2(\rho), \qquad \text{Trace}_1(\rho) \tag{42}$$

We observe a few convenient properties of the trace operator, then give the definitions of *measurement* and *time evolution* for density operators. Once this is done, we can exhibit another interesting phenoma pertaining to Quantum Computing (Example 2.1.10).

**Definition 2.1.4.** The **trace** of an operator $T : \mathcal{H} \longrightarrow \mathcal{H}$ is the trace of any (and hence all) matrix representations of $T$.

The trace of an operator can be computed using a unit vector.

**Lemma 2.1.5.** *Let $A$ be an operator on a Hilbert space $\mathbb{H}$ and let $|\psi\rangle \in \mathbb{H}$ be a unit vector in $\mathbb{H}$. We have the following formula:*

$$\text{Trace}(A |\psi\rangle \langle\psi|) = \langle\psi| A |\psi\rangle \tag{43}$$

*Proof.* In general, if $|v_1\rangle, ..., |v_n\rangle$ is an orthogonal basis for $\mathbb{H}$, and let $A$ is an operator on $\mathbb{H}$ if we write $A|v_j\rangle = a_{1j}|v_1\rangle + \ldots + a_{nj}|v_n\rangle$, then we have

$$\langle v_i| A |v_j\rangle = a_{ij} \tag{44}$$

and so

$$\text{Trace}\, A = \sum_{i=1}^{n} \langle v_i| A |v_i\rangle \tag{45}$$

Applying this to the current statement to be proved, let $|\psi\rangle$ be a unit vector in $\mathbb{H}$ and let $\{|\psi\rangle, |v_2\rangle, ..., |v_n\rangle\}$ be an orthogonal basis for $\mathbb{H}$ (using Gram-Schmidt, say). Then

$$\begin{aligned}\text{Trace}(A|\psi\rangle\langle\psi|) &= \langle\psi| A |\psi\rangle \langle\psi|\psi\rangle + \sum_{i=2}^{n} \langle v_i| A |\psi\rangle \langle\psi|v_i\rangle \\ &= \langle\psi| A |\psi\rangle\end{aligned}$$

□

We see now that Section 1 concerned itself with pure states, where we identify a state vector $|\psi\rangle$ with the operator $|\psi\rangle\langle\psi|$ (that is, we identify the vector $|\psi\rangle$ with the projection onto this vector). We now describe how to generalise the Definitions of Section 1 to the case of mixed states.

**Definition 2.1.6.** A **measurement** on a state space $\mathcal{H}$ is a family of linear operators $\{M_m : \mathcal{H} \longrightarrow \mathcal{H}\}_{m\in\mathcal{M}}$ satisfying

$$\sum_{m\in\mathcal{M}} M^\dagger M = I \tag{46}$$

Associated to every measurement and density operator $\rho$ there is a value

$$p(m) = \text{Trace}(M_m^\dagger M_m \rho) \tag{47}$$

which is understood as the probability $p(m)$ of outcome $m$ on measurement $\{M_m\}_{m\in\mathcal{M}}$.

Also, there is a **resulting density operator**,:

$$\rho_m := \frac{M_m^\dagger \rho M_m}{\text{Trace}(M_m^\dagger M_m \rho)} \tag{48}$$

Definition 2.1.6 becomes more transparent when we pick a diagonalisation of $\rho$. Say

$$\rho = \sum_{i=1}^{n} p_i \left|\psi_i\right\rangle \left\langle\psi_i\right| \tag{49}$$

We have

$$\begin{aligned} p(m) &= \sum_{i=1}^{n} p_i p(m \mid i) \\ &= \sum_{i=1}^{n} p_i \left\langle\psi_i\right| M_m^\dagger M_m \left|\psi_i\right\rangle \\ &= \sum_{i=1}^{n} p_i \,\mathrm{Trace}(M_m^\dagger M_m \left|\psi_i\right\rangle \left\langle\psi_i\right|) \\ &= \mathrm{Trace}(M_m^\dagger M_m \rho) \end{aligned}$$

where the last equality follows from linearity of the Trace.

The resulting density operator is:

$$\rho_m := \sum_{i=1}^{n} p(i \mid m) \frac{M \left|\psi_i\right\rangle \left\langle\psi_i\right| M^\dagger}{p(m \mid i)} \tag{50}$$

we then use Bayes Theorem:

$$p(i \mid m)/p(m \mid i) = p_i/p(m) \tag{51}$$

to obtain:

$$\rho_m = \sum_{i=1}^{n} p_i \frac{M_m \left|\psi_i\right\rangle \left\langle\psi_i\right| M_m^\dagger}{p(m)} = \sum_{i=1}^{n} \frac{M_m \rho M_m^\dagger}{\mathrm{Trace}(M_m^\dagger M \rho)} \tag{52}$$

**Lemma 2.1.7.** *For a pure state density operator $\left|\psi\right\rangle \left\langle\psi\right|$ Definitions 2.1.6 and 1.0.5 agree once $\left|\psi\right\rangle \left\langle\psi\right|$ has been identified with $\left|\psi\right\rangle$.*

**Definition 2.1.8.** Let $\mathcal{H}$ be a state space. A **single step time evolution** on $\mathbb{H}$ is a unitary operator $U : \mathcal{H} \longrightarrow \mathcal{H}$. A **single step time evolution** of a density operator $\rho$ with respect to $U$ is the pair $(\rho, U\rho U^\dagger)$.

An **evolution** of $\mathcal{H}$ is a sequence of unitary operators $(U_1, ..., U_n)$ on $\mathcal{H}$, an **evolution** of a density operator $\rho$ with respect to the evolution $(U_1, ..., U_n)$ is the sequence $(\rho, U\rho U^\dagger, ..., U_n \ldots U_1 \rho U_1^\dagger \ldots U_n^\dagger)$.

**Definition 2.1.9.** Let $\mathcal{H}_1, \mathcal{H}_2$ be two state spaces. The **composite state space** is $\mathcal{H}_1 \otimes \mathcal{H}_2$. We often describe a state of $\mathcal{H}_1 \otimes \mathcal{H}_2$ using the terminology "$\mathcal{H}_1$ is in state $\rho_1$ and $\mathcal{H}_2$ is in state $\rho_2$", this simply describes the state $\rho_1 \otimes \rho_2 \in \mathcal{H}_1 \otimes \mathcal{H}_2$.

With this higher level of fidelity in our theory, we can show another interesting phenomena pertaining to quantum computation.

**Example 2.1.10** (Quantum teleportation)**.** Superposition states clearly have "awareness" of each other, because if we had a pair of qubits which we prepared to the state

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}} \tag{53}$$

and the first qubit was measured and found to be in the state $|0\rangle$, then we know with certainty that the second qubit is in state $|00\rangle$, as the combined state $|01\rangle$ is not an option. Again, this fact can be leaned on to provide an application. The following example shows how a qubit can be sent from one party Alice to another Bob, by only sending a pair of classical bits, provided Alice and Bob are in possession of another pair of qubits which together are in a superposition state.

Consider a pair of qubits which together as a composite system are in the Bell state

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}} \tag{54}$$

Assume that Alice is in posession of the first of these qubits, and Bob is in posession of the second.

Introduce a new qubit $\mathcal{H}$ which is in some state $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$. The system consisting of all three qubits is in state

$$\begin{aligned}
&\frac{1}{\sqrt{2}}(\alpha |0\rangle + \beta |1\rangle)(|00\rangle + |11\rangle) \\
&= \frac{1}{\sqrt{2}}\big(\alpha |0\rangle (|00\rangle + |11\rangle) + \beta |1\rangle (|00\rangle + |11\rangle)\big)
\end{aligned}$$

Alice applies the following unitary matrix (which is written with respect to the ordered basis $|00\rangle, |01\rangle, |10\rangle, |11\rangle$)

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{55}$$

to her pair of qubits, resulting in the following state:

$$\frac{1}{\sqrt{2}}\big(\alpha\,|0\rangle\,(|00\rangle+|11\rangle)+\beta\,|1\rangle\,(|10\rangle+|01\rangle)\big) \tag{56}$$

Then she applies the following unitary matrix to her first qubit

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} \tag{57}$$

resulting in the following state.

$$\frac{1}{2}\big(\alpha(|0\rangle+|1\rangle)(|00\rangle+|11\rangle)+\beta(|0\rangle-|1\rangle)(|10\rangle+|01\rangle)\big) \tag{58}$$

This state can be rewritten as follows.

$$\begin{aligned}&\frac{1}{2}\big(\,|00\rangle\,(\alpha\,|0\rangle+\beta\,|1\rangle)+|01\rangle\,(\alpha\,|1\rangle+\beta\,|0\rangle)\\&+|10\rangle\,(\alpha\,|0\rangle-\beta\,|1\rangle)+|11\rangle\,(\alpha\,|1\rangle-\beta\,|0\rangle)\big)\end{aligned}$$

Now, Alice can perform a measurement on her two qubits, and depending on which outcome $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ the state of Bob's qubit is respectively $\alpha\,|0\rangle+\beta\,|1\rangle, \alpha\,|1\rangle+\beta\,|0\rangle, \alpha\,|0\rangle-\beta\,|1\rangle, \alpha\,|1\rangle-\beta\,|0\rangle$.

Alice can then send Bob the classical bits $00, 01, 10, 11$ respectively indicating that Bob should apply the Unitary matrix respectively $I, X, Z, ZX$, recovering Alice's qubit $|\psi\rangle=\alpha\,|0\rangle+\beta\,|1\rangle$.

## 2.2 Environment

A system interacting with an environment can be modelled as a composite system where one of the systems is the original one in question and the other is the environment.

However, there are particulars we want to consider. We do not want to allow for non-pure states between the system and the environment, and we specifically want to trace over the environment each time. The correct definition is that of a *quantum operation* given below.

**Definition 2.2.1.** An **open quantum system** is the tensor product of two state spaces $\mathcal{H}_p\otimes\mathcal{H}_e$ where $\mathcal{H}_p$ is the **principal system** and $\mathcal{H}_e$ is the **environment**.

**Definition 2.2.2.** The **time evolution of an open quantum system** is that of Definition 1.0.10 where an open quantum system is thought of as a composite system (Definition 1.0.3).

A **quantum operation** on an open quantum system $\mathcal{H}_p \otimes \mathcal{H}_e$ is a triple $(\mathcal{E}, \rho_e, U)$ consisting of an operator

$$\mathcal{E} : \mathrm{Hom}(\mathcal{H}_p, \mathcal{H}_p) \longrightarrow \mathrm{Hom}(\mathcal{H}_p, \mathcal{H}_p) \tag{59}$$

a state $\rho_e \in \mathcal{H}_e$ and a unitary operator $U$ on the entire open quantum system $\mathcal{H}_p \otimes \mathcal{H}_e$. This data is required to satisfy the following for all $\rho \in \mathrm{Hom}(\mathcal{H}_p, \mathcal{H}_p)$.

$$\mathcal{E}(\rho) = \mathrm{Trace}_{\mathcal{H}_e}(U(\rho \otimes \rho_e)U^\dagger) \tag{60}$$

**Remark 2.2.3.** Let $\mathcal{H}_p \otimes \mathcal{H}_e$ be an open quantum system and let $|1\rangle, ..., |n\rangle$ be a basis for $\mathcal{H}_e$. We think of $|1\rangle, ..., |n\rangle$ as operators $\mathbb{C} \longrightarrow \mathcal{H}_e$ and write $\mathrm{id} \otimes |i\rangle$ for the composite $\mathcal{H}_p \longrightarrow \mathcal{H}_p \otimes \mathbb{C} \longrightarrow \mathcal{H}_p \otimes \mathcal{H}_e$. We have for any operator $f : \mathcal{H}_p \otimes \mathcal{H}_e \longrightarrow \mathcal{H}_p \otimes \mathcal{H}_e$ that

$$\mathrm{Trace}_{\mathcal{H}_e}(f) = \sum_{i=1}^{n} (\mathrm{id} \otimes \langle i|) f (\mathrm{id} \otimes |i\rangle) \tag{61}$$

See [2] for background.

We let $|1\rangle, ..., |n\rangle$ be a basis for $\mathbb{H}_e$ and use (61) to rewrite (60) in the special case where $\rho_e = |j\rangle \langle j|$.We have

$$\mathcal{E}(\rho) = \mathrm{Trace}_{\mathbb{H}_e}(U(\rho \otimes |j\rangle \langle j|)U^\dagger) \tag{62}$$

$$= \sum_{i=1}^{n} (\mathrm{id} \otimes \langle i|)(U(\rho \otimes |j\rangle \langle j|)U^\dagger)(\mathrm{id} \otimes |i\rangle \tag{63}$$

Now we make the observation that

$$\rho \otimes |j\rangle \langle j| = (\rho \otimes \mathrm{id})(\mathrm{id} \otimes |j\rangle)(\mathrm{id} \otimes \langle j|) \tag{64}$$

$$= (\mathrm{id} \otimes |j\rangle)(\rho \otimes \mathrm{id})(\mathrm{id} \otimes \langle j|) \tag{65}$$

Substituting (65) into (63) we obtain:

$$\begin{aligned}
&\sum_{i=1}^{n}(\mathrm{id}\otimes\langle i|)(U(\rho\otimes|j\rangle\langle j|)U^{\dagger})(\mathrm{id}\otimes|i\rangle)\\
=&\sum_{i=1}^{n}(\mathrm{id}\otimes\langle i|)U(\mathrm{id}\otimes|j\rangle)(\rho\otimes\mathrm{id})(\mathrm{id}\otimes\langle j|)U^{\dagger}(\mathrm{id}\otimes|i\rangle)\\
=&\sum_{i=1}^{n}(\mathrm{id}\otimes\langle i|)U(\mathrm{id}\otimes|j\rangle)(\mathrm{id}\otimes\langle j|)U^{\dagger}(\mathrm{id}\otimes|i\rangle)\\
=&\sum_{i=1}^{n}E_i\rho E_i^{\dagger}
\end{aligned}$$

where $E_i = (\mathrm{id}\otimes\langle i|)(U((\mathrm{id}\otimes|j\rangle)$.

We thus have a second Definition of a quantum operation:

**Definition 2.2.4.** Given a state space $\mathcal{H}$ (notice, we do not ask for an open quantum system), a **quantum operation** is a pair $(\mathcal{E},\{E_1,...,E_n\})$ consisting of an operator

$$\mathcal{E}:\mathrm{Hom}(\mathcal{H},\mathcal{H})\longrightarrow\mathrm{Hom}(\mathcal{H},\mathcal{H}) \tag{66}$$

and a finite set $\{E_1,...,E_n\}$ of operators on $\mathcal{H}$ subject to the following conditions, where $\rho\in\mathrm{Hom}(\mathcal{H},\mathcal{H})$ is arbitrary.

$$\mathcal{E}(\rho)=\sum_{i=1}^{n}E_i\rho E_i^{\dagger},\qquad\sum_{i=1}^{n}E_i^{\dagger}E_i=I \tag{67}$$

We now have two different definitions of quantum operations, Definition 2.2.2 and Definition 2.2.4. We have already seen in Remark 2.2.3 how to obtain a quantum operation in the sense of Definition 2.2.4 given a quantum operation in the sense of Definition 2.2.2, now we show the converse.

**Remark 2.2.5.** Let $\mathcal{H}$ be a state space and $\{E_1,...,E_n\}$ a quantum operation on $\mathcal{H}$. We introduce the Hilbert space $(\mathbb{C}^2)^{\otimes n}$ which we denote by $\mathcal{H}_e$ and define the following unitary operator.

$$U:\mathcal{H}\longrightarrow\mathcal{H}\otimes\mathcal{H}_e \tag{68}$$

$$|\psi\rangle\longmapsto\sum_{i=1}^{n}E_i|\psi\rangle\otimes|i\rangle \tag{69}$$

We show that this is unitary.

$$\begin{aligned}
\langle\psi| U^\dagger U |\psi\rangle &= \sum_{j=1}^{n} \langle\psi| E_i^\dagger \otimes \langle j|) \sum_{i=1}^{n} E_i |\psi\rangle \otimes |i\rangle \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} \langle\psi| E_j^\dagger E_i |\psi\rangle \langle j| |i\rangle \\
&= \sum_{k=1}^{n} \langle\psi| E_k^\dagger E_k |\psi\rangle \\
&= \langle\psi| |\psi\rangle
\end{aligned}$$

We identify the space $\mathcal{H}$ with the subspace $\mathcal{H} \otimes \operatorname{Span} |1\rangle \subseteq \mathcal{H} \otimes \mathcal{H}_e$, where $|1\rangle \in \mathbb{H}$ is an arbitrarily chosen vector in $\mathcal{H}_e$, and hence by Lemma 1.0.12 the operator $U$ extends to a unitary operator on all of $\mathcal{H} \otimes \mathcal{H}_e$. The next step is to show the following for density operator $\rho := \sum_{i=1}^{m} p_i |\psi_i\rangle \langle\psi_i|$:

$$\operatorname{tr}_{\mathcal{H}_e}(U(\rho \otimes |1\rangle \langle 1|)U^\dagger) = \sum_{k=1}^{n} E_k \rho E_k^\dagger \tag{70}$$

This is shown by the following calculation.

$$\begin{aligned}
\operatorname{tr}_{\mathcal{H}_e}(U(\rho \otimes |1\rangle \langle 1|)U^\dagger) &= \operatorname{tr}_{\mathcal{H}_e}(U(\sum_{i=1}^{n} p_i |\psi_i\rangle \langle\psi_i| \otimes |1\rangle \langle 1|)U^\dagger) \\
&= \operatorname{tr}_{\mathcal{H}_e}(\sum_{i,j,k=1}^{n} p_i E_j |\psi_i\rangle \langle\psi_i| E_k^\dagger \otimes |j\rangle \langle k|) \\
&= \operatorname{tr}_{\mathcal{H}_e}(\sum_{j,k=1}^{n} E_j \rho E_k^\dagger \otimes \langle j| |k\rangle) \\
&= \operatorname{tr}_{\mathcal{H}_e}(\sum_{l=1}^{n} E_l \rho E_l^\dagger \otimes |l\rangle \langle l|) \\
&= \sum_{l=1}^{n} E_l \rho E_l^\dagger
\end{aligned}$$

# 3 Error correction

The more informed two parties are, the more communication may be prone to error while still sustaining certainty on the intended message. This is

because both parties can "error correct" the other.

Throughout, $\mathbb{H}$ denotes a qubit $\mathbb{C}^2$, that is, the complex Hilbert space $\mathbb{C}^2$.

**Definition 3.0.1.** A **message** is a state $|\psi\rangle \in \mathbb{H}^{\otimes n}$, for some $n$. An **error** is a pair of states $(|\varphi\rangle, |\psi\rangle)$ where $|\varphi\rangle, |\psi\rangle \in \mathbb{H}^{\otimes n}$ for some $n$, note that an error may be such that $|\varphi\rangle = |\psi\rangle$. The message $|\varphi\rangle$ is the **intended message** and $|\psi\rangle$ is the **received message**.

**Definition 3.0.2.** An **$n$-encoding of a single state** (sometimes just an **encoding**) is an injective linear map $\iota : \mathbb{H} \longrightarrow \mathbb{H}^{\otimes n}$. An **$n$-encoding of a message** $|m\rangle \in \mathbb{H}^{\otimes k}$ is an $n$-encoding $\iota$ along with a message $|m\rangle \in \mathbb{H}^{\otimes nk}$ for which there exists $|m'\rangle \in \mathbb{H}^{\otimes k}$ satisfying $\iota^{\otimes k} |m'\rangle = |m\rangle$.

**Definition 3.0.3.** A **quantum error correcting code (QECC)** is a pair $\mathcal{Q} = (\mathcal{H}, S)$ consisting of a state space $\mathcal{H}$ along with a set of operators $S$ on $\mathcal{H}$. The elements of $S$ are the **stabilisers**. The **codespace** $\mathcal{H}^S$ of $\mathcal{Q}$ is the maximal subspace of $\mathcal{H}$ invariant under all the operators in $S$.

In Section 5 we will present a method for proving when a set of vectors generate the codespace of a quantum error correction code.

## 3.1 Examples

Throughout, $\mathbb{H}$ denotes a qubit $\mathbb{C}^2$, that is, the complex Hilbert space $\mathbb{C}^2$.

**Definition 3.1.1.** We define the following operators:

$$X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad Y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$
$$Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad H := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

The matrices $X, Y, Z$ are the **Pauli matrices**, and $H$ is the **Hadamard matrix**.

We make the passing observation that all of $X, Y, Z, H$ square to the identity matrix. The basis vectors

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \tag{71}$$

are the **Bell states** and are denoted $|+\rangle, |-\rangle$ respectively. Notice that as already stated, $H^2 = I$, so $H|+\rangle = |0\rangle$ and $H|-\rangle = |1\rangle$.

**Definition 3.1.2.** The standard basis $|0\rangle, |1\rangle$ of $\mathbb{H}$ induces a basis of $\mathbb{H}^{\otimes n}$, we denote $|0\rangle \otimes \ldots \otimes |0\rangle$ by $|0 \ldots 0\rangle$, etc.

**Notation 3.1.3.** Given a Pauli matrix $W \in \{X, Y, Z\}$ the operator on $\mathbb{H}^{\otimes n}$ given by the tensor product consisting of $W$ in the $i^{\text{th}}$ slot (for $i \leq n$) and the identity operator in all other slots by $W_i$. For example, the operator $Z_1$ on $\mathbb{H}^{\otimes 3}$ is the operator $Z \otimes I \otimes I$.

Given a collection of Pauli matrices $W_{i_1}, ..., W_{i_m} \in \{X, Y, Z\}$ where $0 < i_1 < \ldots < i_m \leq n$ we denote by $W_{i_1} \ldots W_{i_m}$ the composition $W_{i_1} \circ \ldots \circ W_{i_m}$. For example, the operator $Z_1 Z_2$ on $\mathbb{H}^{\otimes 3}$ is the operator

$$(Z \otimes I \otimes I) \circ (I \otimes Z \otimes I) = Z \otimes Z \otimes I : \mathbb{H}^{\otimes 3} \longrightarrow \mathbb{H}^{\otimes 3} \tag{72}$$

Consider the **bit flip encoding**

$$\text{BitFlip} : \mathbb{H} \longrightarrow \mathbb{H}^{\otimes 3} \tag{73}$$
$$|0\rangle \longmapsto |000\rangle \tag{74}$$
$$|1\rangle \longmapsto |111\rangle \tag{75}$$

then an encoding of a message with respect to this encoding might be $|000111000\rangle$, but could not be $|000111001\rangle$. We call Encoding 73 the **bit flip encoding**. As another example, we consider the **phase flip encoding**.

$$\begin{aligned} \text{PhaseFlip} : \mathbb{H} &\longrightarrow \mathbb{H}^{\otimes 3} \\ |0\rangle &\longmapsto |+++\rangle \\ |1\rangle &\longmapsto |---\rangle \end{aligned}$$

**Definition 3.1.4.** A **bit flip error** is an error $(|\varphi\rangle, |\psi\rangle)$ where $|\varphi\rangle$ is an encoding of a message with respect to the encoding $\text{BitFlip}^{\otimes m}$ for some $m$, such that $X_i |\varphi\rangle = |\psi\rangle$ for some $i$.

A **phase flip error** is an error $(|\varphi\rangle, |\psi\rangle)$ where $|\varphi\rangle$ is an encoding of a message with respect to the the encoding $\text{PhaseFlip}^{\otimes m}$, such that $Z_i |\varphi\rangle = |\psi\rangle$ for some $i$.

Let $(|\varphi\rangle, |\psi\rangle)$ be a bit flip error. The following algorithm takes as input $|\psi\rangle$ and reconstructs $|\varphi\rangle$: [Bit flip correction] Input: a received message $|\psi\rangle$,

1. perform the following projective measurements:

$$\langle\psi| Z_1 Z_2 |\psi\rangle \text{ with resulting state } |\psi'\rangle, \tag{76}$$

followed by

$$\langle\psi'|\, Z_2Z_3\, |\psi'\rangle \tag{77}$$

let $(r_1, r_2)$ be the pair of results from these measurements.

2. It will be shown that $r_1, r_2 \in \{1, -1\}$, and the resulting state of the second measurement is $|\psi\rangle$.

3. Now retrieve $|\varphi\rangle$ based on the values of $r_1, r_2$:

   - if $(r_1, r_2) = (1, 1)$, return $|\psi\rangle$,
   - if $(r_1, r_2) = (-1, 1)$, return $X_1\,|\psi\rangle$,
   - if $(r_1, r_2) = (1, -1)$, return $X_3\,|\psi\rangle$,
   - if $(r_1, r_2) = (-1, -1)$, return $X_2\,|\psi\rangle$

We now prove correctness of Algorithm 3.1:

*Proof.* It will be helpful to first notice:

$$\begin{aligned}
Z_1Z_2\,|000\rangle &= |000\rangle & Z_1Z_2\,|001\rangle &= |001\rangle \\
Z_1Z_2\,|010\rangle &= -\,|010\rangle & Z_1Z_2\,|011\rangle &= -\,|011\rangle \\
Z_1Z_2\,|100\rangle &= -\,|100\rangle & Z_1Z_2\,|101\rangle &= -\,|101\rangle \\
Z_1Z_2\,|110\rangle &= |110\rangle & Z_1Z_2\,|111\rangle &= |111\rangle
\end{aligned}$$

Let $|\psi\rangle := a\,|010\rangle + b\,|101\rangle$ be a state, ie, an element of $\mathbb{H}^{\otimes 3}$. We perform the measurement $Z_1Z_2$ followed by $Z_2Z_3$:

$$\begin{aligned}
\langle\psi|\, Z_1Z_2\,|\psi\rangle &= (a\,\langle 010| + b\,\langle 101|)Z_1Z_2(a\,|010\rangle + b\,|101\rangle) \\
&= (a\,\langle 010| + b\,\langle 101|)(-a\,|010\rangle - b\,|101\rangle) \\
&= -a^2 - b^2 = -1
\end{aligned}$$

and

$$\begin{aligned}
\langle\psi|\, Z_2Z_3\,|\psi\rangle &= (a\,\langle 010| + b\,\langle 101|)Z_1Z_2(a\,|010\rangle + b\,|101\rangle) \\
&= (a\,\langle 010| + b\,\langle 101|)(-a\,|010\rangle - b\,|101\rangle) \\
&= -a^2 - b^2 = -1
\end{aligned}$$

We can infer from the fact that both of these came out as $-1$ that it was the second bit which was flipped, and so we can correct this. However, what is the impact of this measurement on the state? Again we calculate:

$$\begin{aligned} Z_1 Z_2(a\,|010\rangle + b\,|101\rangle) &= Z_1(-a\,|010\rangle + b\,|101\rangle) \\ &= -a\,|010\rangle - b\,|101\rangle \end{aligned}$$

and

$$\begin{aligned} Z_2 Z_3(-a\,|010\rangle - b\,|101\rangle) &= Z_2(-a\,|010\rangle + b\,|101\rangle) \\ &= a\,|010\rangle + b\,|101\rangle \end{aligned}$$

and so the measurements (in the end) did not impact our state. □

Later, using the theory of *stabiliser codes*, we will show that in fact single bit flip errors form the full set of correctable errors using $Z_1Z_2, Z_1Z_3, Z_2Z_3$.

Let $(|\varphi\rangle, |\psi\rangle)$ be a phase flip error. The following algorithm takes as input $|\psi\rangle$ and reconstructs $|\varphi\rangle$: [Phase flip correction] Input: a received message $|\psi\rangle$:

1. perform the following projective measurements:

$$\langle\psi|\, X_1X_2\,|\psi\rangle \text{ with resulting state } |\psi'\rangle \tag{78}$$

   followed by

$$\langle\psi'|\, X_2X_3\,|\psi'\rangle \tag{79}$$

   let $(r_1, r_2)$ be the pair of results from these measurements.

2. It will be shown that $r_1, r_2 \in \{1, -1\}$ and the resulting state of the second measurement is $|\psi\rangle$.

3. Now retrieve $|\varphi\rangle$ based on the values of $r_1, r_2$:

   (a) if $(r_1, r_2) = (1, 1)$, return $|\psi\rangle$,

   (b) if $(r_1, r_2) = (-1, 1)$, return $Z_1\,|\psi\rangle$

   (c) if $(r_1, r_2) = (1, -1)$, return $Z_3\,|\psi\rangle$,

   (d) if $(r_1, r_2) = (-1, -1)$, return $Z_2\,|\psi\rangle$

*Proof.* In fact all our work is already done. We simply note that $Z|+\rangle = |-\rangle, Z|-\rangle = |+\rangle$ (and so phase flip acts like bit flip for $|+\rangle, |-\rangle$), and that in general

$$H^{\otimes n} Z_{i_1} \dots Z_{i_j} H^{\otimes n} = X_{i_1} \dots X_{i_j} \tag{80}$$

The result then follows from the proof of correctness for Algorithm 3.1. □

What if we wanted to correct an error where we knew the received message corresponded to the intended message by either a bit flip *or* a phase flip. This can be done by combining the two approaches above. Define the following encoding:

**Definition 3.1.5.** The **Shor encoding** is:

$$\text{Shor} : \mathbb{H} \longrightarrow \mathbb{H}^{\otimes 9} \tag{81}$$

where

$$\text{Shor}(|\psi\rangle) = \text{BitFlip} \circ \text{PhaseFlip}\,|\psi\rangle \tag{82}$$

On input $|\psi\rangle$:

1. Perform the following projective measurements:

$$\begin{aligned}
&\langle\psi| Z_1 Z_2 |\psi\rangle \text{ with resulting state } |\psi'\rangle \\
&\langle\psi'| Z_2 Z_3 |\psi'\rangle \text{ with resulting state } |\psi\rangle \\
&\langle\psi| Z_3 Z_4 |\psi\rangle \text{ with resulting state } |\psi''\rangle \\
&\langle\psi'| Z_4 Z_5 |\psi''\rangle \text{ with resulting state } |\psi\rangle \\
&\langle\psi'| Z_5 Z_6 |\psi\rangle \text{ with resulting state } |\psi'''\rangle \\
&\langle\psi'| Z_6 Z_7 |\psi'''\rangle \text{ with resulting state } |\psi\rangle \\
&\langle\psi'| Z_7 Z_8 |\psi\rangle \text{ with resulting state } |\psi''''\rangle \\
&\langle\psi'| Z_8 Z_9 |\psi''''\rangle \text{ with resulting state } |\psi\rangle
\end{aligned}$$

   let $(r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8) \in \mathbb{Z}_2^8$ be the results from these measurements. Notice that there are only three possibilities, all entries are 1, exactly one entry is $-1$ in which case it is either $r_1$ or $r_8$ (with the rest equal to 1) or exactly two values are $-1$ and the rest are 1 in which case the two $-1$ entries are neighbours.

2. Then perform the following measurements:

$$\langle\psi| X_1X_2X_3X_4X_5X_6 |\psi\rangle \text{ with resulting state } |\psi'\rangle$$
$$\langle\psi'| X_4X_5X_6X_7X_8X_9 |\psi'\rangle \text{ with resulting state } |\psi\rangle$$

let $(s_1, s_2) \in \mathbb{Z}_2^2$ be the result of these measurements.

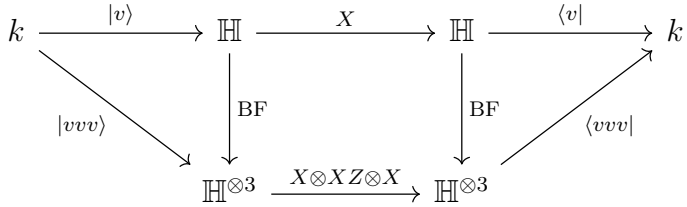
Now retrieve $|\varphi\rangle$ based on the values:

| $(r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8)$ | $(s_1, s_2)$ | Return |
|---|---|---|
| $(1,1,1,1,1,1,1,1)$ | $(1,1)$ | $\|\psi\rangle$ |
| $(-1,1,1,1,1,1,1,1,1)$ | $(1,1)$ | $X_1\|\psi\rangle$ |
| $(-1,-1,1,1,1,1,1,1,1)$ | $(1,1)$ | $X_2\|\psi\rangle$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $(1,1,1,1,1,1,1,-1,-1)$ | $(1,1)$ | $X_8\|\psi\rangle$ |
| $(1,1,1,1,1,1,1,1,-1)$ | $(1,1)$ | $X_9\|\psi\rangle$ |
| $(1,1,1,1,1,1,1,1)$ | $(-1,1)$ | $\|\psi\rangle$ |
| $(-1,1,1,1,1,1,1,1,1)$ | $(-1,1)$ | $Z_1Z_2Z_3X_1\|\psi\rangle$ |
| $(-1,-1,1,1,1,1,1,1,1)$ | $(-1,1)$ | $Z_1Z_2Z_3X_2\|\psi\rangle$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $(1,1,1,1,1,1,1,-1,-1)$ | $(-1,-1)$ | $Z_4Z_5Z_6X_8\|\psi\rangle$ |
| $(1,1,1,1,1,1,1,1,-1)$ | $(-1,-1)$ | $Z_4Z_5Z_6X_9\|\psi\rangle$ |
| $(-1,1,1,1,1,1,1,1,1)$ | $(-1,-1)$ | $Z_4Z_5Z_6X_1\|\psi\rangle$ |
| $(-1,-1,1,1,1,1,1,1,1)$ | $(-1,-1)$ | $Z_4Z_5Z_6X_2\|\psi\rangle$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $(1,1,1,1,1,1,1,-1,-1)$ | $(-1,-1)$ | $Z_4Z_5Z_6X_8\|\psi\rangle$ |
| $(1,1,1,1,1,1,1,1,-1)$ | $(-1,-1)$ | $Z_4Z_5Z_6X_9\|\psi\rangle$ |
| $(1,1,1,1,1,1,1,-1,-1)$ | $(-1,-1)$ | $Z_4Z_5Z_6X_8\|\psi\rangle$ |
| $(1,1,1,1,1,1,1,1,-1)$ | $(-1,-1)$ | $Z_4Z_5Z_6X_9\|\psi\rangle$ |
| $(-1,1,1,1,1,1,1,1,1)$ | $(-1,1)$ | $Z_7Z_8Z_9X_1\|\psi\rangle$ |
| $(-1,-1,1,1,1,1,1,1,1)$ | $(-1,1)$ | $Z_7Z_8Z_9X_2\|\psi\rangle$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $(1,1,1,1,1,1,1,-1,-1)$ | $(-1,1)$ | $Z_7Z_8Z_9X_8\|\psi\rangle$ |
| $(1,1,1,1,1,1,1,1,-1)$ | $(-1,1)$ | $Z_7Z_8Z_9X_9\|\psi\rangle$ |

*Proof.* That the Shor algorithm corrects bit flip errors is obvious.

Now assume a single phase flip error has occurred, and no bit flip error has occurred. The core observation is the commutativity of the following

diagrams for any $|v\rangle \in \{|0\rangle, |1\rangle\}$ (where we think of any "ket" vector $|v\rangle$ as a map $k \longrightarrow |v\rangle$). We have written BF for BitFlip:




(83)

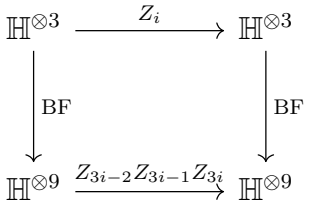
A similar Diagram but with $X \otimes XZ \otimes X$ replaced by $XZ \otimes X \otimes X$ or by $X \otimes X \otimes XZ$ also commutes. Thus, if $|\psi\rangle = Z_i |\varphi\rangle$, defining

$$s(i) = \begin{cases} 1, & i = 1, 2, 3 \\ 2, & i = 4, 5, 6 \\ 3, & i = 7, 8, 9 \end{cases} \tag{84}$$

we have

$$\langle m| \operatorname{BF}^{\otimes 3\dagger} Z_i^\dagger X_1 X_2 X_3 X_4 X_5 X_6 Z_i \operatorname{BF}^{\otimes 3} |m\rangle = \langle m| Z_{s(i)}^\dagger X_1 X_2 Z_{s(i)} |m\rangle \tag{85}$$

and so we can treat each "block" of three states as a single state, so we know how to interpret the measurements $(s_1, s_2)$. The last observation to make is commutativity of the following Diagram for all $i = 1, 2, 3$




(86)

Now say a combination of a bit flip and a phase flip error occurred. That is, say $|\psi\rangle = Z_i X_j |\varphi\rangle$. The error correction will first correct the bit flip which reduces to the previous case. In other words, $X_j^2 |\varphi\rangle = |\varphi\rangle$ is in the image of BitFlip. □

# 4 General error correction

This section is the climax of this document, and Theorem 4.0.9 is the climax of this section. It presents the general error correction conditions as advertised

in the Introduction. That is, Theorem 4.0.9 present necessary and sufficient conditions for a set of operators to be correctable, in the sense made precise by Definition 4.0.4.

To prove Theorem 4.0.9 we go back to definition of a density operator and of quantum operators and make the observation that we did *not* describe a canonical presentation of either. That is, a density operator is an operator which *admits* a certain form, and likewise for quantum operators. Choices of presentations of particular density and quantum operators are not unique. Our first goal is to establish how two distinct presentations relate to each other. See [1, Page 103] for an example of two different diagonalisations of the same density operator. The following Proposition describes the relationship between these different diagonalisations.

**Proposition 4.0.1.** *Let $\rho = \sum_{i=1}^{n} p_i |\psi_i\rangle \langle\psi_i|$, where $|\psi_1\rangle, \ldots, |\psi_n\rangle$ is some explicit choice of vectors, be a positive (and hence diagonalisable) operator on a Hilbert space $\mathcal{H}$. Let $|1\rangle, ..., |m\rangle$ be an orthonormal set of vectors so that $\rho$ written as a matrix with respect to $|1\rangle, ..., |m\rangle$ is diagonal. Let $\lambda_1, ..., \lambda_m$ denote the eigenvalues corresponding to the eigenvectors $|1\rangle, ..., |m\rangle$. Let $r$ denote $\max\{n, m\}$. Then there exists a unitary matrix $A = (a_{ij})_{1 \leq i,j \leq r}$ so that for all $i = 1, ..., n$*

$$p_i |\psi_i\rangle \langle\psi_i| = \sum_{j=1}^{m} a_{ij} \lambda_j |j\rangle \langle j| \tag{87}$$

*Proof.* If $n < m$ then can define $|\psi_{n+1}\rangle = \ldots = |\psi_m\rangle = 0$ and $p_{n+1} = \ldots = p_m = 0$ so that it is sufficient to consider the case when $n \geq m$.

Since $|1\rangle, ..., |m\rangle$ form an orthonormal basis for $\mathcal{H}$ we can write for each $i = 1, ..., n$ the following, where $a_{i1}, ..., \alpha_{im} \in \mathbb{C}$

$$\sqrt{p_i} |\psi_i\rangle = \sum_{j=m}^{n} a_{ij} \sqrt{\lambda_j} |j\rangle \tag{88}$$

Hence we have the following calculation.

$$\rho = \sum_{i=1}^{n} p_i \left|\psi\right\rangle \left\langle\psi\right| \tag{89}$$

$$= \sum_{i=1}^{n} \sqrt{p_i} \left|\psi_i\right\rangle \sqrt{p_i} \left\langle\psi_i\right| \tag{90}$$

$$= \sum_{i=1}^{n} \sum_{j,j'=1}^{m} a_{ij}\overline{a_{j'i}}\sqrt{\lambda_j\lambda_{j'}} \left|j\right\rangle \left\langle j'\right| \tag{91}$$

$$= \sum_{j=1}^{m} \lambda_k \left|k\right\rangle \left\langle k\right| \tag{92}$$

It follows from this that $\sum_{j,j'=1}^{m} a_{ij}\overline{a_{j'i}} = 0$ if $j \neq j'$ and $\sum_{j=1}^{n} a_{ij}\overline{a_{ji}} = 1$. Thus, if $m = n$ the matrix $(a_{ij})_{1\leq i,j\leq n}$ is untary. If $m > n$ then we define $a_{ij} = 0$ for $j = n+1, ..., m$ and $i = 1, ..., m$ and arrive at a square, unitary matrix. □

**Corollary 4.0.2.** *If $\rho = \sum_{i=1}^{n} p_i \left|\psi_i\right\rangle \left\langle\psi_i\right| = \sum_{j=1}^{m} q_j \left|\varphi_j\right\rangle \left\langle\varphi_j\right|$ and $r$ denotes $\max\{n, m\}$ then there is a positive operator then there exists a unitary matrix*

$$A = (a_{ij})_{1\leq i,j\leq r} \tag{93}$$

*so that for all $i = 1, ..., n$*

$$p_i \left|\psi\right\rangle \left\langle\psi_i\right| = \sum_{j=1}^{r} a_{ij}q_j \left|\varphi_j\right\rangle \left\langle\varphi_j\right| \tag{94}$$

Moreover, in Section 2, the choice of operators $\{E_1, \ldots, E_n\}$ for a quantum operation was also not given a canonical form. Indeed, the operators $\{E_1, \ldots, E_n\}$ are not uniquely determined by the operator $\sum_{i=1}^{n} E_i^\dagger \rho E_i$. The following proposition classifies this discrepency.

**Proposition 4.0.3.** *Let $\{E_1, ..., E_n\}$ and $\{F_1, ..., F_m\}$ be two sets of operators on a Hilbert space $\mathcal{H}$ so that for all positive operators $\rho$ on $\mathcal{H}$ we have*

$$\sum_{i=1}^{n} E_i \rho E_i^\dagger = \sum_{i=1}^{m} F_i \rho F_i^\dagger \tag{95}$$

*If $r$ denote $\max\{n, m\}$, then there exists a unitary matrix $(a_{ij})_{1\leq i,j\leq r}$ so that for each $i = 1, ..., n$ we have*

$$E_i = \sum_{j=1}^{m} a_{ij} F_j \tag{96}$$

*The converse also holds.*

*Proof.* Say the dimension of $\mathcal{H}$ is $k$ and $|1\rangle, ..., |k\rangle$ is an orthonormal basis. The proof will proceed by introducing a new Hilbert space, $\mathcal{Q}$, which is freely generated by $|1\rangle, ..., |k\rangle$ and then we define a positive operator $\sigma$ on $\mathcal{H} \otimes \mathcal{Q}$. We then appeal to Corollary 4.0.2.

Let $|\alpha\rangle$ denote the vector $\sum_{i=1}^{k} |i\rangle \otimes |i\rangle \in \mathcal{H} \otimes \mathcal{Q}$. For each $i = 1, ..., k$ define the following vectors in $\mathcal{H} \otimes \mathcal{Q}$.

$$|e_i\rangle := \sum_{j=1}^{k} E_i |j\rangle \otimes |j\rangle \tag{97}$$

$$|f_i\rangle := \sum_{j=1}^{k} F_i |j\rangle \otimes |j\rangle \tag{98}$$

Define the following operators on $\mathcal{H} \otimes |Q\rangle$.

$$\sum_{i=1}^{n} |e_i\rangle \langle e_i| \tag{99}$$

$$\sum_{i=1}^{m} |f_i\rangle \langle f_i| \tag{100}$$

The operators (99), (100) are equal, which we now justify. Notice first that for any $j = 1, ..., k$ the operator $|j\rangle \langle j|$ is positive, as for any $|\psi\rangle \in \mathcal{H}$ we have

$$\langle\psi| |j\rangle \langle j| |\psi\rangle = |\langle\psi| |\psi\rangle|^2 \geq 0 \tag{101}$$

This along with (95) justifies (104) in the following calculation.

$$\sum_{i=1}^{n} |e_i\rangle \langle e_i| = \sum_{i=1}^{n} \sum_{j,j'=1}^{k} E_i |j\rangle \langle j'| E_i^\dagger \otimes |j\rangle \langle j'| \tag{102}$$

$$= \sum_{j,j'=1}^{k} \Big( \sum_{i=1}^{n} E_i |j\rangle \langle j'| E_i^\dagger \Big) \otimes |j\rangle \langle j'| \tag{103}$$

$$= \sum_{j,j'=1}^{k} \Big( \sum_{i=1}^{m} F_i |j\rangle \langle j'| F_i^\dagger \Big) \otimes |j\rangle \langle j'| \tag{104}$$

$$= \sum_{i=1}^{m} |f_i\rangle \langle f_i| \tag{105}$$

Thus, by Corollary 4.0.2, if $r$ denotes $\max\{n, m\}$, there exists a unitary matrix $(a_{ij})_{1 \leq i,j \leq r}$ so that for each $i = 1, ..., n$ we have the following, where if $m > n$ we set $|f_{n+1}\rangle = \ldots = |f_m\rangle = 0$ and if $n > m$ we set $a_{i(n+1)} = \ldots = a_{im} = 0$.

$$|e_i\rangle = \sum_{j=1}^{r} a_{ij} |f_j\rangle \tag{106}$$

It now remains to show that $E_i = \sum_{j=1}^{r} a_{ij} F_j$. To do this, we use the following trick. Let $|\psi\rangle \in \mathcal{H}$ and write $|\psi\rangle = \alpha_1 |1\rangle + \ldots + \alpha_k |k\rangle$. We let consider the linear functional $\sum_{l=1}^{k} \alpha_l \langle l|$. We consider also the linear function $\mathrm{id}_{\mathcal{H}} \otimes \langle j| : \mathcal{H} \otimes \mathcal{Q} \longrightarrow \mathcal{H}$. This has the following property.

$$\Big( \mathrm{id}_{\mathcal{H}} \otimes \big( \sum_{l=1}^{k} \alpha_l \langle l| \big) \Big) |e_i\rangle = \sum_{j=1}^{k} E_i |j\rangle \otimes \sum_{l=1}^{k} \alpha_l \langle l| |j\rangle \tag{107}$$

$$= \sum_{j=1}^{k} \alpha_j E_i |j\rangle \tag{108}$$

$$= E_i |\psi\rangle \tag{109}$$

Combining this calculation with (106) we obtain (96).

Now we prove the converse, this is a simple calculation.

$$\sum_{i=1}^{n} F_i \rho F_i^\dagger = \sum_{i,j,j'=1}^{n} a_{ij} \overline{a}_{ji} E_i \rho E_i^\dagger = \sum_{i=1}^{n} E_i \rho E_i^\dagger \tag{110}$$

□

Necessary conditions for quantum error correction follow as a corollary to Proposition 4.0.3.

**Definition 4.0.4.** Let $\mathcal{H}$ be a Hilbert space and $(\mathcal{E}, \{E_1, ..., E_n\})$ a quantum operation over $\mathcal{H}$ and let $C \subseteq \mathcal{H}$ be a codespace (that is, $C \subseteq \mathcal{H}$ is a subspace). The quantum operation $\mathcal{E}$ is a **correctable set of errors for** $C$ if it satisfies the following condition: let $|1\rangle, ..., |l\rangle$ denote an orthonormal basis for $C$. Let $\mathcal{H}_e$ denote the complex Hilbert space freely generated by $|1\rangle_e, ..., |n\rangle_e$. There exists a quantum operation $\mathcal{R} = \{R_1, ..., R_m\}$ along with a set of complex numbers $\{\alpha_{jk}\}_{1 \leq j \leq n, 1 \leq k \leq m}$ so that for each $i = 1, ..., n$ we have the following, where $|1\rangle_a, ..., |m\rangle_a$ is a basis for the free complex Hilbert space of dimension $m$.

$$\sum_{j=1}^{n} \sum_{k=1}^{m} R_k E_j |i\rangle \otimes |j\rangle_e \otimes |k\rangle_a = |i\rangle \otimes \sum_{k=1}^{n} \sum_{l=1}^{m} \alpha_{jk} |j\rangle_e \otimes |k\rangle_a \tag{111}$$

there, $\rho : C \longrightarrow C$ is an operator on $C$.

**Remark 4.0.5.** By Lemma 4.0.6 below, condition (111) implies that there exists a family of complex numbers $\{\lambda_{jk}\}_{1 \leq j \leq n, 1 \leq k \leq m}$ so that

$$R_k E_j |i\rangle = \sqrt{\lambda_{jk}} |i\rangle \tag{112}$$

In other words,

$$R_k E_j |i\rangle \langle i| E_j^\dagger R_k^\dagger = \lambda_{jk} |i\rangle \langle i| \tag{113}$$

This is what [1] mean when they write the condition

$$\mathcal{R}(\mathcal{E}(\rho)) \propto \rho \tag{114}$$

there, $\rho$ is a positive operator.

**Lemma 4.0.6.** *Let $(H, Q)$ be a pair of Hilbert spaces, let $|1\rangle, ..., |n\rangle$ and $|\bar{1}\rangle, ..., |\bar{m}\rangle$ respectively be orthonormal basis vectors for $H, Q$. Also, for each $j = 1, ..., m$ let $M_j$ be a bounded linear operator on $H$. If there exists $\alpha_1, ..., \alpha_m \in \mathbb{C}$ so that for all $i = 1, ..., n$ we have:*

$$\sum_{j=1}^{m} M_j |i\rangle \otimes |\bar{j}\rangle = |i\rangle \otimes \Big( \sum_{j=1}^{m} \alpha_j |\bar{j}\rangle \Big) \tag{115}$$

*then there exists $\lambda_1, ..., \lambda_m \in \mathbb{C}$ so that for all $j = 1, ..., m$*

$$M_j |i\rangle = \lambda_j |i\rangle \tag{116}$$

*Proof.* For each $j = 1, ..., m$ write

$$M_j |i\rangle = \beta_1^{i,j} |1\rangle + \ldots + \beta_n^{i,j} |n\rangle \tag{117}$$

Then we have

$$\begin{aligned}\sum_{j=1}^{m} M_j |i\rangle \otimes |\bar{j}\rangle &= \sum_{j=1}^{m} (\beta_1^{i,j} |1\rangle + \ldots + \beta_n^{i,j} |n\rangle) \otimes |\bar{j}\rangle \\ &= \sum_{j=1}^{m} \sum_{i'=1}^{n} \beta_{i'}^{i,j} |i'\rangle \otimes |\bar{j}\rangle\end{aligned}$$

which by assumption is equal to

$$|i\rangle \otimes \Big( \sum_{j=1}^{m} \alpha_j |\bar{j}\rangle \Big) = \sum_{j=1}^{m} \alpha_j |i\rangle \otimes |\bar{j}\rangle \tag{118}$$

It follows that $\beta_{i'}^{i,j} = 0$ if $i' \neq i$. The result follows. □

In light of Remark 4.0.5 we may make the following, equivalent definition of a correctable set of operators (Definition 4.0.4).

**Definition 4.0.7.** Let $(\mathcal{E}, \{E_1, ..., E_n\})$ be a quantum operator on a Hilbert space $\mathcal{H}$ and let $C \subseteq \mathcal{H}$ be a codespace. The quantum operator is a **correctable set of errors** if there exists a trace-preserving quantum operator $(\mathcal{R}, \{R_1, ..., R_m\})$ and a complex number $\lambda \in \mathbb{C}$ so that for any positive operator $\rho : C \longrightarrow C$ the following holds.

$$\mathcal{R}(\mathcal{E}(\rho)) = \lambda \rho \tag{119}$$

**Remark 4.0.8.** In Section 3.1 we considered error correction codes which had "multiple steps". For instance, the bitflip error correcting algorithm (Algorithm 3.1) has *two* diagnoses involved, first that from the measurement $Z_1 Z_2$ and then that form the measurement $Z_2 Z_3$. In Definition 4.0.7 we ask for more than this, we ask that there exists a *single* quantum operator $\mathcal{R}$ which in the proof of Theomem 4.0.9 below we will see involves a *single* diagnosis.

Thus, we will not extract the exact algorithms considered in Section 3.1 from the general theory of this section. We can however apply the result

of this section to the cases considered in Section 3.1 to obtain something different, we do this at the end of this section.

The flow of content for this document should thus be read as follows: in Section 3.1 we saw that error correction (in some sense) was possible for some particular examples, and in this section we classify when error correction in the sense of Definition 4.0.7 is possible.

**Theorem 4.0.9.** *Let $\mathbb{H}$ be a qubit and $C \subseteq \mathbb{H}$ a codespace, ie, $\mathbb{H}$ is $(\mathbb{C}^{\otimes 2})^n$ for some $n$ and $C \subseteq \mathbb{H}$ is a subspace. Let $P$ denote the projection onto $C$. Suppose $\mathcal{E}$ is a quantum operation (Definition 2.2.4) with operator elements $\{E_1, ..., E_n\}$. Then there exists a trace preserving quantum operation $\mathcal{R}$ which corrects $\mathcal{E}$ (Definition 4.0.4) if and only if there exists a Hermitian matrix $A = (\alpha_{ij})_{1 \leq i,j \leq n}$ satisfying*

$$\forall i, j = 1, \ldots, n \quad PE_i^\dagger E_j P = \alpha_{ij} P \tag{120}$$

*Proof.* First we show that these conditions are necessary. If $\mathcal{R}$ exists, that is, if there is a collection of operators $\{R_1, ..., R_m\}$ on $\mathbb{H}$ so that for all $\rho \in \mathrm{Hom}(\mathbb{H}, \mathbb{H})$ there exists a family of complex numbers $\{\lambda_{jk}\}_{1 \leq j \leq n, 1 \leq k \leq m}$ so that

$$R_k E_j P \rho P E_j^\dagger R_k^\dagger = \lambda_{jk} P \rho P \tag{121}$$

This is because $P\rho P$ is an operator on $C$. In other words, there exists $\lambda \in \mathbb{C}$ so that

$$\mathcal{R}(\mathcal{E}(P\rho P)) = \lambda P \rho P \tag{122}$$

Let $\mu \in \mathbb{C}$ be a complex number so that $\mu^2 = \lambda$. The two operators induced by the sets $\{R_k E_j P\}_{1 \leq j \leq n, 1 \leq k \leq m}$ and $\{\mu P\}$ together satisfy the hypothesis of Proposition 4.0.3. Consider lexicographic ordering on the set $\{(j,k) \mid 1 \leq j \leq n, 1 \leq k \leq m\}$ induced by the standard order $<$ on the integers. With respect to this indexing, there exists a unitary matrix $(a_{jk,j'k'})_{1 \leq j,j' \leq n, 1 \leq k,k' \leq m}$ subject to the following.

$$R_k E_j P = \sum_{(k',j')} a_{kj,k'j'} \mu P \tag{123}$$

It follows that

$$PE_i^\dagger R_k^\dagger R_k E_j P = \sum_{(k',i')} \sum_{(k'',j')} \overline{a_{k'i',ki}} a_{kj,k''j'} \mu \overline{\mu} P \tag{124}$$

Now we set

$$\alpha_{ij} = \sum_{k=1}^{m} \sum_{(k',i')} \sum_{(k'',j')} \overline{a_{k'i',ki}} a_{kj,k''j'} \mu \overline{\mu}$$

and sum (124) over all $k$ to obtain

$$PE_i^\dagger E_j P = \alpha_{ij} P \tag{125}$$

as required, it is easy to see that $(\alpha_{ij})_{1 \leq i,j \leq n}$ is Hermitian.

Now we prove sufficiency.

First we simplify the error correction conditions by diagonalising the Hermitian matrix $A$. Let $D$ be diagonal and $U$ unitary such that $D = U^\dagger AU$. Denote the entry in row $i$ and column $j$ of $U$ by $u_{ij}$, similary for $u_{ij}^\dagger$. For each $k = 1, ..., n$ we define operators $F_k = \sum_{i=1}^{n} u_{ij} E_i$. Notice that by Proposition 4.0.3 we have

$$\sum_{i=1}^{n} F_i \rho F_i^\dagger = \sum_{i=1}^{n} E_i \rho E_i^\dagger \tag{126}$$

We then calculate, for $k, l \in \{1, ..., n\}$:

$$PF_k^\dagger F_l P = \sum_{i,j=1}^{n} u_{ki}^\dagger u_{jl} P E_i^\dagger E_j P \tag{127}$$

Substituting (120) we have $PF_k^\dagger F_l P = \sum_{i,j=1}^{n} u_{ki}^\dagger \alpha_{ij} u_{jl} P$ and since $D = U^\dagger AU$ we obtain:

$$PF_k^\dagger F_l P = d_{kl} P \tag{128}$$

Now we make use of polar decomposition (Theorem A.2.15). There exists for each $k = 1, ..., m$ a unitary matrix $U_k$ such that $F_k P = U_k \sqrt{PF_k^\dagger F_k P} = \sqrt{d_{kk}} U_k P$. We define $P_k := U_k P U_k^\dagger$, these operators $P_k$ are the syndrome measurement. We will make use of the following observation.

$$P_k = F_k P U_k^\dagger / \sqrt{d_{kk}} \tag{129}$$

We define $\mathcal{R} = \{U_1^\dagger P_1, ..., U_n^\dagger P_n\}$ with corresponding operator $\mathcal{R}(\rho) = \sum_{i=1}^{n} U_i^\dagger P_i \rho P_i U_i$.

We now have the following incredible calculation.

$$\begin{aligned}
\mathcal{R}(\mathcal{E}(\rho)) &= \sum_{i,j=1}^{n} U^\dagger P_j E_i \rho E_i^\dagger P_j U_j \\
&= \sum_{i,j=1}^{n} U_j^\dagger P_j F_i \rho F_i^\dagger P_j U_j && \text{By (126)} \\
&= \sum_{i,j=1}^{n} U_j^\dagger P_j^\dagger F_i P \rho P F_i^\dagger P_j U_j \\
&= \sum_{i,j=1}^{n} U_j^\dagger U_j P F_j^\dagger F_i P \rho P F_i^\dagger F_j P U_j^\dagger U_j / d_{jj} && \text{By (129)} \\
&= \sum_{i,j=1}^{n} d_{ji} \rho d_{ij} / d_{jj} && \text{By (128)} \\
&= \sum_{i=1}^{n} d_{ii} \rho \\
&\propto \rho
\end{aligned}$$

□

**Definition 4.0.10.** The equations (120) are the **error correction conditions**.

In Section 3.1 we looked at some specific examples of quantum error correction codes, in particular we looked at the bit flip algorithm (Algorithm 3.1). We show here how this particular example fits into the general theory presented in this Section.

**Example 4.0.11.** The operator elements in question are $\{I, X_1, X_2, X_3\}$. The appropriate projector $P$ is $P := |000\rangle\langle 000| + |111\rangle\langle 111|$. We let $E_1 = I, E_2 = X_1, E_3 = X_2, E_4 = X_3$ and notice that for $i, j = 1, ..., 4$ we have $PE_i^\dagger E_j P = \delta_{ij} P$. So the identity matrix $I$ can be taken as the appropriate Hermitian operator $A$.

We now run through the proof of 4.0.9 and see how it works in this particular setting. We have that $A = I$ is already diagonal so $F_k = E_k$. Moreover, we have that $\sqrt{PF_k^\dagger F_k P} = \sqrt{PP} = P$ and so the polar decomposition

of $F_kP$ is $F_kP$ (as $F_k = E_k$ is unitary). We thus have:

$$\begin{aligned}
P_1 &= IPI^\dagger = P = |000\rangle\langle 000| + |111\rangle\langle 111| \\
P_2 &= X_1PX_1 = |100\rangle\langle 100| + |011\rangle\langle 011| \\
P_3 &= X_2PX_2 = |010\rangle\langle 010| + |101\rangle\langle 101| \\
P_4 &= X_3PX_3 = |001\rangle\langle 001| + |110\rangle\langle 110|
\end{aligned}$$

Thus $\mathcal{R} = \{IP_1, X_1P_2, X_2P_3, X_3P_4\}$:

$$\begin{aligned}
IP_1 &= PI^\dagger = P = |000\rangle\langle 000| + |111\rangle\langle 111| \\
X_1P_2 &= PX_1 = |000\rangle\langle 100| + |111\rangle\langle 011| \\
X_2P_3 &= PX_2 = |000\rangle\langle 010| + |111\rangle\langle 101| \\
X_3P_4 &= PX_3 = |000\rangle\langle 001| + |111\rangle\langle 110|
\end{aligned}$$

and so

$$\mathcal{R}(\rho) = P_1\rho P_1 + X_1P_2\rho P_2X_1 + X_2P_3\rho P_3X_2 + X_3P_4\rho P_4X_3 \tag{130}$$

As anticipated by Remark 4.0.8, we see that (130) is distinct from Algorithm 3.1.

# 5 Stabilisers

We provide a means for determining when a vector subspace consisting of correctable errors is the largest such. That is, we establish a method for proving that a set of vectors span the codespace of a QECC (Definition 3.0.3).

Throughout, $\mathbb{H}$ denotes a qubit $\mathbb{C}^2$, that is, the complex Hilbert space $\mathbb{C}^2$.

Recall the Pauli operators of Definition 3.1.1.

$$X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad Y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{131}$$

Recall also our notation that, for example, $Z_1Z_2$ on $\mathbb{H}^{\otimes 3}$ denotes the operator $Z \otimes Z \otimes I$, see Notation 3.1.3.

**Definition 5.0.1.** Let $n > 0$. The $n^{\text{th}}$-**Pauli Group**, denoted $G_n$, is the set of operators $\mathbb{H}^{\otimes n} \longrightarrow \mathbb{H}^{\otimes n}$ generated by of all operators $\pm I, iI, X_j, Y_j, Z_j$ for $j = 1, ..., n$.

**Definition 5.0.2.** Given a subgroup $S \subseteq G_n$ of the Pauli group $G_n$, we denote by $V^S$ the subspace of $\mathbb{H}^{\otimes n}$ which is invariant under the operators $S$. That is, $|\psi\rangle \in V^S$ if and only if

$$\forall W \in S, W |\psi\rangle = |\psi\rangle \tag{132}$$

Denote by $\mathscr{X}$ the following Pauli operators

$$\mathscr{X} := \{I, X, Y, Z\} \tag{133}$$

For an arbitrary element $g \in G_n$, let $g_1, ..., g_n \in \mathscr{X}$ be such that

$$g = \alpha g_1 \otimes \ldots \otimes g_n, \quad \alpha \in \{1, -1, i, -i\} \tag{134}$$

then the sequence $g_1, ..., g_n$ is the unique such, and we denote a length $2n$ sequence $x = (x_1, ..., x_{2n})$ in $\mathbb{Z}_2^{2n}$ by $r(g)$ defined by the following schemata:

- $x_i = 1$ if and only if $g_i = X$,
- $x_{i+n} = 1$ if and only if $g_i = Z$,
- $x_i = x_{i+n} = 1$ if and only if $g_i = Y$.

Given a set $\{g_1, ..., g_k\}$ of elements of the Pauli group, the **check matrix** is the $k \times 2n$ matrix whose $j^{\text{th}}$ row is $r(g_j)$. The check matrix is denoted $\text{Check}(g_1, ..., g_k)$.

Let $(g, h)$ be a pair of elements of $G_n$ and let $g_1, ..., g_n, h_1, ..., h_n \in \mathscr{X}$ be such that

$$\begin{aligned} g &= \alpha g_1 \otimes \ldots \otimes g_n, \quad \alpha \in \{1, -1, i, -i\} \\ h &= \beta h_1 \otimes \ldots \otimes h_n, \quad \beta \in \{1, -1, i, -i\} \end{aligned}$$

we see that $g$ and $h$ commute if and only if the number of times $g_j$ and $h_j$ are distinct matrices with neither equal to the identity is even. Defining

$$\Lambda := \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \tag{135}$$

we have the following Lemma.

**Lemma 5.0.3.** *Let $(g_1, g_2) \in G_n$. Then $g_1, g_2$ commute if and only if*

$$r(g_1)\Lambda r(g_2)^T = 0 \tag{136}$$

*Rough sketch.* The form of $r(g_1)$:

$$r(g_1) = \big(X \text{ or } Y \text{ in } g_1 \quad | \quad Z \text{ or } Y \text{ in } g_1\big) \tag{137}$$

and similarly for $r(g_2)$. Thus we have

$$r(g_1)\Lambda r(g_2)^T = \big(X \text{ or } Y \text{ in } g_1 \quad | \quad Z \text{ or } Y \text{ in } g_1\big) \begin{pmatrix} Z \text{ or } Y \text{ in } g_2 \\ X \text{ or } Y \text{ in } g_2 \end{pmatrix} \tag{138}$$

This contains the data of the requirements specified by Observation 5. □

The Check matrix is useful for more:

**Definition 5.0.4.** A set of elements $g_1, ..., g_r \in G_n$ of the Pauli group $G_n$ are **independent** if the for any $j$ we have, where we write $\hat{g}_i$ for the omission of $g_i$:

$$\langle g_1, ..., g_r \rangle \neq \langle g_1, ..., \hat{g}_j, ..., g_n \rangle \tag{139}$$

(here, the notation $\langle g_1, ..., g_n \rangle$ denotes the group generated by these elements).

**Lemma 5.0.5.** *Let $g_1, ..., g_r \in G_n$ be a set of elements such that $-I \notin \langle g_1, ..., g_r \rangle$, then the elements $g_1, ..., g_r$ are independent if and only if $r(g_1), ..., r(g_r)$ and linearly independent (over the field $\mathbb{Z}_2$).*

*Proof.* See [1, Page 457, Proposition 10.3] □

The following Lemma will be used to calculate the dimension of $V^S$:

**Lemma 5.0.6.** *Let $g_1, ..., g_k$ be independent elements of the Pauli group $G_n$ and denote by $S$ the group they generate. Assume $-I \notin S$. Then for each $i = 1, ..., k$ there exists $g \in G_n$ such that $g$ anti-commutes with $g_i$ and commutes with all $g_j$ satisfying $i \neq j$.*

*Proof.* The set $r(g_1), ..., r(g_k)$ is linearly independent by Lemma 5.0.5, thus the check matrix of $g_1, ..., g_k$ has $k$ linearly independent columns. So, there exists a vector $x \in \mathbb{Z}_2^k$ such that

$$\text{Check}(g_1, ..., g_n)\Lambda x = e_i \tag{140}$$

where $e_i$ is the $i^{\text{th}}$ standard basis vector of $\mathbb{Z}_2^k$. Let $g$ be such that $r(g)^T = x$. The result follows from Lemma 5.0.5. □

**Theorem 5.0.7.** *Let $S = \langle g_1, ..., g_k \rangle \subseteq G_n$ and say $-I \notin S$. Then* $\dim V^S = 2^{n-k}$.

*Proof.* We notice that $(1/2)(I + g_j)$ is the projector onto the $+1$-Eigenspace of $g_j$. We let $x = (x_1, ..., x_k) \in \mathbb{Z}_2^k$ and define the operator

$$P_S^x := 1/2^k \prod_{j=1}^{k} (I + (-1)^{x_j} g_j) \tag{141}$$

By Lemma 5.0.6 we have for each $g_j$ there exists $g_{x_j}$ such that $g_{x_j} g_j g_{x_j}^{-1} = -g_j$. Let $g_x = g_{x_1} \dots g_{x_k}$, then

$$\begin{aligned} g_x P_S^{(0,...,0)} g_x^{-1} &= 1/2^k \prod_{j=1}^{k} (g_{x_j} g_{x_j}^{-1} + g_{x_j} g_j g_{x_j}^{-1}) \\ &= P_S^x \end{aligned}$$

Thus there is an isomorphism

$$\operatorname{im} P_S^x \cong \operatorname{im} P_S^{(0,...,0)} \tag{142}$$

Since $\operatorname{im} P_S \cong V_S$ we have $\dim \operatorname{im} P_S^x = \dim V_S$. Finally we note that

$$I = \sum_{x \in \mathbb{Z}_2^k} P_S^x \tag{143}$$

The operator $I$ is a projector onto an $n$-dimensional space, and $\sum_{x \in \mathbb{Z}_2^k} P_S^x$ is a sum of $2^k$ orthogonal projectors all of the same dimension as $V_S$, thus the only possibility is $\dim V_S = 2^{n-k}$. □

In the context of the bitflip error correction, we have:

$$S = \langle Z_1 Z_2, Z_2 Z_3 \rangle \subseteq G_3 \tag{144}$$

It is clear that

$$V^S \supseteq \operatorname{Span}\{|000\rangle, |111\rangle\} \tag{145}$$

Now we want to use Theomre 5.0.7 to prove that in fact (145) holds up to equality.

Since $Z_1 Z_2, Z_2 Z_3$ are 2 independent generators for $S$, it follows from Theorem 5.0.7 that

$$\dim V^S = 2^{3-2} = 2 = \dim \big( \operatorname{Span}\{|000\rangle, |111\rangle\} \big) \tag{146}$$

# A Operator Theory

## A.1 Adjoint operators

We will be chiefly concerned with the Hilbert space $\ell^2$ but we work in a more general setting for now. A *Hilbert space* will always mean over $\mathbb{C}$. Associated to every operator between Hilbert spaces is an operator between their *dual spaces*:

In general, if $\mathcal{I}$ is any inner product space over $\mathbb{C}$ and we have two vectors $x, y \in I$ then we can consider the projection of $y$ onto $x$ which is given by

$$\text{Proj}_y(x) := \frac{\langle x, y\rangle}{||y||}\frac{y}{||y||} \tag{147}$$

Thus, if $U \subseteq \mathcal{I}$ is a one dimensional subspace spanned by a unit vector $u \in U$ then the projection of any $x \in \mathcal{I}$ onto $u$ is given by the simple formula $\langle x, u\rangle u$. The following Lemma shows what we can say when the subspace is of arbitrary dimension but with $U$ closed:

**Lemma A.1.1.** *Let $\mathbb{H}$ be a Hilbert space and $U \subseteq \mathbb{H}$ a closed subspace. Then*

$$\mathbb{H} = U \oplus U^{\perp}$$

*Proof.* We will define a projection

$$\begin{aligned} P_U : \mathbb{H} &\longrightarrow U \\ x &\longmapsto \inf\{||x - y|| \mid y \in U\} \end{aligned}$$

We let $d$ denote $\inf\{||x - y|| \mid y \in U\}$. By definition of inf there exists a sequence $(x_n)_{n=0}^{\infty}$ of elements in $U$ such that $\lim_{n\to\infty} ||x - x_n|| = d$. Since $U$ is closed it is complete and the norm is continuous so it suffices to show that the sequence $(x_n)_{n=0}^{\infty}$ is Cauchy. This can be done for example using the parallelogram identity: for all $n, m \geq 0$:

$$||x_n - x_m||^2 + ||(x - x_n) + (x - x_m)||^2 = 2||x - x_n||^2 + 2||x - x_m||^2 \tag{148}$$

As given $\epsilon > 0$ there exists $N \geq 0$ such that $||x - x_n||^2 < d^2 + \epsilon^2/4$, for $n \geq N$. Thus

$$\begin{aligned} ||x_n - x_m||^2 &= 2||x - x_n||^2 + 2||x - x_m||^2 - 4||x = ||1/2(x_n + x_m)||^2 \\ &\leq 4d^2 + \epsilon^2 - 4||x - 1/2(x_n + x_m)||^2 \end{aligned}$$

which since $1/2(x_n + x_m) \in C$ we have $d \leq ||x - 1/2(x_n + x_m)||$, proving $(x_n)_{n=0}^{\infty}$ is Cauchy. This also shows linearity.

It remains to show $x - P_U(x) \in U^{\perp}$. To do this, we will consider the family of vectors $c(t) = (1-t)P_U(x) + ty$, $(t \in \mathbb{R})$ and analyse the derivative of $||x - y_t||^2$ at $t = 0$.

Consider the composition

$$\gamma : \mathbb{R} \longrightarrow \mathbb{R} \tag{149}$$

$$t \longmapsto ||x - c(t)||^2 \tag{150}$$

We can write $\gamma$ in a more explicit form:

$$\begin{aligned}\gamma(t) &= ||x - P_U(x) + t(y - P_U(x))||^2 \\ &= \big\langle x - P_U(x) + t(y - P_U(x)), x - P_U(x) + t(y - P_U(x)) \big\rangle \\ &= ||x - P_U(x)||^2 - 2t\,\mathrm{Re}\langle x - P_U(x), y - P_U(x)\rangle + t^2||y - P_U(x)||\end{aligned}$$

which is clearly differentiable and has derivative $-2\,\mathrm{Re}\langle x - P_U(x), y - P_U(x)\rangle$ at $t = 0$. Since $P_U(x)$ (which equals $c(0)$) is a minimum of $\gamma(t)$ we have that $\mathrm{Re}\langle x - P_U(x), y - P_U(x)\rangle = 0$. This holds true for arbitrary $y \in U$ and lastly we have

$$\{y - P_U(x) \mid y \in U\} = U$$

thus for all $y \in U$:

$$\mathrm{Re}\langle x - P_U(x), y\rangle = 0 \tag{151}$$

This shows that $x - P_U(x) \in U^{\perp}$. □

Given a Hilbert space $\mathbb{H}$ there is a map

$$\Phi : \mathbb{H} \longrightarrow \mathbb{H}^* \tag{152}$$

$$b \longmapsto \langle \_, b\rangle \tag{153}$$

Notice that in order to produce a *linear* functional, it was important we put $b$ in the second argument, we must define $\Phi$ so that $\Phi(b) \neq \langle b, \_\rangle$. By anti-linearity of the second argument of the inner product we have that $\Phi$ is anti-linear, and moreover is injective as

$$\begin{aligned}\Phi(b) = \Phi(b') &\Longrightarrow \langle \_, b\rangle = \langle \_, b'\rangle \\ &\Longrightarrow \forall b'' \in \mathbb{H}, \langle b'', b - b'\rangle = 0 \\ &\Longrightarrow \text{ in particular, } \langle b - b', b - b'\rangle = 0 \\ &\Longrightarrow b - b' = 0\end{aligned}$$

In the special case where $\mathbb{H}$ is finite dimensional, we automatically have that this map is surjective as it is injective, and any anti-linear, injective map between two finite dimensional spaces of equal dimension is automatically surjective. More generally, if $\mathbb{H}$ has arbitrary dimension, then for any $y \in \mathbb{H}$ the map $\langle \_, y \rangle$ is bounded (see Remark A.1.5) so the image of $\Phi$ is contained in the set of continuous linear functionals, the following establishes the reverse inequality:

**Theorem A.1.2** (Riesz Representation Theorem)**.** *Let $\mathbb{H}$ be a Hilbert space. For every continuous linear functional $\varphi \in \mathbb{H}^*$ there exists a unique element $h_\varphi \in \mathbb{H}$ such that*

$$\varphi = \langle \_, h_\varphi \rangle \tag{154}$$

*Moreover, we have*

$$||\varphi||_{\mathbb{H}^*} = ||h_\varphi||_{\mathbb{H}} \tag{155}$$

We will use the following Lemma:

**Lemma A.1.3.** *Let $\mathbb{H}$ be a Hilbert space and $\varphi \in \mathbb{H}^*$ be non-zero and continuous. Then $(\ker \varphi)^\perp$ is one dimensional.*

*Proof.* Since $\varphi$ is continuous the set $\ker \varphi$ is closed and so by Lemma A.1.1 we have $\mathbb{H} = \ker \varphi \oplus (\ker \varphi)^\perp$, which since $\varphi \neq 0$ implies there exists $v \neq 0 \in (\ker \varphi)^\perp$, so $\dim(\ker \varphi)^\perp > 0$. Now, say $v_1, v_2 \in (\ker \varphi)^\perp$ so that $\varphi(v_1) \neq 0$ and $\varphi(v_2) \neq 0$. These are complex numbers and so there exists $\lambda \in \mathbb{C}$ such that

$$0 = \lambda\varphi(v_1) - \varphi(v_2) = \varphi(\lambda v_1 - v_2)$$

which means $\lambda v_1 - v_2 \in \ker \varphi \cap (\ker \varphi)^\perp = \{0\}$. □

*Proof of Theorem A.1.2.* Clearly if $\ker \varphi = \mathbb{H}$ we can take $h_\varphi = 0$ so assume this is not the case. Since $\varphi$ is continuous its kernel $\ker \varphi$ is a closed subset of $\mathbb{H}$. Thus, by Lemma A.1.1 the Hilbert space $\mathbb{H}$ decomposes: $\mathbb{H} = \ker \varphi \oplus (\ker \varphi)^*$. Since $\ker \varphi$ is a proper subset it then follows that there exists a non-zero element $v \neq 0 \in (\ker \varphi)^*$, by normalising we may assume that $v$ is a unit vector. We will show that $\overline{\varphi(v)}v$ is the appropriate unique choice for $h_\varphi$.

By Lemma A.1.3 the subspace $(\ker \varphi)^\perp$ is one dimensional, hence we can use formula (147) for the projection of arbitrary $x$ onto $(\ker \varphi)^\perp$. Observe

the following calculation:

$$\begin{aligned}\varphi(x) &= \varphi(x - \langle x, v\rangle v + \langle x, v\rangle v) \\ &= \varphi(x - \langle x, v\rangle v) + \varphi(\langle x, v\rangle v) \\ &= 0 + \langle x, v\rangle \varphi(v) \\ &= \langle x, \overline{\varphi(v)} v\rangle\end{aligned}$$

For uniqueness, say $h'_\varphi$ was another such element. Then

$$\begin{aligned}&\forall x \in \mathbb{H}, \langle x, h_\varphi\rangle = \langle x, h'_\varphi\rangle \\ &\implies \forall x \in \mathbb{H}, \langle x, h_\varphi - h'_\varphi\rangle = 0 \\ &\implies ||h_\varphi - h'_\varphi|| = 0 \\ &\implies h_\varphi = h'_\varphi\end{aligned}$$

For the second claim, we use the Cauchy-Schwartz inequality:

$$|\varphi(x)| = |\langle x, \overline{\varphi(v)} v\rangle| \leq ||x||||\overline{\varphi(v)}||v|| = ||x|||\varphi(v)|$$

and so if $x$ has unit norm $|\varphi(x)| \leq |\varphi(v)|$, in other words, $||\varphi||_{\mathbb{H}^*} \leq |\varphi(v)|$ however $v$ has unit norm itself, so $||\varphi||_{\mathbb{H}^*} = |\varphi(v)|$. The proof is now complete once it is noted that $||h_\varphi||_{\mathbb{H}} = |\varphi(v)|$. □

**Corollary A.1.4.** *There existgs an antilinear, isometric injection:*

$$\mathbb{H} \longrightarrow \mathbb{H}^* \tag{156}$$

$$v \longmapsto \langle v, \_\rangle \tag{157}$$

*and hence a bijection when $\mathbb{H}$ is finite dimensional.*

**Remark A.1.5.** Let $y \in \mathbb{H}$ be an element of a Hilbert space $\mathbb{H}$ and consider the function $\langle \_, y\rangle$. This is bounded, as by Cauchy-Schwartz:

$$|\langle x, y\rangle| \leq ||x||||y||$$

thus $|\langle \_, y\rangle|/||x|| \leq ||y||$ and in fact this is equality as $|\langle y/||y||, y\rangle| = ||y||$.

Given an operator $u : \mathbb{H}_1 \longrightarrow \mathbb{H}_2$ there is for each $y \in \mathbb{H}_2$ an associated linear functional $x \longmapsto \langle u(x), y\rangle$ which we denote by $\langle u(\_), y\rangle$. By Theorem

A.1.2 there is thus an element $y^* \in \mathbb{H}_1$ such that $\langle u(\_), y\rangle = \langle \_, y^*\rangle$. The assignment $y \mapsto y^*$ is in fact linear, we show additivity:

$$\langle u(\_), y_1 + y_2\rangle = \langle \_, (y_1 + y_2)^*\rangle$$

and

$$\begin{aligned}\langle u(\_), y_1 + y_2\rangle &= \langle u(\_), y_1\rangle + \langle u(\_), y_2\rangle \\ &= \langle \_, y_1^*\rangle + \langle \_, y_2^*\rangle \\ &= \langle \_, y_1^* + y_2^*\rangle\end{aligned}$$

which implies $(y_1 + y_2)^* = y_1^* + y_2^*$. We define:

**Definition A.1.6.** The **adjoint operator** associated to an operator $u : \mathbb{H}_1 \longrightarrow \mathbb{H}_2$ is the linear map:

$$\begin{aligned}u^* : \mathbb{H}_2 &\longrightarrow \mathbb{H}_1 \\ y &\longmapsto y^*\end{aligned}$$

Its existence is established by the Riesz Representation Theorem (A.1.2) and it is uniquely determined by the property:

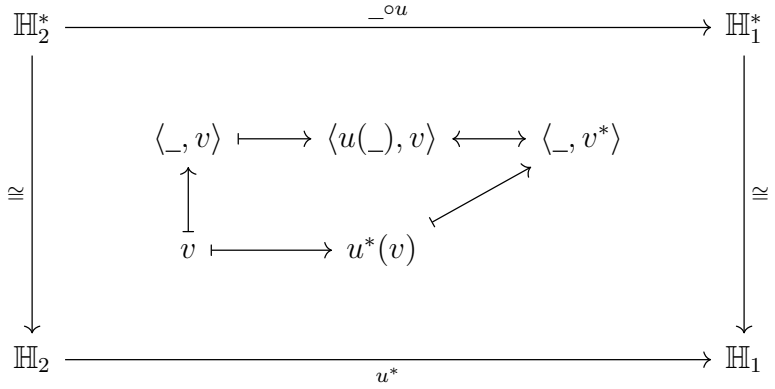
$$\forall x \in \mathbb{H}_1, y \in \mathbb{H}_2, \langle u(x), y\rangle = \langle x, u^*(y)\rangle \tag{158}$$

**Remark A.1.7.** Let $\big(\mathbb{H}_1, \langle\cdot,\cdot\rangle_B\big), \big(\mathbb{H}_2, \langle\cdot,\cdot\rangle\big)_C$ be Hilbert spaces and let $u : \mathbb{H}_1 \longrightarrow \mathbb{H}_2$ be an operator. The **adjoint** to $u$, denoted $u^*$ is the operator:

$$\_ \circ u : \mathbb{H}_2^* \longrightarrow \mathbb{H}_1^* \tag{159}$$

$$\varphi \longmapsto \varphi \circ u \tag{160}$$

The following diagram commutes:




(161)

which explains the overloading of terminology.

**Notation A.1.8.** Given a complex matrix $A$, the matrix given by conjugating each element $a \in A$ and then transposing the result, ie, the **conjugate transpose** is denoted $A^\dagger$. Due to Proposition A.1.9 below, the conjugate transpose of a matrix is often referred to as the **adjoint**.

**Proposition A.1.9.** *Let $\mathbb{H}_1, \mathbb{H}_2$ be finite dimensional, and let $v_1, ..., v_n \in \mathbb{H}_1, w_1, ..., w_m \in \mathbb{H}_2$ be orthonormal bases for $\mathbb{H}_1, \mathbb{H}_2$ respectively. If $\varphi : \mathbb{H}_1 \longrightarrow \mathbb{H}_2$ is a linear transformation and $A$ its matrix representation with respect to these bases, then the matrix representation of the adjoint $\varphi^*$ is $A^\dagger$, the conjugate transpose of $A$.*

*Proof.* ' For each $j = 1, ..., m$ write $w_j^* = \alpha_1 v_1 + \ldots + \alpha_n v_n$ and each $i = 1, ..., n$ write $\varphi(v_i) = \beta_1 w_1 + \ldots + \beta_m w_m$. We calculate:

$$\langle \varphi(v_i), w_j \rangle = \beta_m \langle w_1, w_j \rangle + \ldots + \beta_m \langle w_m, w_j \rangle = \beta_j \tag{162}$$

and

$$\langle v_i, w_j^* \rangle = \bar{\alpha}_1 \langle v_i, v_1 \rangle + \ldots + \bar{\alpha}_n \langle v_i, v_n \rangle = \bar{\alpha}_i \tag{163}$$

Since by definition $\langle \varphi(v_i), w_j \rangle = \langle v_i, w_j^* \rangle$ the proof is complete. □

## A.2 Hermitian and unitary operators

Throughout, $V$ is a complex vector space.

**Definition A.2.1.** A square, complex matrix $A$ is **Hermitian** if it is self-adjoint, that is $A^\dagger = A$.

A matrix is **normal** if $AA^\dagger = A^\dagger A$

An operator $\varphi : V \longrightarrow V$ is **Hermitian** (**normal**) if a (and hence all) matrix representation(s) of $V$ is Hermitian (normal).

Clearly, all Hermitian matrices are normal.

**Theorem A.2.2** (Spectral decomposition)**.** *Let $V$ be a finite dimensional complex inner product space and $A$ a matrix representation of an operator on $V$. The matrix $A$ is normal if and only if it is diagonalisable with respect to some orthonormal basis for $V$.*

*Proof.* We prove that normal matrices are diagonalisable.

We proceed by induction on the size of the matrix. If the matrix is $1 \times 1$ then there is nothing to prove. Now for the inductive step. Let $\lambda$ be an

eigenvalue of $A$, and $P$ the matrix which projects onto the $\lambda$-eigenspace. We let $Q$ denote $I - P$, the projector onto the complement subspace. We notice that

$$A = (P+Q)A(P+Q) = PAP + QAP + PAQ + QAQ \tag{164}$$

We have that $QAP = 0$ because $A$ maps the $\lambda$-eigenspace onto itself, and we claim moreover that $PAQ = 0$. To see this, let $v$ be an eigenvector with eigenvalue $\lambda$, then

$$AA^\dagger v = A^\dagger A v = A^\dagger \lambda v = \lambda A^\dagger v \tag{165}$$

which means $A^\dagger$ maps the $\lambda$-eigenspace onto itself. This implies $QA^\dagger P = 0$, taking the transpose of which we end at $PAQ = 0$ as claimed.

Thus $A = PAP + QAQ$. The matrix $PAP$ is diagonalisable with respect to some orthonormal basis for $P$. Since $P \cap Q = 0$ it remains to show that $QAQ$ is diagonalisable with respect to some orthonormal basis for $Q$. The space $Q$ has strictly smaller size than $A$ and so this follows by induction once we have shown that $QAQ$ is normal. This is a simple calculation:

$$\begin{aligned} QAQQA^\dagger Q &= QAQA^\dagger Q \\ &= QA(P+Q)A^\dagger Q \\ &= QAA^\dagger Q \\ &= QA^\dagger AQ \\ &= QA^\dagger (P+Q)AQ \\ &= QA^\dagger QAQ \\ &= QA^\dagger QQAQ \end{aligned}$$

□

**Definition A.2.3.** Let $\mathbb{H}$ be a possibly inifinite dimensional Hilbert space, an operator $U : \mathbb{H} \longrightarrow \mathbb{H}$ is **unitary** if $U^\dagger U = UU^\dagger = \mathrm{Id}_n$.

**Definition A.2.4.** A matrix $U$ is **unitary** if $U^\dagger U = I$.

**Lemma A.2.5.** *A square, unitary matrix $U$ satisfies $UU^\dagger = I$.*

*Proof.* Let $u_{ij}$ denote the entry of $U$ in row $i$ and column $j$. The entry in row $i$ and column $j$ of $U^\dagger U$ is $\sum_{k=1}^n \overline{u}_{ik} u_{kj}$ which by hypothesis is equal to $\delta_{ij}$. Hence, $\sum_{k=1}^n \overline{u}_{ki} u_{kj}$ is equal to $\sum_{k=1}^n u_{ik} \overline{u}_{jk}$ which is the entry in row $i$ and column $j$ of $UU^\dagger$. □

**Corollary A.2.6.** *If $\mathbb{H}$ is a finite dimensional Hilbert space and $U : \mathbb{H} \longrightarrow \mathbb{H}$ is an operator on $\mathbb{H}$, then $U$ is unitary if and only if for all $u, v \in \mathbb{H}$ we have $\langle Uu, Uv\rangle = \langle u, v\rangle$.*

*Proof.* First we observe the following calculation, where $u \in \mathbb{H}$ is arbitary.

$$\begin{aligned}
||U^\dagger Uu - u|| &= \langle U^\dagger Uu - u, U^\dagger Uu - u\rangle \\
&= \langle U^\dagger Uu, U^\dagger Uu\rangle - \langle U^\dagger Uu, u\rangle - \langle u, U^\dagger Uu\rangle + \langle u, u\rangle \\
&= \langle UU^\dagger Uu, Uu\rangle - \langle Uu, Uu\rangle - \langle Uu, Uu\rangle + \langle u, u\rangle \\
&= \langle U^\dagger Uu, u\rangle - \langle u, u\rangle - \langle u, u\rangle + \langle u, u\rangle \\
&= \langle Uu, Uu\rangle - \langle u, u\rangle \\
&= \langle u, u\rangle - \langle u, u\rangle \\
&= 0
\end{aligned}$$

Hence $U^\dagger Uu = u$ for all $u \in \mathbb{H}$ and so $U^\dagger U = \mathrm{Id}_{\mathbb{H}}$.

Let $u_1, ..., u_n$ be an orthonormal basis for $\mathbb{H}$ and let $\underline{U}$ denote the matrix of $U$ written with respect to this basis. Since $U$ is unitary we have that $\underline{U}$ is unitary and so $\underline{U}^\dagger \underline{U} = I$ and by Lemma A.2.5 we have $\underline{U}\underline{U}^\dagger = I$. It follows from this that $UU^\dagger = \mathrm{Id}_{\mathbb{H}}$ and so $U$ is unitary.

The converse is obvious. □

In fact, it is sufficient to check even less.

**Lemma A.2.7.** *Let $U : \mathbb{H} \longrightarrow \mathbb{H}$ be an operator on a finite dimensional Hilbert space. If $\langle Uu, Uu\rangle = \langle u, u\rangle$ for all $u \in \mathbb{H}$, then for all $u, v \in \mathbb{H}$ we have $\langle Uu, Uv\rangle = \langle u, v\rangle$.*

*Proof.* It suffices to prove that if $C : \mathbb{H} \longrightarrow \mathbb{H}$ is an operator on $\mathbb{H}$ such that for all $x \in \mathbb{H}$ we have $\langle Cx, x\rangle = 0$ then $C = 0$.

We let $x, y \in \mathbb{H}$ be arbitrary and consider $\langle C(x + y), x + y\rangle$. Since this is 0 it follows that $\langle Cx, y\rangle = -\langle Cy, x\rangle$. On the other hand, $\langle C(x + iy), x + iy\rangle$ is also 0, which implies $\langle Cx, y\rangle = \langle Cx, y\rangle$. Hence $\langle Cx, y\rangle = \langle Cy, x\rangle = 0$. □

**Corollary A.2.8.** *If $U : \mathbb{H} \longrightarrow \mathbb{H}$ is an operator and $\mathbb{H}$ is finite dimensional, then $U$ is unitary if and only if $\forall u \in \mathbb{H}, \langle Uu, Uu\rangle = \langle u, u\rangle$.*

*Proof.* Immediate from Corollary A.2.6 and Lemma A.2.7. □

Notice that the spectral decomposition (A.2.2) states that the matrix $A$ is such that $A = U^\dagger DU$ for a diagonal matrix $D$ and a unitary matrix $U$.

**Corollary A.2.9.** *A normal matrix $A$ is Hermitian if and only if its eigenvalues are real.*

*Proof.* First notice that if a matrix is Hermitian then for any eigenvector $v$ with eigenvalue $\lambda$:

$$\lambda|v|^2 = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \bar{\lambda}|v|^2 \tag{166}$$

Now we prove the other direction. Let $D$ be diagonal and $U$ a unitary matrix such that $A = U^{-1}DU$. Then

$$A^\dagger = U^\dagger D^\dagger U^{-1^\dagger} = U^{-1}DU = A \tag{167}$$

□

**Definition A.2.10.** An operator $\varphi : V \longrightarrow V$ is **positive** if:

$$\forall v \in V, \langle v, \varphi v\rangle \geq 0 \tag{168}$$

which means, $\langle v, \varphi v\rangle$ is real and non-negative. If the inequality is strict, then $\varphi$ is **positive definite**.

**Example A.2.11.** Let $A$ be any operator. Then for any $v \in V$:

$$\langle v, A^\dagger Av\rangle = \langle Av, Av\rangle = ||Av||^2 \geq 0 \tag{169}$$

Thus $A^\dagger A$ is positive.

**Proposition A.2.12.** *A positive operator on a finite dimensional vector space is necessarily Hermitian.*

*Proof.* Let $A$ be a matrix representation of the positive operator. Notice the following calculation:

$$\begin{aligned} 0 \leq \langle v, (A - A^\dagger)v\rangle &= \langle (A^\dagger - A)v, v\rangle \\ &= \overline{\langle v, (A^\dagger - A)v\rangle} \\ &= \langle v, (A^\dagger - A)v\rangle \\ &= -\langle v, (A - A^\dagger)v\rangle \geq 0 \end{aligned}$$

and so for all $v \in V$ we have $\langle v, (A - A^\dagger)v\rangle = 0$.

Moreover, we notice that $A - A^\dagger$ is normal and hence diagonalisable, by the Spectral decomposition. It follows from these two observations that $A - A^\dagger = 0$. □

**Definition A.2.13.** Let $A, B$ be matrices, then the **commutator** is $[A, B] := AB - BA$. The **anticommutator** is $\{A, B\} = AB + BA$.

**Theorem A.2.14** (Simultaneous Diagonalisation Theorem)**.** *Let $A, B$ be Hermitian operators. Then $[A, B] = 0$ if and only if $A$ and $B$ are simultaneously diagonalisable.*

*Proof.* If $A$ and $B$ are simultaneously diagonalisable, then let $U$ be a unitary matrix and $D_1, D_2$ diagonal matrices such that

$$A = U^{-1} D_1 U, \qquad B = U^{-1} D_2 U \tag{170}$$

We then have:

$$\begin{aligned} AB &= U^{-1} D_1 U U^{-1} D_2 U \\ &= U^{-1} D_1 D_2 U \\ &= U^{-1} D_2 D_1 U \\ &= U^{-1} D_2 U U^{-1} D_1 U \\ &= BA \end{aligned}$$

Conversely, say $[A, B] = 0$. We have that $A$ is Hermitian and so admits a spectral decomposition. Let $a_1, ..., a_n$ be the eigenvalues corresponding to this decomposition and let $V_{a_i}$ denote the $a_i$-eigenspace. We first notice that $B$ maps $V_{a_i}$ into itself: for any $v \in V_{a_i}$

$$ABv = BAv = a_i Bv \tag{171}$$

Now, since $B$ is Hermitian, it follows that $B_{V_{a_i}} : V_{a_i} \longrightarrow V_{a_i}$ is and so there exists a spectral decomposition of $B_{V_{a_i}}$ for each vector space $V_{a_i}$. Denote by $b_1^{a_i}, ..., b_{k_{a_i}}^{a_i}$ an orthonormal basis for $V_{a_i}$. We then have that

$$\{b_1^{a_i}, ..., b_{k_{a_i}}^{a_i}\}_{i=1}^{n} \tag{172}$$

is a basis of eigenvectors of both $A$ and $B$ for the whole space $V$. □

There is another decomposition which is often helpful: Let $T : V \longrightarrow V$ be a linear operator on a finite dimensional vector space $V$. We could ask if $T$ can be factored $T = UT'$ where $U$ is unitary? Say this was possible, then

$$T^\dagger T = T'^\dagger U^\dagger U T' \tag{173}$$

so if $T'$ were Hermitian we would have $T^\dagger T = T'^2$ which would imply $T' = \sqrt{T^\dagger T}$, in fact $T^\dagger T$ is Hermitian (indeed it is positive) and thus so is $\sqrt{T^\dagger T}$ and so our assumption that $T'$ be Hermitian is not too much to ask for, and if $U$ were to exist it must be that $T' = \sqrt{T^\dagger T}$. Thus we are prompted to make the following calculation: let $v_1, ..., v_n$ be a basis for $V$ such that (we write $P_{v_i}$ for the projection onto $v_i$)

$$\sqrt{T^\dagger T} = \sum_{i=1}^{n} \lambda_i P_{v_i} \tag{174}$$

then

$$\sqrt{T^\dagger T} v_i \lambda_i \tag{175}$$

and indeed we want $U$ such that $\lambda_i U v_i = T v_i$. One might suggest defining $U v_i = T v_i / \lambda_i$ at this point, however there is no reason for this to be unitary. Instead we define

$$U = \sum_{j=1}^{n} T v_j P_{v_j} / \sqrt{\lambda_j} \tag{176}$$

which indeed is unitary. In fact we read off from this that $\{Tv_1/\sqrt{\lambda_1}, ..., Tv_n/\sqrt{\lambda_n}\}$ is an orthonormal basis for $V$. Notice however that this assumes $\lambda_i \neq 0$ for all $i$. This can be fixed by doing this process first for all $\lambda_i \neq 0$, and to construct an orthonormal set $\{Tv_1/\sqrt{\lambda_1}, ..., Tv_j/\sqrt{\lambda_j}\}$ and then extending this to an orthonormal basis for $V$ via the Gram-Schmidt process.

We have proven the first half of:

**Theorem A.2.15** (Polar decomposition)**.** *Let $T : V \longrightarrow V$ be a linear operator on an $n$-dimensional vector space $V$. Then there exists a unitary operator $U$ and positive operators $J, K$ such that*

$$T = UJ = KU \tag{177}$$

*with $J = \sqrt{T^\dagger T}, K = \sqrt{T T^\dagger}$.*

To obtain $K$ we simply notice

$$A = JU = UJU^\dagger U \tag{178}$$

so we set $K = UJU^\dagger$, which is a positive operator. Then $AA^\dagger = KUU^\dagger K = K^2$. If we have such a decomposition $T = UJ$, then $J$ is diagonalisable, being positive, thus $T = USDS^\dagger$ for unitary $S$ and diagonal $D$. Setting $V = S^\dagger$ we obtain:

**Corollary A.2.16** (Singular value decomposition)**.** *Let $T : V \longrightarrow V$ be a linear operator on an n-dimensional vector space, then there exists unitary operators $U, V$ and a diagonal operator $D$ such that*

$$T = UDV \tag{179}$$

**Remark A.2.17.** We make a remark on notation. Given a vector $v \in \mathbb{H}$ in some Hilbert space $\mathbb{H}$ (which we assume to be finite dimensional for simplicity), the linear functional which we have been notating as $\langle v, \_ \rangle$ can also be written simply as $\langle v|$. Symmetrically, the vector $v$ can be identified with the linear map $k \longrightarrow \mathbb{H}$ sending $1 \longmapsto v$, we notate this map by $|v\rangle$. Hence, given two vectors $v, u \in V$, the notation $\langle v| \, |u\rangle$ denotes the linear map $k \longrightarrow k$ sending $1 \longmapsto \langle v, u \rangle$. We now describe how some of the concepts introduced in this Section and the last are written using this notation.

- The linear map given in Corollary A.1.4 can be written as $|v\rangle \longmapsto \langle v|$.
- Let $U : \mathbb{H} \longrightarrow \mathbb{H}$ be an operator. We have for any $v \in \mathbb{H}$ that:

$$\langle Uv| = \langle Uv, \_ \rangle = \langle v, U^\dagger \_ \rangle = \langle v| \, U^\dagger \tag{180}$$

  Hence, in light of Corollary A.2.6 we have that $U$ is unitary if and only if for all $v \in \mathbb{H}$ we have $\langle v| \, U^\dagger U \, |v\rangle = \langle v| \, |v\rangle$. This is the condition which is checked throughout the body of this paper.

# References


[1] M. Nielsen, I. Chuang *Quantum Computation and Quantum Information tenth anniversary* Cambridge University Press, 09 December 2010

[2] W. Troiani *Elementary Commutative Algebra* `https://williamtroiani.github.io/pdfs/CommutativeAlgebraWithLin.pdf`

[3] J. Baez. `https://math.ucr.edu/home/baez/quantum/node4.html`