Isoclinic Subspaces and Quantum Error Correction

by
David Mammarella

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Masters of Science
in
Mathematics

Guelph, Ontario, Canada

ABSTRACT

ISOCLINIC SUBSPACES AND QUANTUM ERROR CORRECTION

Advisor:

David Mammarella — Dr. D.W. Kribs

University of Guelph, 2020 — Dr. R. Pereira

This thesis studies the classical notion of canonical angles to explore isoclinic subspaces on a complex inner product space and equivalent conditions are developed for a set of subspaces to be isoclinic. A connection between isoclinic subspaces and quantum error correction will be identified. We will show that every quantum error correcting code is associated with a family of isoclinic subspaces and a partial converse for pairs of such subspaces will be proved. It will also be shown how the canonical angles for isoclinic subspaces arise in the structure of the higher rank numerical ranges of the corresponding orthogonal projections. An examination of how this connection could be used to fuel other ideas in quantum error correction and quantum information theory in general will be discussed to conclude this work.

Dedication

To my parents, John and Emma and my sister, Julia. You are my greatest inspirations and the reason I strive to excel.

Acknowledgements

I want to express my appreciation and gratitude to Dr. Pereira and Dr. Kribs. Your knowledge, guidance and expertise is an inspiration to my academic endeavours and something I will carry with me and remember forever.

To my office mates Alireza, Katrina and especially Eric. I would not want to work beside anyone else day in and day out.

Finally, to my housemates at 93 Moss Pl, I could not ask for a better group of brothers.

Contents

List of Figures

Chapter 1

Introduction

Although quantum computation is studied by Physicists, Computer Scientists and Mathematicians, this work will focus on the mathematics. Quantum error correction is a focal point of research within quantum information and computation. A major breakthrough in the mathematical theory of quantum error correction was the quantum error correction conditions, more specifically the Knill-Laflamme conditions derived in [26]. The Knill-Laflamme conditions give a way to determine if an error correcting code is correctable for an arbitrary quantum channel. The error operators of a quantum channel act on a larger Hilbert space that contains the code subspace. The Knill-Laflamme conditions show if a code is correctable or not by testing if the subspaces corresponding to the error operators are proportional to a unitary transformation on the code space with a certain orthogonality condition. In particular, the ranges of the code subspace corresponding to different error operators must be orthogonal to each other in order to be distinguishable. This in turn allows us to discover the error and further how to correct the error.

The work presented in this thesis takes a slightly different approach to exploring the Knill-Laflamme conditions. Precisely, we show the conditions are linked with a different subject in linear algebra, namely isoclinic subspaces. We can calculate the angle between different subspaces of a larger vector space. If all of the angles between two subspaces are equal to each other then we call these subspaces isoclinic to each other. This was originally formalized by Wong in [51]. Isoclinicity is a subject that is not as well-known, but useful when studying the geometry of subspaces in $\mathbb{R}^n$. In this work we go further and discuss isoclinicity primarily in the complex spaces $\mathbb{C}^n$. The connection between the two fields stems from the use of subspace geometry in quantum error correction. It is a natural step to describe quantum error correction with a more calculated method such as angles between these subspaces. Further, this work will show when isoclinicity arises from the range of the code space corresponding to arbitrary error operators of a correctable code.

The thesis is organized as follows: In chapter 2, we give relevant mathematical preliminaries that will be important in understanding the rest of this thesis. Hilbert spaces are discussed, as well as the different operators that can act on Hilbert spaces. Other concepts relevant will be discussed such as tensor products and an introduction to quantum information theory. In chapter 3, a detailed description of quantum error correction is given beginning with classical error correction. We will state the quantum error correction conditions with a full proof. Lastly, this chapter will conclude with a description of the higher rank numerical range and its relevance to quantum error correction. Chapter 4 begins with canonical angles (angles between subspaces) then progresses towards a full definition of isoclinic subspaces that will be used throughout the rest of the work. Additionally, isoclinic subspaces are explored in greater depth and equivalent characterizations are given from other authors. An explicit equation is derived that determines if two subspaces are in fact isoclinic. Lastly, in chapter 5 a connection between quantum error correction and isoclinic subspaces is presented. This chapter will give equivalent conditions in which subspaces are isoclinic and state when correctable models determine isoclinic subspaces. This chapter provides some examples to build intuition on this connection as well. Discussions on higher rank numerical ranges and their connections with isoclinic subspaces are also included. To end chapter 5 and the thesis, a conclusion and further works section is presented. The presentation of material covered in chapters 2, 3, and 4 is motivated in part by the following references [54, 8, 19, 43, 28, 24, 22, 7, 50, 23, 53, 1, 42, 51, 52, 10, 40, 39, 37, 18, 41, 26, 30, 4, 45, 25, 46, 47, 5, 17, 32, 27, 14, 12, 11, 33, 34, 6, 29]. The results of chapter 5 are contained in the paper [31].

Chapter 2

Background of Quantum Information and Matrix Theory

This chapter will cover the background material needed for this thesis. The next section will discuss operators in Hilbert space and matrix decompositions that will be useful in this thesis. The third section is a review of tensor products, an operation often used in quantum information. The last section will be a brief review of quantum information. We will focus on the concepts necessary for the thesis starting from qubits, mentioning Dirac Notation, and quantum channels.

2.1 Hilbert Space

Hilbert spaces are important for the mathematics of quantum and quantum information theory. We will begin with inner product spaces.

Definition 1. Let X be a complex vector space. An *inner product on* X is a function $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{C}$ such that for all $x, y, z \in X$ and $\alpha, \beta \in \mathbb{C}$ the following properties are satisfied,

1. $\langle x, x \rangle \geq 0$,

2. $\langle x, x\rangle = 0$ if and only if $x = 0$,

3. $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$,

4. $\langle x, y\rangle = \overline{\langle y, x\rangle}$.

A real or complex space X with an inner product $\langle ., .\rangle$ is called an inner product space.

An inner product space satisfies two important properties.

Theorem 2. *Let X be an inner product space and let $x, y \in X$. Then:*

*1. $|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle$, $x, y \in X$, (**The Cauchy-Schwarz Inequality**).*

2. The function $\|\cdot\| : X \to \mathbb{C}$ defined by $\|x\| = \sqrt{\langle x, x\rangle}$, is a norm on X.

The proof for this theorem will be omitted however, this shows that every inner product space can also be regarded as a normed space as well. In an inner product space we are able to define notions such as the length of vectors and angles between these vectors.

Definition 3. Let X be an inner product space and $x, y \in X \setminus \{0\}$ then the angle θ between the vectors x, y are

$$\theta = \cos^{-1}\left(\frac{\langle x, y\rangle}{\|x\|\|y\|}\right). \tag{2.1}$$

Since we have a notion of what an angle would be, we are able to define what it means mathematically to have vectors be perpendicular to each other.

Definition 4. Let X be an inner product space, the vectors $x, y \in X$ are said to be *orthogonal* if $\langle x, y\rangle = 0$.

We are already well accustomed to the concept of orthonormal sets from real Euclidean space, $\mathbb{R}^n$ with the dot product. This can be generalized to arbitrary inner product spaces.

Definition 5. Let X be an inner product space. The set $\{x_1, \ldots, x_k\} \subset X$ is said to be *orthonormal* if $\langle x_i, x_j \rangle = \delta_{ij}$, where

$$\delta_{i,j} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

is the Kronecker Delta function.

If X is a real or complex inner product space that is also *complete*, then X is called a *Hilbert Space*. If X is a finite dimensional inner product space then it is already complete and hence a Hilbert space. If the reader is unfamiliar with the definition of completeness they may refer to [19]. In this thesis we will be focusing on finite dimensional Hilbert spaces.

The following concepts will be useful later on, especially when discussing quantum information.

Definition 6. Let X be an inner product space and let $S \subset X$, then the *orthogonal complement of* S is the set

$$S^{\perp} = \{x \in X : \langle x, s \rangle = 0, \text{ for all } s \in S\}.$$

$S^{\perp}$ is always a closed subspace of X. Orthogonal complements are important due to the following proposition.

Proposition 7. *Let $\mathcal{H}$ be a Hilbert space and let S be a closed subspace of $\mathcal{H}$, then for any $x \in \mathcal{H}$, there exists a unique $y \in S$ and $z \in S^{\perp}$ such that $x = y + z$.*

We will also make use of the following notation.

Definition 8. A subset Y of a vector space X is *convex* if, for all $x, y \in Y$ and $0 \leq \lambda \leq 1$, $\lambda x + (1 - \lambda) y \in Y$.

Definition 9. Let X be a subset of a convex Euclidean space, then the *convex hull of the set* X is the intersection of all convex sets containing the set X. We denote the convex hull of X, $conv(X)$.

For more details on these subjects of this subsection the interested reader can refer to [19, 43, 28].

2.2 Operators on a Hilbert Space

A natural continuation from the last chapter is to describe some of the operators that act on a Hilbert Space. These appear quite frequently in quantum information and will be used continuously through out the course of this work. Denote $B(X, Y)$ to be the space of bounded linear maps from X to Y. A linear map is bounded if and only if it is continuous on X and continuous on the origin. If $X = Y$ then we denote $B(X, X)$ as $B(X)$. In the finite dimensional case, every linear map is bounded and hence belongs to $B(X, Y)$. Given a specific basis (in the operator norm) on a finite dimensional Hilbert space, linear operators can then be represented as a matrix. Throughout the course of this thesis assume all Hilbert spaces are finite dimensional, hence unless otherwise stated, it can be assumed that an operator is equivalent (in the linear algebraic sense) to a matrix, with the understanding that an orthonormal basis for the underlying Hilbert space has been fixed.

Theorem 10. *Let $\mathcal{H}$ and $\mathcal{K}$ be complex Hilbert Spaces and let $T \in B(\mathcal{H}, \mathcal{K})$, then there exists a unique operator $T^* \in B(\mathcal{K}, \mathcal{H})$ such that*

$$\langle Tx, y \rangle = \langle x, T^* y \rangle$$

for all $x \in \mathcal{H}$ and $y \in \mathcal{K}$.

The proof for this theorem will be skipped as it is fairly elementary and not needed for this work.

Definition 11. If $\mathcal{H}$ and $\mathcal{K}$ are complex Hilbert Spaces and $T \in B(\mathcal{H}, \mathcal{K})$, the operator T^* is called the *adjoint of* T. Further, T is said to be *self-adjoint* or *Hermitian* if $T = T^*$.

The next definition is the matrix analogue of the previous definition.

Definition 12. If $A = [a_{i,j}]$, a $m \times n$ matrix then the matrix $A^* = [\overline{a_{j,i}}]$.

Definition 13. Let P be an operator in $B(\mathcal{H})$. The operator P is called an *projection* if $P^2 = P$. We denote the image of P as $P\mathcal{H}$. Further, P is an *orthogonal projection* if it satisfies $\langle Px, x\rangle = \langle x, Py\rangle$, for all x, y in $\mathcal{H}$. This condition equivalent to P being self-adjoint, $P = P^*$.

From here on orthogonal projections will be used so they will be referred to as projections.

Definition 14. Suppose T is an operator in $B(\mathcal{H})$. An operator T that satisfies $\|Tx\| = \|x\|$ is called an *isometry*, namely, T is a distance preserving operator.

Proposition 15. *The following are equivalent:*

1. $T^*T = I$.
2. *T is an isometry.*
3. $\|Tx - Ty\| = \|x - y\|$, *for all* x, y *in* $\mathcal{H}$.

Proof. 2. $\implies$ 3. By definition 14, if T is an isometry, then $\|Tx\| = \|x\|$ for all $x \in \mathcal{H}$. As T is linear, it follows that for all $x, y \in \mathcal{H}$,

$$\|Tx - Ty\| = \|T(x - y)\| = \|x - y\|.$$

3. $\implies$ 1. Assume $\|Tx - Ty\| = \|x - y\|$, so T is inner product preserving. Let $y = 0$, now

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, x\rangle = \|x\|^2.$$

Hence,

$$\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2 = \|x\|^2 = \langle x, x\rangle$$
$$\implies \langle (T^*T - I)x, x\rangle = 0, \text{ for all } x.$$

Thus $TT^* = I$.

1. $\implies$ 2. This follows from the definition of an isometry. □

Definition 16. Let $T \in B(\mathcal{H})$, if T satisfies $TT^* = I$, then T is called a *co-isometry.*

Corollary 17. *T is an isometry if and only if T^* is a co-isometry.*

Proof. First, assume T is an isometry. Then, $T^*T = I$. Then using the fact that $(T^*)^* = T$, we have

$$T^*(T^*)^* = T^*T = I.$$

Thus, T^* is a coisometry.

Conversely, by the same logic, T is an isometry if T^* is a coisometry. □

Proposition 18. *Let T be an isometry, then TT^* is a projection.*

Proof. Assume T is an isometry, then $T^*T = I$. Squaring TT^* we get:

$$(TT^*)^2 = TT^*TT^* = T(I)T^* = TT^*.$$

Also,

$$(TT^*)^* = (T^*)^*(T)^* = TT^*.$$

Thus, TT^* is a projection. □

By the same logic it can be shown if T is a co-isometry then T^*T will be the projection.

Definition 19. Let $U \in B(\mathcal{H})$, if U satisfies $UU^* = U^*U = I$ then U is called a *unitary operator*.

In a finite dimensional space every isometry is a unitary. Now that the above operators have been defined, a partial isometry $T \in B(\mathcal{H})$ will be defined. All isometries, projections and unitaries are also partial isometries. Two equivalent definitions are presented below. The range and kernel of an operator T will be denoted $\text{ran}(T)$ (or $\text{Im}T$) and $\text{ker}T$, respectively.

Definition 20. A *partial isometry* is an operator T such that $T = TT^*T$. Further, a partial isometry T has an initial projection $P = T^*T$ and its range projection $Q = TT^*$. That is, T maps $P\mathcal{H}$ isometrically onto $Q\mathcal{H}$, and vanishes on $P^{\perp}\mathcal{H}$.

Alternatively, a partial isometry also has the following characterization. We will give the argument for one direction of the proof below.

Proposition 21. *If T is a partial isometry then $\|Tx\| = \|x\|$ for all $x \in (\text{ker}T)^{\perp}$.*

Proof. Assume T is a partial isometry, then $T = TT^*T$. Firstly, $x \in (\text{ker}T)^{\perp}$ if and only if $x \in \text{ran}(T^*)$, i.e. there exists some v such that $T^*v = x$. Now,

$$T^* = (TT^*T)^* = T^*(T^*)^*T^* = T^*TT^*.$$

Thus, we have,

$$x = T^*v = T^*TT^*v = T^*T(T^*v) = T^*Tx.$$

This implies:

$$\|x\|^2 = \langle x, x\rangle = \langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2,$$

for all $x \in (\text{ker}T)^{\perp}$. Thus, $\|Tx\| = \|x\|$ for all $x \in (\text{ker}T)^{\perp}$. □

Remark 22. We call the space $(\text{ker}T)^{\perp}$ the *initial space* of T and the range of T the *final space* of T.

Every partial isometry is necessarily a contraction, (i.e. $\|Tx\| \leq \|x\|$ for all $x \in \mathcal{H}$). It is also important to state the difference between partial isometries on finite and infinite dimensional Hilbert Spaces. If $\mathcal{H}$ is finite dimensional, every isometry $T \in B(\mathcal{H})$ is surjective and hence a unitary operator [54]. However, if $\mathcal{H}$ is infinite dimensional, then this is not necessarily true. The following are examples of partial isometries. Example 23 is an infinite dimensional isometry that is not surjective where as Example 24 is in fact surjective.

Example 23. Let $\mathcal{H}$ be a separable infinite dimensional Hilbert Space (i.e. $\mathcal{H}$ has a countable orthonormal basis). Let $\{e_n\}$ be an orthonormal basis for $\mathcal{H}$. Define operator T on H as follows:

$$T\left(\sum_{n=1}^{\infty} a_n e_n\right) = \sum_{n=1}^{\infty} a_n e_{n+1}, \text{ where } \{a_n\} \in l^2.$$

Then T is an isometry which is not onto (since $\mathcal{H}$ is infinite). This operator is called the forward shift [54].

Example 24. Let $\mathcal{H}$ be a separable infinite dimensional Hilbert space with an orthonormal basis $\{e_n\}$ for $\mathcal{H}$. Now, let the operator T be defined by:

$$T\left(\sum_{n=1}^{\infty} a_n e_n\right) = \sum_{n=1}^{\infty} a_{n+1} e_n, \text{ where } \{a_n\} \in l^2.$$

Then T is a partial isometry with the initial space being the orthogonal complement of the vector e_1. This operator is called the backward shift [54] and is a coisometry.

Remark 25. In general, isometries and co-isometries are partial isometries. Operators are unitary if and only if they are both isometries and co-isometries. There are partial isometries that are not isometries nor coisometries.

Now we characterize what it means for an operator to be positive as well as the equivalent condition for matrices.

Definition 26. a) Let $\mathcal{H}$ be a complex Hilbert space, let $S \in B(\mathcal{H})$, S is *positive* if it is self-adjoint and

$$\langle Sx, x\rangle \geq 0, \text{ for all } x \in \mathcal{H}.$$

b) If A is a self-adjoint $n \times n$ matrix then A is *positive (semidefinite)* if

$$\langle Ax, x\rangle \ \geq 0 \text{ for all } x \in \mathbb{C}^n.$$

2.2.1 Spectrum of an Operator

The spectrum of a matrix provides some important insight in many applications of finite dimensional linear algebra. Speaking loosely, spectral theory characterizes when an operator or matrix is able to be diagonalized.

Definition 27. Let $\mathcal{H}$ be a complex Hilbert space, $I \in B(\mathcal{H})$ be the identity operator and $T \in B(\mathcal{H})$. The *spectrum of* T, denoted $\beta(T)$, is the set

$$\beta(T) = \{\lambda \in \mathbb{C} : T - \lambda I \text{ is not invertible}\}.$$

From this definition it is easy to see that if $T \in B(\mathcal{H})$, and if λ is an eigenvalue of T then $\lambda \in \beta(T)$.

The next theorem will not be proved as it can be found in any linear algebra text, this is the Spectral Theorem for Hermitian matrices.

Theorem 28. *Let A be a complex $n \times n$ Hermitian matrix, then the following are true:*

1. *All of matrix A's eigenvalues are real.*
2. *The eigenvectors of A corresponding to distinct eigenvalues are orthogonal to each other.*
3. *There exists an orthonormal basis of the Hilbert space consisting entirely of eigenvectors of A.*
4. *Matrix A is diagonalizable, that is, there exists a unitary U such that*

$$A = U\Lambda U^*,$$

where Λ is a real diagonal matrix. The diagonal entries of Λ are the eigenvalues of A and the columns of U are an orthonormal basis of the eigenvectors of A.

2.2.2 Important Decompositions

Theorem 29. *Let $\mathcal{H}$ be a complex Hilbert space.*

1. *If $\mathcal{V}$ is a non-zero subspace of $\mathcal{H}$ there is an orthogonal projection $P_{\mathcal{V}} \in B(\mathcal{H})$ with* $\operatorname{ran}(P_{\mathcal{V}}) = \mathcal{V}$ *and* $\ker(P_{\mathcal{V}^\perp}) = \mathcal{V}^\perp$ *and* $\|P_{\mathcal{V}}\| = 1$.

2. *If $Q \in B(\mathcal{H})$ is an orthogonal projection then the range of Q is a linear subspace and* $Q = P_{\operatorname{ran}(Q)}$.

Proof. 1. Assume $\mathcal{V} \neq 0$ is a linear subspace of $\mathcal{H}$. Let $x \in \mathcal{H}$ then by Proposition 7 we are able to write these elements as $x = y + z$ where $y \in \mathcal{V}$ and $z \in \mathcal{V}^\perp$. Let $P_{\mathcal{V}} : \mathcal{H} \to \mathcal{H}$ be defined by $P_{\mathcal{V}}(x) = y$. We first show $P_{\mathcal{V}}$ is a linear transformation so, let $x_1, x_2 \in \mathcal{H}$ with $x_1 = y_1 + z_1$ and $x_2 = y_2 + z_2$ where $y_1, y_2 \in \mathcal{V}$, $z_1, z_2 \in \mathcal{V}^\perp$ and $\lambda, \mu \in \mathbb{C}$.

Since $\mathcal{V}$ and $\mathcal{V}^\perp$ are subspaces of $\mathcal{H}$ we know that $\lambda y_1 + \mu y_2 \in \mathcal{V}$ and $\lambda z_1 + \mu z_2 \in \mathcal{V}^\perp$, so $\lambda x_1 + \mu x_2 = (\lambda y_1 + \mu y_2) + (\lambda z_1 + \mu z_2)$. Hence,

$$P_{\mathcal{V}}(\lambda x_1 + \mu x_2) = \lambda y_1 + \mu y_2 = \lambda P_{\mathcal{V}} x_1 + \mu P_{\mathcal{V}} x_2.$$

Thus, $P_{\mathcal{V}}$ is a linear transformation.

Next, $P_{\mathcal{V}}$ must be shown to be bounded and self adjoint.

$$\|P_{\mathcal{V}} x\|^2 = \|y\|^2 \leq \|x\|^2 = \|y\|^2 + \|z\|^2,$$

$P_{\mathcal{V}}$ is bounded and $\|P_{\mathcal{V}}\| \leq 1$. Also, for $0 \neq x \in \mathcal{V}$, $P_{\mathcal{V}} x = x$, and so $\|P_{\mathcal{V}}\| = 1$. Furthermore,

$$\langle P_{\mathcal{V}} x_1, x_2 \rangle = \langle y_1, y_2 + z_2 \rangle = \langle y_1, y_2 \rangle + \langle y_1, z_2 \rangle = \langle y_1, y_2 \rangle.$$

Since, $z_2 \in \mathcal{V}^\perp$ and similarly for $\langle x_1, P_\mathcal{V} x_2 \rangle = \langle y_1, y_2 \rangle$. Thus, $P_\mathcal{V}$ is self adjoint.

Now we checking if $P_\mathcal{V}$ is an orthogonal projection with range $\mathcal{V}$ and kernel $\mathcal{V}^\perp$. If $v \in \mathcal{V}$ then the orthogonal decomposition is $v = v + 0$ so, $P_\mathcal{V} v = v$. Hence, $\mathcal{V} \subseteq \text{ran}(P_\mathcal{V})$. However, $\text{ran} P_\mathcal{V} \subseteq \mathcal{V}$ by definition of $P_\mathcal{V}$. Thus, $\text{ran} P_\mathcal{V} = \mathcal{V}$.

For all $x \in \mathcal{H}$,

$$(P_\mathcal{V})^2(x) = P_\mathcal{V}(P_\mathcal{V} x) = P_\mathcal{V} y = y = P_\mathcal{V} x,$$

since, $y \in \mathcal{V}$. Therefore, $P_\mathcal{V}^2 = P_\mathcal{V}$ and $P_\mathcal{V}$ is an orthogonal projection.

Lastly,

$$\ker(P_\mathcal{V}) = (\text{ran} P_\mathcal{V}^*)^\perp = \text{ran}(P_\mathcal{V})^\perp = \mathcal{V}^\perp.$$

2. Let $\mathcal{V} = \text{ran}(Q)$, as Q has already been proved to be a linear transformation, $\mathcal{V}$ is a linear subspace. Let $y \in \mathcal{V}$ then $Qx = y$ for some $x \in \mathcal{H}$, so $Qy = Q^2 x = Qx = y$. If $z \in \mathcal{V}^\perp$ then

$$\|Qz\|^2 = \langle Qz, Qz \rangle = \langle z, Q^2 z \rangle = \langle z, Qz \rangle = 0,$$

by orthogonality. Thus, $Qz = 0$.

Hence, if $x \in \mathcal{H}$ and $x = y + z$ we get $x = Qy + z$ so,

$$P_\mathcal{V} x = y = Qy = Qx,$$

as $Qz = 0$. Hence, $Q = P_{(\text{ran} Q)} = P_\mathcal{V}$. □

The next proposition gives the explicit form of a projection onto a subspace.

Proposition 30. *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{V} = \text{span}\{e_1, \ldots, e_k\}$ be a subspace of $\mathcal{H}$*

where $\{e_j\}$ is an orthonormal basis for $\mathcal{V}$. Now, let $x \in \mathcal{H}$ then

$$P_{\mathcal{V}}(x) = \sum_{j=1}^{k} \langle x, e_j \rangle e_j. \tag{2.2}$$

Proof. Let $x \in \mathcal{H}$ then $x = P_{\mathcal{V}}(x) + P_{\mathcal{V}^{\perp}}(x)$. Now, $P_{\mathcal{V}}(x) = \sum_{j=1} a_j e_j$ for some $a_j \in \mathbb{C}$. Now we need to show that $a_j = \langle x, e_j \rangle$. For a fixed j_o,

$$a_{j_o} = \langle P_{\mathcal{V}} x, e_j \rangle = \langle x, P_{\mathcal{V}}^{*} e_j \rangle = \langle x, P_{\mathcal{V}} e_j \rangle = \langle x, e_j \rangle.$$

Thus, $P_{\mathcal{V}}(x) = \sum_{j=1} \langle x, e_j \rangle e_j$. □

Lemma 31. *If $\mathcal{H}$ is a complex Hilbert space, $\mathcal{V}$ is a closed linear subspace of $\mathcal{H}$ and P is the orthogonal projection of $\mathcal{H}$ onto $\mathcal{V}$, then $I - P$ is the orthogonal projection of $\mathcal{H}$ onto $\mathcal{V}^{\perp}$.*

Proof. As I and P are both self adjoint, so $I - P$ is as well. Also, we check the idempotency of $(I - P)$,

$$(I - P)^2 = I^2 - 2P + P^2 = I - 2P + P = I - P,$$

showing $(I - P)$ is in fact an orthogonal projection.

Now, let $x \in \mathcal{H}$ then $x = y + z$ for $y \in \mathcal{V}$ and $z \in \mathcal{V}^{\perp}$, where $Px = y$. So,

$$(I - P)(x) = (I - P)(x) = Ix - Px = x - y = y + z - y = z.$$

Hence, $(I - P)$ is the orthogonal projection onto $\mathcal{V}^{\perp}$. □

The next decompositions will use the definition of a positive operator.

Proposition 32. *Let $\mathcal{H}$ be a complex Hilbert space and let $T \in B(\mathcal{H})$. A square root of T is a unique positive operator $R \in B(\mathcal{H})$ such that $R^2 = T$.*

Theorem 33. *Let $\mathcal{H}$ be a complex Hilbert space and let $T \in B(\mathcal{H})$ be invertible. Then $T = UR$, where U is unitary and R is positive. This is called the polar decomposition of T.*

Proof. Since T is invertible, so is T^* and T^*T. Now, T^*T is positive since $\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle \geq 0$.

So T^*T has a positive square root $R = (T^*T)^{\frac{1}{2}}$. Since T^*T is invertible, so is R. Let $U = TR^{-1}$. Then U is invertible and so the range of U is $\mathcal{H}$. Also,

$$U^*U = (R^{-1})^*T^*TR^{-1} = (R^{-1})^*R^2R^{-1} = R^{-1}R^2R^{-1} = I.$$

Since U is invertible, this shows that U is in fact unitary. □

The next two theorems will be stated but not proven.

Theorem 34. *Let A be a Hermitian positive semi-definite matrix, then the Cholesky decomposition of A is:*

$$A = L^*L,$$

where L is a lower triangular matrix with real positive diagonal entries. Every Hermitian positive semi-definite matrix is able to be decomposed into this form uniquely.

Theorem 35. *Let A be an $m \times n$ matrix of rank r. Then A can be factorized in the form*

$$A = U\Sigma V,$$

This is called the singular value decomposition. Here U is an $m \times m$ unitary matrix, Σ is an $m \times n$ diagonal matrix with non-negative real numbers on the diagonal and V is an $n \times n$ unitary matrix. The non-negative real numbers on the diagonal of Σ are σ_i and are called the singular values of A. $(\sigma)_i^{\downarrow}$ is the notation for the singular values of a matrix arranged in a vector in decreasing order of value.

The singular value decomposition is a generalization of eigenvalue decomposition for general $m \times n$ matrices.

Lemma 36. *A non-negative real number σ is a singular value of a matrix A if and only if there exist unit vectors $u \in \mathbb{C}^m$ and $v \in \mathbb{C}^n$ such that*

$$Av = \sigma u, \tag{2.3}$$

and

$$A^* u = \sigma v. \tag{2.4}$$

Where u and v are called the left and right singular vectors for σ, respectively.

The next proposition discusses the behaviour of singular values of isometries and will be useful later in this thesis.

Proposition 37. *T is an $m \times n$ partial isometry if and only if the singular values of T are all in the set $\{0, 1\}$.*

Proof. Let T be a $m \times n$ partial isometry. From Definition 20 we have that $T = TT^*T$. We will begin by using this definition of a partial isometry with the equations (2.3) and (2.4) from the above lemma.

$$\sigma u = Tv = TT^*Tv = \sigma(TT^*u) = \sigma^2(Tv) = \sigma^3 u.$$

Thus, $\sigma^3 - \sigma^2 = 0$ and $\sigma(\sigma^2 - 1) = 0$, which implies $\sigma \in \{0, \pm 1\}$. But since singular values are always non-negative, $\sigma_i \in \{0, 1\}$.

Conversely, assume the singular values of T are in $\{0, 1\}$. We start with Theorem 71, so $T = U\Sigma V$, where Σ is a diagonal matrix of ones and zeros as per our assumption.

$$TT^*T = U\Sigma(VV^*)\Sigma(U^*U)\Sigma V = U\Sigma I\Sigma I\Sigma V = U\Sigma^3 V = U\Sigma V = T.$$

Since Σ is a diagonal matrix of ones and zeros $\Sigma^3 = \Sigma$. Thus, T is a partial isometry. □

Remark 38. In the finite dimensional case, if T is a $m \times n$ partial isometry, we have 3 special cases:

1. If $m = n$ and all n singular values of the unitary are 1 then T is unitary.

2. If $m \leq n$ then T is a co-isometry. All m singular values will be 1, and the rest will be 0.

3. If $m \geq n$ then T is an isometry. All n singular values will be 1, and the rest will be 0.

These are special cases of partial isometries but of course they are not exhaustive in the sense that there are more possibilities. The references for the results in this section can be found in [43, 16]

2.3 Tensor Products

The next definition is called the Kronecker product. It is the matrix version of the more general tensor product. Let the set of all $m \times n$ matrices with complex entries be denoted $M_{m,n}(\mathbb{C})$. If $m = n$ then we write $M_m(\mathbb{C})$.

Definition 39. Let $A \in M_{n,k}(\mathbb{C})$ and $B \in M_{m,l}(\mathbb{C}^{m\times l})$, then $A \otimes B \in M_{n,k}(\mathbb{C}) \otimes M_{m,l}(\mathbb{C})$ is defined as:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \ldots a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nn}B \end{bmatrix}, \text{ where } a_{ij} \text{ denotes the } ij^{th} \text{ entry of } A.$$

It is often beneficial to merge matrices in order to create a new $nm \times kl$ matrix (a partitioned matrix consisting of $m \times l$ blocks) that will share certain desirable properties with the previously combined matrices.

The next Lemma states the properties of tensor products.

Lemma 40. *Let A be any $m \times n$ matrix, B be any $m' \times n'$ matrix, C be any $n \times r$ matrix and D be any $n' \times r'$ matrix then,*

1. *If α is a scalar then $\alpha \otimes A = A \otimes \alpha = \alpha A$*

2. $(A \otimes 0) = (0 \otimes A) = 0$

3. $(A \otimes B)^* = A^* \otimes B^*$

4. $A \otimes (B \otimes C) = (A \otimes B) \otimes C$ *(associativity)*

5. $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$ *(left distribution)*

6. $(A + B) \otimes C = (A \otimes C) + (B \otimes C)$ *(right distribution)*

7. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$,

where addition and multiplication of such matrices in any of the above properties imply that their matrix dimensions are compatible.

A series of theorems will now follow this lemma to show the mathematical usefulness of a tensor product.

Theorem 41. *Let A and B be two $n \times n$ and $m \times m$ matrices respectively. Let λ be an eigenvalue of A corresponding to eigenvector u and let ω be an eigenvalue of B corresponding to eigenvector v. Then $u \otimes v$ is an eigenvector of $A \otimes B$ with eigenvalue $\lambda\omega$.*

Proof.

$$(A \otimes B)(u \otimes v) = Au \otimes Bv = \lambda u \otimes \omega v = \lambda\omega(u \otimes v).$$

□

Proposition 42. *If $A \in M_n(\mathbb{C})$ and $B \in M_n(\mathbb{M})$ are positive semidefinite then $A \otimes B \in M_{m,n}(\mathbb{C})$ is positive.*

Proof. Assume A and B are positive semidefinite. It is known that A and B are Hermitian.

$$\text{Since } (A \otimes B)^* = \begin{bmatrix} \bar{a}_{11}B & \bar{a}_{12}B & \dots \\ \vdots & \ddots & \vdots \\ \bar{a}_{n1}B & \cdots & \bar{a}_{nn}B \end{bmatrix} = \begin{bmatrix} a_{11}B & a_{12}B & \dots \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nn}B \end{bmatrix} = (A \otimes B)$$

Thus, $(A \otimes B)$ is also Hermitian. By the spectral theorem, there exists an orthonormal sets of eigenvectors for A and B, call them x_i and y_j respectively. These eigenvectors correspond to eigenvalues λ_i and μ_j (for $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$). It will be taken as an obvious verification that:

$$(A \otimes B)(x_i \otimes y_j) = \lambda_i \mu_j (x_i \otimes y_j).$$

So, $x_i \otimes y_j$ are eigenvectors of $A \otimes B$ with eigenvalues $\lambda_i \mu_j$. Since A and B are positive semidefinite matrices, their eigenvectors are λ_i and μ_j. Thus, $\lambda_i \mu_j \geq 0$. Since the eigenvalues of $A \otimes B$ are non-negative, $A \otimes B$ is positive semidefinite. □

Corollary 43. *Let A and B be two $n \times n$ and $m \times m$ matrices respectively. Then*

$$\text{Tr}(A \otimes B) = \text{Tr}(A)\text{Tr}(B).$$

Proof. Let the of eigenvalues of A be $\lambda_1, ..., \lambda_n$ and eigenvalues of B be $\omega_1, ..., \omega_m$. Then $Tr(A) = \lambda_1 + \cdots + \lambda_n$ and $\text{Tr}(B) = \omega_1 + \cdots + \omega_m$. By Theorem 4 the eigenvalues of $(A \otimes B)$ are

$$\lambda_1\omega_1, ..., \lambda_1\omega_m, ..., \lambda_i\omega_1, ..., \lambda_n\omega_1, ..., \lambda_n\omega_m.$$

Therefore, $\text{Tr}(A \otimes B) = \lambda_1(\omega_1 + \cdots + \omega_m) + \cdots + \lambda_n(\omega_1 + \cdots + \omega_m) = \lambda_1 \text{Tr}(B) + \cdots + \lambda_n \text{Tr}(B) = \text{Tr}(A)\text{Tr}(B)$. □

Corollary 44. *Let A and B be two $n \times n$ and $m \times m$ matrices respectively. Then*

$$\det(A \otimes B) = (\det A)^m (\det B)^n.$$

Proof. Let the eigenvalues of A be $\lambda_1, ..., \lambda_n$ and the eigenvalues of B be $\omega_1, ..., \omega_m$. Then $\det A = \lambda_1 \cdots \lambda_n$ and $\det B = \omega_1 \cdots \omega_m$. Furthermore, the eigenvalues of $(A \otimes B)$ are

$$\lambda_1\omega_1, ..., \lambda_1\omega_m, ..., \lambda_i\omega_1, ..., \lambda_n\omega_1, ..., \lambda_n\omega_m.$$

Therefore, $\det(A \otimes B) = \lambda_1^m \cdots \lambda_n^m (\omega_1 \cdots \omega_m)^n = (\lambda_1 \cdots \lambda_n)^m (\omega_1 \cdots \omega_m)^n = (\det A)^m (\det B)^n$.

□

Corollary 45. *Let $a \in \mathbb{C}^n$, $b \in \mathbb{C}^m$, $c \in \mathbb{C}^p$, and $d \in \mathbb{C}^q$. Then $\langle a \otimes c, b \otimes d \rangle = \langle a, b \rangle \langle c, d \rangle$*

Proof.

$$\langle a \otimes c, b \otimes d \rangle = (a^* \otimes c^*)(b \otimes d) = a^*b \otimes c^*d = \langle a, b \rangle \otimes \langle c, d \rangle$$

Since the tensor of two scalars is a scalar $\implies \langle a, b \rangle \langle c, d \rangle$. □

2.4 Quantum Information

The last section to this preliminary material is a brief description of some main topics in quantum information that will be treated as assumed knowledge later in the work. For the interested reader, introductions that go into greater detail on quantum information can be referenced in [40, 39, 41, 45, 46].

2.4.1 Qubits and Dirac Notation

It is natural to start at the most basic object of quantum information, the qubit. The qubit is the quantum analogue of the classical bit. The defining difference between the two objects is fundamentally that a classical bit can only have a value of a 0 or 1, where as a qubit is a quantum state, meaning that it can be represented as a linear combination of a 0 or a 1 with probability amplitudes attached to these states. Both bits and qubits have physical realizations in order to perform computations however we will focus on the abstract mathematical description of a qubit and its linear algebraic representation as a unit vector in $\mathbb{C}^2$:

$$|\phi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle, \text{ where } \alpha_0, \alpha_1 \in \mathbb{C}.$$

This linear combination of classic states is called a quantum superposition with $|\alpha_0|^2+|\alpha_1|^2 = 1$. Denote $|i\rangle \in \mathbb{C}^n$ as the ith position with a 1 in the vector and the rest of the elements in the vector being 0, for example in the 2-qubit case the two classical bits 0 and 1 are represented as

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Before going further the notation used above will be discussed, namely, Dirac Notation. Using the Dirac Notation, every column vector a is written as $|a\rangle$ and pronounced a "ket". The dual of vector a is written as $\langle a|$ and called "bra". Now that we have these we can write the inner product as $\langle a|a\rangle = a^*a$. Further, we write the outer product as $|a\rangle\langle a| = aa^*$. These are the very basics of Dirac Notation, it will be used interchangeably with regular linear algebra notation as we go forward. If new notation is used it will be stated.

As stated above, Hilbert space is used to study quantum mechanics abstractly. Hilbert space allows us to generalize common practices of linear algebra and calculus and extend it to higher finite or infinite dimensions. Since Hilbert space is also a generalized vector space it also has an inner product allowing the length and angle to be measured. These features of Hilbert space are necessary when studying quantum information.

Due to the nature of classical information (0 or 1), a Hilbert space $\mathcal{H}$ of interest usually will have a dimension of 2^n, where n is the number of bits. Alternatively, the Hilbert space can written as $\mathcal{H}^{\otimes n} := \bigotimes_{i=1}^{n} \mathcal{H}$, where $\mathcal{H}$ is a 2 dimensional Hilbert space. The standard basis is a convenient basis to work with to define a Hilbert space. The elements of this Hilbert space can be thought of as ket vectors of string length n where each character representing a 2×1 vector is separated by the tensor product. Observe:

$$|0 \cdots 0\rangle = |0\rangle \otimes \cdots \otimes |0\rangle = \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix},$$

where there are n $|0\rangle$'s tensored together. The same is true for $|1\rangle$. The standard basis for $\mathcal{H}^{\otimes n}$ is all the permutations of 0's and 1's, that is:

$$|00 \cdots 00\rangle \ , \ |01 \cdots 00\rangle \ , \ \ldots \ , \ |11, \cdots 10\rangle \ , \ |11 \cdots 11\rangle.$$

An arbitrary vector in $\mathcal{H}$ can also be written as a weighted sum of the basis vectors of $\mathcal{H}$, this of course allows for the mathematical representation of superposition states in quantum information/mechanics. This fact allows us to define another common basis for 2 dimensional Hilbert space, the Hadamard basis:

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \ , \ |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

The next definition allows us to characterize the basis of a Hilbert space using Dirac notation [25]:

Definition 46. Let $\mathcal{H}$ be a Hilbert space of dimension 2^n. A set of vectors $B = \{|b_i\rangle\}_{i=1}^{2^n} \subseteq \mathcal{H}$ is called an *orthonormal basis* for $\mathcal{H}$ if

$$\langle b_i | b_j \rangle = \delta_{i,j},$$

and every state $|\phi\rangle \in \mathcal{H}$ can be written as

$$|\phi\rangle = \sum_{b_i \in B} \phi_i |b_i\rangle, \text{ for some } \phi \in \mathbb{C}.$$

The values of ϕ_i satisfy $\phi_i = \langle b_i | \phi \rangle$, and are called the coefficients of $|\phi\rangle$ with respect to basis

B.

A density operator (matrix) can be described as the description of a mixed quantum state (i.e. probability distributions on quantum states). It is defined mathematically as follows:

Definition 47. Let $\{|\phi_i\rangle\}$ be a set of arbitrary quantum states each with a probability p_i, then the *density operator (matrix)* that describes this quantum state can be written as:

$$\rho = \sum_i p_i \, |\phi_i\rangle\langle\phi_i|, \tag{2.5}$$

where $\sum_i p_i = 1$ and $p_i \geq 0$. Alternatively, ρ is positive and the trace of ρ is 1.

Now that we have formal definitions of the basic parameters we will be working with in quantum information, the next subsection will display how these parameters evolve through certain conditions.

2.4.2 Quantum Channels

A quantum channel is a description of the evolution of a quantum state. This is very applicable in quantum information as this could potentially describe a message being sent over a communication channel, where the original message is ρ and the evolution of the message is $\mathcal{E}(\rho)$, where $\mathcal{E}$ is the quantum channel. Before formally stating a quantum channel, we must have a description of positive maps.

Definition 48. A linear map $\phi : M_n(\mathbb{C}) \longrightarrow M_m(\mathbb{C})$ is called *positive* if $\phi(A)$ is positive, for all positive $A \in M_n(\mathbb{C})$.

The next definition uses new notation namely, ϕ_m, so we will introduce it here first. For $m \geq 1$, $\phi_m : M_{n,m}(\mathbb{C}) \longrightarrow M_{k,m}(\mathbb{C})$ defined by $\phi_m = \phi \otimes \mathrm{id}_m$ where id_m is the identity map on $M_m(\mathbb{C})$.

Definition 49. A linear map ϕ is called *m-positive* if ϕ_m is positive.

Example 50. ϕ is called 1-positive if $\phi(A)$ is positive semidefinite whenever A is.

ϕ is called 2-positive if a $2n \times 2n$ matrix

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \text{ is positive semidefinite } \Longrightarrow \begin{bmatrix} \phi(A) & \phi(B) \\ \phi(B^*) & \phi(C) \end{bmatrix} \text{ is positive semidefinite.}$$

Definition 51. A linear map ϕ is called *completely positive* if ϕ_m is positive $\forall m \in \mathbb{N}$.

Originally due to Man-Duen Choi, in [10], the next theorem presented is appropriately known as Choi's Theorem. To be efficient, this theorem is written as a compact overview of the results Choi showed in [10]. The proof for this theorem will be omitted for this work.

Theorem 52. *Let $\phi : M_n(\mathbb{C}) \longrightarrow M_m(C)$ be a linear map. The following are equivalent:*

1. *ϕ is n-positive.*
2. *The matrix $C_\phi = (I_n \otimes \phi)(\sum_{ij} E_{ij} \otimes E_{ij}) = \sum_{ij} E_{ij} \otimes \phi(E_{ij}) \in M_{nm}(\mathbb{C})$ is positive. $E_{ij} \in M_n(\mathbb{C})$ is the elementary matrices with 1 in the i, j^{th} entry and 0's elsewhere.*
3. *$\phi(A) = \sum_{j=1}^{nm} V_j A V_j^*$ where $V_j \in M_{n,m}(\mathbb{C})$, $1 \leq j \leq nm$, $\forall A \in M_m(\mathbb{C})$.*
4. *ϕ is completely positive.*

Remark 53. The trace-preserving condition of a quantum channel, $\mathcal{E}$, refers to the trace of the original state being equal to the trace of output after the state goes through the quantum channel, that is, $\mathrm{Tr}(\rho) = \mathrm{Tr}(\mathcal{E}(\rho))$, for all ρ.

It is now possible to formally define a quantum channel.

Definition 54. Let $\mathcal{H}_a$ and $\mathcal{H}_b$ be two Hilbert spaces. A map $\mathcal{E}$: $B(\mathcal{H}_a) \longrightarrow B(\mathcal{H}_b)$ is a *quantum channel* if it is a linear, trace-preserving, completely positive map (TPCP).

More explicitly, Kraus operators can be used to define the operator sum representation of a quantum channel, that is a written output of a state from a quantum channel. Let ϕ be a quantum channel and $\rho \in B(\mathcal{H})$, then

$$\phi(\rho) = \sum_i V_i \rho V_i^*, \tag{2.6}$$

for all i, where the operators $\{V_i\}$ are the Kraus operators. Equation (2.6) is called the Kraus representation. The Kraus operators must satisfy the completeness relation which is equivalent to the map being trace persevering, that is, $\sum_i^m V_i^* V_i = I_m$, where I_m is the identity matrix of size $m \times m$. Kraus operators are also known as error operators in quantum error correction, we will see why in later chapters. Next are some properties of Kraus operators.

Proposition 55. *Let $\mathcal{H}$ be a Hilbert space and ϕ_1, ϕ_2 be two quantum channels with Kraus decompositions*

$$\phi_1(\rho) = \sum_{i=1} A\rho A^* \quad \textit{and} \quad \phi_2(\tau) = \sum_{i=1} B\tau B^*,$$

respectively.

i) The composition $\phi_2 \circ \phi_1$ is a quantum channel. Its Kraus decomposition is given by

$$\phi_2\phi_1(\rho) = \sum_{i,j=1} B_j A_i \rho A_i^* B_i^*.$$

ii) Any convex combination $\lambda\phi + (1-\lambda)\phi_2$ $(0 \leq \lambda \leq 1)$ is a quantum channel. The Kraus decomposition

$$(\lambda\phi_1 + (1-\lambda)\phi_2)(\rho) = \sum_{i=1} (\sqrt{\lambda}A_i)\rho(\sqrt{\lambda}A_i)^* + \sum_{i=1} (\sqrt{1-\lambda}B_i)\rho(\sqrt{1-\lambda}B_i)^*.$$

Chapter 3

Quantum Error Correction

A main goal of this thesis is to identify a connection between quantum error correction and isoclinic subspaces. Isoclinic subspaces will be introduced next chapter but in this section the background material of quantum error correction that motivates the relationship between these topics will be discussed. Quantum error correction is a broad area of research with much of the work being done in the last two decades. Mathematically, there are many approaches one can take to adequately describe quantum error correction. It will benefit us to give a linear algebraic description as isoclinic subspaces are also a topic from linear algebra.

We begin by giving a treatment of classical error correction. This seems to be a necessary precursor as classical error correction is a well understood theory and is a good starting place. In fact, quantum error correction was built from classical error correction. Next, the basics of quantum error correction will be discussed and some important codes will be presented. In the third section of this chapter a more detailed description of the theory will be shown defining some key parts of quantum error correction to familiarize readers with its formalization. We then present the main focus of this section: the quantum error correction conditions or the Knill-Laflamme conditions. This will detail how the Knill-Laflamme conditions arise and will be shown along side a full proof of the statement. Lastly, the higher rank k-numerical range will be defined and some important results are stated.

3.1 Classical Error Correction

Error correction is a fairly intuitive idea, it is concerned with two things, the detection of errors and the reconstruction of the original data, with no errors. Error correction is absolutely necessary in order for a computer to be able to perform calculations as well as for communication methods/systems to become a reality. While there is still research going on in classical error correction to further improve upon it, classical error correction is a much better understood theory and much better implemented practically. The error correction codes we are able to apply during classical calculations allow the probability of a an error occurring to be negligible. However, when employing communication methods things do become more complicated. A good model to describe how classical communication works is Figure 3.1. This figure shows that the original message going in does not necessarily come back the same. It becomes encoded, goes through the channel, decoded and then estimated at the other side of the channel with the hope that a low probability of error occurred on the original message. The more noise there is in a system, the greater probability of error on a message. Noise is defined as the phenomena in a communication channel (the medium in which information travels over) that may cause errors. The probability of an error occurring when sending information over a large distance is increased. There is a greater potential for noise to cause problems in the system and on the information being transmitted [39]. The general idea of error correction is to protect against errors and noise. One method that is able to protect against errors is to send redundant information so that even if the information sent is corrupted, there will still be enough to recover the original information.

To describe this we identify a simple example of a noisy classical communication channel called the binary symmetric channel. This channel is described by assigning a value of p to be the probability of error (a bit flip) in a transmitted message. Since p is the

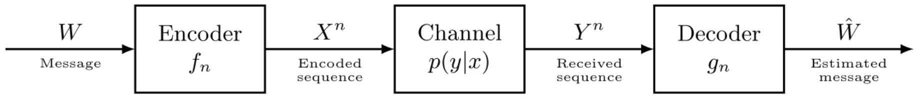

Figure 3.1: Model of a Classical Channel

probability of an error, $p > 0$ and of course the probability of there being no error is $1 - p$. Figure 3.2 depicts this.

To protect against a bit flip like mentioned above, a repetition code is used. This entails replacing a bit with copies of itself and sending it through a channel. In the case of the binary symmetric channel, 0 becomes 000 and 1 turns into 111. Once the repetition bit is sent through the channel it is decoded by a process known as majority voting. If p is not too high of a value then it is likely that whatever bit value appears more times when the message is decoded is the actually channel output. Essentially, if two or more bits are flipped when put through the channel then majority voting fails. Although it will not be shown here, this encoding can be proved to succeed by the probability of error, $p_e = 3p^2 - 2p^3$. This encoding is more reliable if $p_e < p$, thus, if $p = \frac{1}{2}$, this inequality is attained [40].

This is only a brief introduction to classical error correction in order to build intuition on how quantum error correction may work, some supplementary references are provided. These references provide a much deeper insight into classical information and error correction for the interested reader. These are some of the first papers that introduce error correction [44, 20] and next are some textbooks that go into greater detail and describe what is known about classical error correction in greater detail [36, 9].

Classical error correction is much simpler considering there are limited types of errors, most importantly a bit flip. In quantum error correction things get much more difficult.

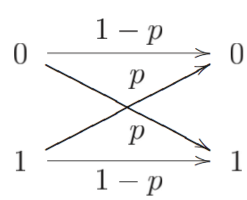

Figure 3.2: Binary Symmetric Channel Probability Description (from [40]).

3.2 Quantum Error Correction and Shor's Code

3.2.1 Limitations of QEC

Quantum computation changes the way information is processed. With potential innovation there are always a plethora of issues that go along with it. The nature of quantum mechanics poses some limitations on quantum error correction that were just not an issue in classical error correction. Reference [40] gives a nice description of these limitations and how they can be overcome. This section will give a brief synopsis of these limitations.

Firstly, there is a well known theorem in quantum mechanics called the no cloning theorem that in essence states it is impossible to create an identical copy of an arbitrary unknown quantum state. The problem the no cloning theorem creates is that quantum error correction codes cannot consist of arbitrary repeating codes like the aforementioned classical error correction method of majority voting employs. More can be read about this shortcoming of quantum error correction in [40, 37, 18, 41, 26, 45].

Moving on to the next major issue is the fact that errors have the potential to be continuous within quantum information. Many errors could occur on one single qubit and this would take infinite precision to detect and therefore an infinite amount of resources would be required to correct the information. Below, a way around this will be discussed.

Lastly, the very essence of quantum mechanics becomes a problem for error correction. In classical error correction the output can be observed from a channel and if incorrect (from a process like majority voting), it can be corrected back to its original state. If we take the same approach with quantum information, the observation after leaving the quantum channel will destroy the information forever, thus, recovery will become impossible.

3.2.2 Shor Code

One of the most ground breaking realizations that gave hope to many theorists that quantum error correction could be a possibility was the Shor Code. The error correcting code that Shor built for quantum information shows that arbitrary errors on a single qubit can be corrected. The interesting part about this is that this solves the continuum of errors problem on quantum information mentioned above. It compresses the infinite resources the quantum error correction process should need into a finite set of the Pauli operators. Before diving deeper into Shor code, it is beneficial to describe the types of errors on a qubit and how they are able to be fixed. They actually make up the basis of the Shor code.

Two examples of potential errors that can be found in quantum information processing are bit flips ($|0\rangle \to |1\rangle$ or $|1\rangle \to |0\rangle$) and phase flips ($|+\rangle \to |-\rangle$ or $|-\rangle \to |+\rangle$), and combinations of the two are possible as well. The bit flip is described by the quantum Pauli operator X. The action of a bit on a qubit is: $X|0\rangle = |1\rangle$ and $X|1\rangle = |0\rangle$. The matrix form of X is:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = |0\rangle\langle 1| + |1\rangle\langle 0|.$$

The phase flip operation is described by the Pauli operator Z. The action of Z on a qubit is: $Z|+\rangle = |-\rangle$ and $Z|-\rangle = |+\rangle$. The matrix form of Z is:

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = |0\rangle\langle 0| + |1\rangle\langle 1|.$$

Finally, both the bit and phase flip are described by XZ (or iY). The matrix form of Y is:

$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = i|0\rangle\langle 1| - i|1\rangle\langle 0|.$$

While we know an arbitrary quantum state can not be repeated, the way to protect against bit or phase flips is similar to a repetition code, it is called the three qubit flip/phase code. If a qubit has a probability of error of p (the error is bit **or** phase flip). Each can be encoded by a similar circuit changing one qubit state into a 3 qubit state, namely for bit flips:

$$a|0\rangle + b|1\rangle \to a|000\rangle + b|111\rangle \tag{3.1}$$

And the phase flip encoding:

$$a|+\rangle + b|-\rangle \rightarrow a|+++\rangle + b|---\rangle. \tag{3.2}$$

Quantum Error Correction is a two step process. Detection is the first step and then secondly recovering (or correcting) the information, so that it is changed back to its originally sent state. The way the error detection process works is by performing a syndrome measurement on the encoded state to determine if any errors took place. It is important to mention that the syndrome measurement does not change the quantum state or give any information about the probability of what the state may be, it only determines if any errors are present. This is why it is a viable option for detection, because the syndrome measurement does not destroy the quantum information. For the recovery part of the process, the error syndrome value measured will tell which qubit has an error. Depending on that value and what type of error occurred, a corresponding operation will be applied to correct that error. For a more intricate description and deeper understanding the reader is pointed to [18, 26, 41, 40]. An example will follow to outline this method further.

Example 56. Some potential errors on 3 qubit encoding for phase flips ($|+++\rangle$):

- $|+++\rangle$, syndrome measurement: $\rightarrow$ No errors, apply I.
- $|++-\rangle$, syndrome measurement: $\rightarrow$ Error on 3rd qubit, apply $Z_3 = I \otimes I \otimes Z$.
- $|-++\rangle$, syndrome measurement: $\rightarrow$ Error on 1st qubit, apply $Z_1 = Z \otimes I \otimes I$.

This is similar for the 3 qubit bit flip encoding except with the X_i Pauli being the error/recovery operation, where i is the ith qubit being operated on.

The bit and phase flip qubit codes work very similar to each other that is why they are presented together although they are both separate error correcting codes. Their main difference is the circuit used to create them, and the basis that they are measured in, the standard basis for bit flips, $\{|0\rangle, |1\rangle\}$ and the Hadamard basis for phase flips, $\{|+\rangle, |-\rangle\}$.

Now that it is understood how bit flip and phase flip errors can successfully be protected against separately, it is time to become more general and discuss a code that can protect from each of these errors or both of them. This code is called the Shor Code [45]. Shor Code actually combines both these codes to make it impervious to arbitrary errors on a single qubit. In this case the qubits are encoded to become the following 9-qubit states:

$$|0\rangle \rightarrow |0_L\rangle = \frac{1}{2\sqrt{2}}(|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle), \quad (3.3)$$

$$|1\rangle \rightarrow |1_L\rangle = \frac{1}{2\sqrt{2}}(|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle). \quad (3.4)$$

As shown here this is very similar to three-qubit repetition codes but now has much more power then any before. Like mentioned above, Shor's Code is so useful because it is able to narrow down the unknown set of potential errors on a single qubit into the finite set of the Pauli operators. To demonstrate this, we will define an arbitrary error to be named U. U will act on a single qubit of state $|\phi\rangle$. Since U is an arbitrary error $|\phi\rangle$, it can be described as a linear combination of all the possible errors, namely,

$$U = f_0 I + f_1 X + f_2 Y + f_3 Z, \quad (3.5)$$

where f_0, f_1, f_2, f_3 are complex constants. Now that we have written U like this, it means that:

- $U = I$ if there are no errors,
- $U = X$ if bit flip errors occur,
- $U = Z$ if phase flip errors occur,
- $U = XZ = iY$ if both bit and phase flip errors occur.

It turns out that Shor's 9-qubit Code can correct against arbitrary single qubit errors. This is a defining example for Quantum Error Correction, because Shor's Code can correct errors on single qubits (i.e. $U \in \{I, X, Y, Z\}$). Furthermore, it showed there was a possibility of being able to create different codes to correct more complex errors on a large number of qubits, which of course must happen in order to make quantum information processing a reality.

3.3 Operators and the Theory of Quantum Error Correction

Thus far it has been made clear that it is possible to protect against arbitrary errors on quantum information, at least when using a small number of qubits. This is a good start but the goal is to protect against errors on a large number of qubits in order to be able to process quantum information reliably. Although this is a multi-discipline problem considering there are all sorts of physical and experimental considerations to take into account, the mathematical approach to quantum error correction is an avenue to explore theoretical solutions on how a general quantum error correcting code may work in practice. This thesis relies on the fact that quantum error correction can be modeled using subspaces within a Hilbert space, hence that will be a focal point of discussion. As a note, $\mathcal{H}^{\otimes n}$ will be written as $\mathcal{H}$ in this subsection.

Firstly, like classical error correction, information needs to be encoded in order to be theoretically protected from errors when exposed to noise. Mathematically, unitary operators are the encoders used to map quantum states into a code space $\mathcal{C}$. Formally, $\mathcal{C}$ is defined next.

Definition 57. The *code space* $\mathcal{C}$ is a vector subspace of Hilbert space $\mathcal{H}$.

Quantum codes must be subspaces to allow for linear combinations of code words, which correspond to superpositions of classical states. We let $P_{\mathcal{C}}$ denote the projection of $\mathcal{H}$ onto the code subspace $\mathcal{C}$. Also, let $P_{\mathcal{C}}^{\perp}$ be the projection of $\mathcal{H}$ onto the orthogonal complement of the code space. In order to detect which errors may have occurred, a measurement must be performed. The measurement will determine if a given state belongs to the code space and thus has an error that needs to be corrected.

A code-word from the subspace $\mathcal{C}$ could be affected by noise in the transmission of the quantum information so a general error correction theory is desirable in order to detect and fix these potential errors. To do this, some assumptions need to be made about this noise and the recovery of the original information. It is a standard practice to define these assumptions as operations when describing the theory behind quantum error correction as [40] does so eloquently. These assumptions will now be briefly described.

Remark 58. Quantum Noise and Recovery Assumptions

1. Any noise that is found to change the original quantum state can be described by the quantum operation $\mathcal{E}$. $\mathcal{E}$ can be looked at as a quantum channel described by the operators $\mathcal{E} = \{E_i\}$. This must include the operator $E_0 = I$, which is the case in which there are no errors on the information. As a linear map, the noise model is given by the quantum channel
$$\mathcal{E}(\rho) = \sum_j E_j \rho E_j^*.$$

2. To counteract any errors on the system from $\mathcal{E}$, the trace-preserving quantum operation $\mathcal{R}$ is defined. This operation is an arbitrary error correcting procedure which constitutes the detection of an error as well as the recovery of original information.

Now, with these quantum operators defined we are able to state when error correction is successful from [30].

Definition 59. Let $\mathcal{E}$ be a quantum channel from Remark 58 and let $\mathcal{C}$ be a quantum code space from Definition 57 with the projection $\mathcal{P}_{\mathcal{C}}$. Then $\mathcal{C}$ is *correctable* for $\mathcal{E}$ if there is a quantum channel $\mathcal{R}$ such that
$$(\mathcal{R} \circ \mathcal{E})(\rho) \ = \ \rho \tag{3.6}$$
for all ρ supported on $\mathcal{C}$; that is all ρ with $P_{\mathcal{C}} \rho P_{\mathcal{C}} = \rho$.

What makes this process fruitful is being able to actually distinguish which error has occurred on the code. The reason this is possible is because of the error operators. The error operators correspond to undeformed orthogonal subspaces that are variants of the original code space. The subspaces are restricted to the code space $\mathcal{C}$ and can be denoted $S_k = \text{ran}(E_k|_\mathcal{C})$. The subspaces are an important part of quantum error correcting codes as we will see.

As stated above the subspaces must be orthogonal in order to be measured and distinguish an error in the encoded state. This is a restriction the definition above does not state explicitly but is necessary for quantum error correction. This brings us to a central theorem of interest in this thesis, the quantum error correction conditions. Originally the quantum error correction conditions were developed separately by Knill and Laflamme [26] and Bennett, DiVincenzo, Smolin and Wooters [4]. Before explicitly stating the quantum error correction conditions, a theorem is needed to prove them. This theorem appears in many works such as [30, 40] and it is a result that details mathematically how different sets of noise operators for the same quantum channel are related. Only sufficiency will be shown for this theorem as it is the simpler direction. The full proof can be found in [40] for the interested reader.

Theorem 60. *Suppose $\{E_1, ..., E_m\}$ and $\{F_1, ..., F_n\}$ are the noise operations giving rise to quantum operations $\mathcal{E}$ and $\mathcal{F}$, respectively. By appending zero operators to the shorter list of operation elements we ensure that $m = n$. Then $\mathcal{E} = \mathcal{F}$ if and only if there exist a complex unitary matrix $U = (u_{ij})$ such that*

$$E_i = \sum_j u_{ij} F_j, \text{ for all } i.$$

Proof. Let $\mathcal{E}(\rho) = \sum_i E_i \rho E_i^*$ and $\mathcal{F}(\rho) = \sum_j F_j \rho F_j^*$. Now, assume there exists there exists a complex unitary matrix $U = (u_{ij})$ such that $E_i = \sum_j u_j F_j$.

$$\begin{aligned}
\mathcal{E}(\rho) &= \sum_i E_i \rho E_i^* \\
&= \sum_i (\sum_j u_{ij} F_j) \rho (\sum_k u_{ik} F_j)^* \\
&= \sum_i (\sum_j u_{ij} F_j) \rho (\sum_k \bar{u}_{ki} F_k^*) \\
&= \sum_{ikj} \bar{u}_{ki} u_{ij} F_j \rho F_k^* \\
&= \sum_{kj} \delta_{kj} F_j \rho F_k^* \\
&= \sum_j F_j \rho F_j^* \\
&= \mathcal{F}(\rho).
\end{aligned}$$

Since matrix entries $\bar{u}_{ki}$ and u_{ij} are part of unitary matrices, they will reduce to the delta function (from Definition (46)) over the sum $\sum_{kj}$. This is because a unitary matrix's rows and columns are orthonormal sets. □

The following is also known as the Knill-Laflamme Theorem,

Theorem 61. ***(Quantum error correction Conditions)***
Let $\mathcal{C}$ be a quantum code space of $\mathcal{H}$with the projection $P_{\mathcal{C}}$ onto $\mathcal{C}$. Suppose $\mathcal{E}$ is a quantum operation with operation elements $\{E_i\}$. A necessary and sufficient condition for the existence of an error operator $\mathcal{R}$ correcting $\mathcal{E}$ on $\mathcal{C}$ is that for all i, j,

$$P_{\mathcal{C}} E_i^* E_j P_{\mathcal{C}} = \alpha_{ij} P_{\mathcal{C}}, \tag{3.7}$$

for some Hermitian matrix (α_{ij}) of complex numbers.

It is helpful to consider the Knill-Laflamme conditions in a distinguished special case. Suppose (3.7) is satisfied with $\alpha = (\alpha_{ij})$ such that $\alpha_{ij} = 0$ whenever $i \neq j$. Let

$|\phi\rangle, |\psi\rangle \in \mathcal{C}$. Then, $P_{\mathcal{C}}|\phi\rangle = |\phi\rangle$ and $P_{\mathcal{C}}|\psi\rangle = |\psi\rangle$, and for $i \neq j$ we have

$$\langle P_{\mathcal{C}} E_i^* E_j P_{\mathcal{C}} \phi | \psi \rangle = \alpha_{ij} \langle P_{\mathcal{C}} \phi | \psi \rangle = 0,$$

but also,

$$\begin{aligned} \langle P_{\mathcal{C}} E_i^* E_j P_{\mathcal{C}} \phi | \psi \rangle &= \langle E_j P_{\mathcal{C}} \phi | E_i P_{\mathcal{C}} \psi \rangle \\ &= \langle E_j \phi | E_i \psi \rangle. \end{aligned}$$

In other words, observe in this case that the subspaces $E_i\mathcal{C}$ and $E_j\mathcal{C}$ are orthogonal whenever $i \neq j$. This special case gives primary motivation for the proof that follows.

Proof. Let $\mathcal{C}$ be a quantum code space as described in Definition (57) and $P_{\mathcal{C}}$ be a projection onto this code space. First we prove the necessity of the quantum error correction conditions. Assume that $\mathcal{C}$ is correctable for $\mathcal{E}$ by the quantum operation $\mathcal{R}$ with noise operators $\{R_j\}$. Now we define a compressed channel by $\mathcal{E}_{\mathcal{C}}(\rho) = \mathcal{E}(P_{\mathcal{C}} \rho P_{\mathcal{C}})$. By our assumption, $\mathcal{R}(\mathcal{E}_{\mathcal{C}}) = \mathcal{R}(\mathcal{E}(P_{\mathcal{C}} \rho P_{\mathcal{C}})) = P_{\mathcal{C}} \rho P_{\mathcal{C}}$. Further,

$$\sum_{i,j} R_j E_i P_{\mathcal{C}} \rho P_{\mathcal{C}} E_i^* R_j^* = P_{\mathcal{C}} \rho P_{\mathcal{C}} \text{ for all } \rho.$$

Thus, by Theorem 60 there exist scalars λ_{ki} such that

$$R_k E_i P_{\mathcal{C}} = \lambda_{ki} \text{ for all } i, k.$$

Hence,

$$P_{\mathcal{C}} E_i^* R_k^* R_k E_j P_{\mathcal{C}} = \bar{\lambda}_{ki} \lambda_{kj} P_{\mathcal{C}} \text{ for all } i, j, k. \tag{3.8}$$

Since $\mathcal{R}$ preserves trace, so that $\sum_k R_k^* R_k = I$ and thus when we sum (3.8) over k we obtain

$$P_{\mathcal{C}} E_i^* E_k P_{\mathcal{C}} = \alpha_{ij} \text{ for all } i, j,$$

where $\alpha_{ij} = \sum_k \bar{\lambda}_{ki}\lambda_{kj}$, which is just a Hermitian matrix. These are the quantum error correction conditions.

Conversely, we again assume $\mathcal{E}$ is a quantum channel with errors $\{E_i\}$ and $\mathcal{C}$ a code space with projection $P_{\mathcal{C}}$ such that the quantum error correction conditions (i.e. equation (3.7)) hold. Since (3.7) tell us that α is a Hermitian positive semi-definite matrix, so we can diagonalize it to give us a diagonal matrix $D = U^*\alpha U$, where U is a unitary operator.
By Theorem (60), let $F_k = \sum_i u_{ik}E_i$ and thus $\{F_k\}$ is also a set of operation elements for $\mathcal{E}$. Now, using substitution equation, (3.7) and $D = U^*\alpha U$,

$$P_{\mathcal{C}}F_k^*F_lP_{\mathcal{C}} = \sum_{i,j} u_{ik}^*u_{jl}P_{\mathcal{C}}E_i^*E_jP_{\mathcal{C}} = \sum_{i,j} u_{ik}^*\alpha_{ij}u_{jl}P_{\mathcal{C}} = d_{kl}P_{\mathcal{C}}, \text{ for all k,l.}$$

This can be thought of as a simplification of the quantum error correction conditions. Now, using the polar decomposition for $F_KP_{\mathcal{C}}$ this gives

$$F_kP_{\mathcal{C}} = U_k\sqrt{P_{\mathcal{C}}F_k^*F_kP_{\mathcal{C}}} = \sqrt{d_{kk}}U_kP_{\mathcal{C}} \tag{3.9}$$

for some unitary U_k. Define projections $P_k = U_kP_{\mathcal{C}}U_k^* = \frac{F_kP_{\mathcal{C}}U_k^*}{\sqrt{d_{kk}}}$ (the second equated expression is due to (3.9)). We can assume that $d_{kk} \neq 0$ with loss of generality because these matrix entries will not in fact evolve into subspaces. Then, for $l \neq k$ we are able to deduce that the range of P_k or to be more clear, these subspaces are orthogonal to each other.

$$P_lP_k = P_l^*P_k = \frac{U_lP_{\mathcal{C}}F_l^*F_kP_{\mathcal{C}}U_k^*}{\sqrt{d_{ll}d_{kk}}} = 0.$$

It is important to assume $\sum_k P_k = I$, if we cannot make this assumption then the projection can be added onto the orthogonal complement and let $U_k = I$, then the recovery operation for Quantum Error Correction is

$$\mathcal{R}(\rho) = \sum_k U_k^*P_k\rho P_kU_k.$$

For all ρ with $\rho = P_{\mathcal{C}}\rho P_{\mathcal{C}}$ in the code $\mathcal{C}$ is shown to be correctable by the following algebraic manipulations:

$$U_k^* P_k F_l \sqrt{\rho} = U_k^* P_k^* F_l \sqrt{\rho} = \frac{U_k^* U_k P_{\mathcal{C}} F_k^* F_l P_{\mathcal{C}} \sqrt{\rho}}{\sqrt{d_{kk}}} = \delta_{kl}\sqrt{d_{kk}} P_{\mathcal{C}}\rho = \delta_{kl}\sqrt{d_{kk}}\rho$$

Hence,

$$\mathcal{R}(\mathcal{E}(\rho)) = \sum_{k,l} U_k^* P_k F_l \rho F_l^* P_k U_k = \sum_{k,l} \delta_{kl} d_{kk} \rho = \rho$$

for all $\rho = P_{\mathcal{C}}\rho P_{\mathcal{C}}$. □

This proof is included because the quantum error correction conditions are a centre piece of this paper. Further, the new results in the following chapters use Theorem 61. To complement this proof, next are some examples of the quantum error correction conditions being used to emphasize their practicality.

Example 62. For a 2-qubit space, the Hilbert space is defined as $\mathcal{H}_4 = \mathbb{C}^2 \otimes \mathbb{C}^2$. The basis of this Hilbert space is, $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Suppose the error operators are as follows $\{E_1 = \frac{1}{\sqrt{2}} I, E_2 = \frac{1}{\sqrt{2}} X_1 = \frac{1}{\sqrt{2}} X \otimes I\}$ where the projector onto the code space is $P = |00\rangle\langle 00| + |01\rangle\langle 01|$. Check to see if the quantum error correction conditions are satisfied, so $\mathcal{C} = \text{span}\{|00\rangle, |01\rangle\}$.

Firstly, $i = j$, for i $\in \{1, 2\}$:

$$PE_i^* E_i P = P\frac{1}{\sqrt{2}} I \frac{1}{\sqrt{2}} IP = \frac{1}{2} PP = \frac{1}{2} P.$$

Now check $i = 1, j = 2$,

$$\begin{aligned}
PE_1^* E_2 P &= P\frac{1}{\sqrt{2}} I \frac{1}{\sqrt{2}} X_1 P \\
&= \frac{1}{2}(|00\rangle\langle 00| + |01\rangle\langle 01|) I (X \otimes I)|00\rangle\langle 00| + (X \otimes I)|01\rangle\langle 01|) \\
&= \frac{1}{2}(|00\rangle\langle 00| + |01\rangle\langle 01|)(|10\rangle\langle 10| + |11\rangle\langle 11|) \\
&= |00\rangle\langle 00|10\rangle\langle 10| + |00\rangle\langle 00|11\rangle\langle 11| + |01\rangle\langle 01|10\rangle|10\rangle + |10\rangle\langle 10|11\rangle\langle 11| \\
&\text{By orthogonality, each inner product is equal to 0} \\
&= 0
\end{aligned}$$

Similarly, for $i = 2$ and $j = 1$,

$$PE_2^* E_1 P = P\frac{1}{\sqrt{2}} X_2 \frac{1}{\sqrt{2}} IP = 0$$

Thus, the quantum error correction conditions are satisfied for the Hermitian matrix

$$\alpha = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

Now using the Knill-Laflamme conditions and the proof we will easily be able to show that the Shor code is correctable for any single qubit error. Since there are so many examples of errors that could be shown in this nine qubit code, we will stick with a small error operator set.

Example 63. Let $\mathcal{H}$ be a Hilbert space defined as $\mathcal{H} = \mathbb{C}^{2^9}$ since Shor code is the nine qubit code. Refer back to (3.3) and (3.4) for the encoding of the $|0\rangle$ and $|1\rangle$ qubits and let

$$|\phi_0\rangle = \frac{|000\rangle + |111\rangle}{\sqrt{2}} \text{ and } |\phi_1\rangle = \frac{|000\rangle - |111\rangle}{\sqrt{2}}.$$

Now that we have Equation 3.7, the Knill-Laflamme conditions, we know that if we have an error that transforms the code subspace to an another orthogonal subspace, we have that this error is correctable.

Now, let the errors for a particular quantum channel be that nothing happens with some probability and a phase flip on the 3rd qubit block happen with another probability. Then, the error operators for the Shor nine qubit code will be $E_1 = I^{\otimes 9}$ and $E_2 = I^{\otimes 2} \otimes Z \otimes I^{\otimes 6}$. Observe the effect that these error operators have on

$$|\phi\rangle = \alpha|0_L\rangle + \beta|1_L\rangle = \alpha|\phi_0\rangle|\phi_0\rangle|\phi_0\rangle + \beta|\phi_1\rangle|\phi_1\rangle|\phi_1\rangle :$$

$$E_1|\phi\rangle \to \alpha|0_L\rangle + \beta|1_L\rangle.$$

Since these states are the same, this will satisfy Equation 3.7. For E_2,

$$E_2(\alpha|\phi_0\rangle|\phi_0\rangle|\phi_0\rangle + \beta|\phi_1\rangle|\phi_1\rangle|\phi_1\rangle) \to \alpha|\phi_0\rangle|\phi_0\rangle|\phi_1\rangle + \beta|\phi_1\rangle|\phi_1\rangle|\phi_0\rangle.$$

Since $E_2|\phi\rangle$ is still orthogonal to $|\phi\rangle$ by Theorem 61 satisfies equation 3.7 and thus the Shor nine qubit code is correctable for these error operators. This is a a specific example but it holds true for any single qubit error operator that belongs to the Pauli set as mentioned in previous sections.

3.4 Higher Rank Numerical Range and QEC

The higher rank numerical range of a matrix or operator has been an intensely investigated subject for over a decade now in matrix theory and beyond [33, 49, 38, 11, 35, 34, 32, 15]. There is a connection between the higher rank numerical range and quantum error correction originally considered in [14, 12]. It will be further explored in this work in the following chapter, so as a primer, a brief treatment of the higher rank numerical range will be given in this section.

Definition 64. Let $\mathcal{P}_k$ be the set of all rank k orthogonal projections of size $n \times n$. The *joint rank k-numerical range* of a m-tuple of $n \times n$ matrices $\mathbf{A} = (A_1, \ldots, A_m)$ is the subset of the complex space $\mathbb{C}^n$ given by

$$\Lambda_k(\mathbf{A}) = \{\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{C}^m | PA_iP = \lambda_i P \text{ for all } i \text{ and for some rank } k \text{ projection } P \text{ on } \mathbb{C}^n\}. \tag{3.10}$$

An interesting apparent connection the rank k-numerical range has with quantum error correction stated in [32] is as follows.

Proposition 65. *A quantum channel has error correcting code of k-dimension if and only if $\Lambda_k(E_1^* E_1, \ldots, E_r^* E_r) \neq \emptyset$.*

This proposition follows from the definition of the higher rank k-numerical range and the quantum error correction conditions. The next theorem is named Tverberg's theorem, a useful theorem that allows one to take a set in Euclidean space and partition it into subsets that have intersecting convex hulls. This theorem will be used in the proof of the Theorem 67. Tverberg's theorem will not be proved but the proof and other details can be found in [48].

Theorem 66. *Let r be a positive integer and let T be a set whose cardinality exceeds $(d+1)(r-1)$ that belongs to d-dimensional Euclidean space, then T can be partitioned into r subsets, $\{T_i\}_{i=1}^r$ whose convex hulls have a non-empty intersection:*

$$T = \bigcup_{i=1}^{r} T_i, \quad \text{for } i \neq j, S_i \cup S_j \neq \emptyset, \quad \bigcap_{i=k}^{r} \operatorname{conv}(T_k) \neq \emptyset.$$

Next is a property of the rank k-numerical range, it is essentially a different proof for the main result of [27]. The theorem's full proof is given next from [32].

Theorem 67. *Let be $\boldsymbol{A}$ an m-tuple of $n \times n$ self adjoint matrices $\boldsymbol{A} = (A_1, \ldots, A_m)$. For $m \geq 1$ and $k > 1$. If $dim\mathcal{H} = n \geq (k-1)(m+1)^2$, then $\Lambda_k(\boldsymbol{A}) \neq \emptyset$.*

Proof. Assume $dim\mathcal{H} = n = (k-1)(m+1)^2$. Otherwise, replace each A_j by $U^* A_j U$ for some U such that $U^*U = I_n$.

Let $q = (m+1)(k-1)+1$.

Choose an eigenvector x_1 of A_1 with $\|x_1\|$. Then choose a unit vector $x_2 \perp \text{span}\{x_1, A_2x_1, ..., A_mx_1\}$. By assumption of n, we can choose an orthonormal set $\{x_1, ..., x_q\}$ of q vectors in $\mathbb{C}^n$ such that for $1 < r < q$, x_r is orthogonal to A_jx_i, for all $1 < i < r$ and $1 \geq j \geq m$. Let X be an $n \times q$ matrix with x_i as the ith column. Then X^*A_jX is a diagonal matrix for $1 \geq j \geq m$. By Theorem 66, we can partition the set $\{i : 1 \geq i \geq q\}$ into k disjoint subsets R_j, $1 \geq j \geq k$ such that $R = \cap_{j=1}^{k} conv\{\langle Ax_i, x_i\rangle : i \in R_j\} \neq \emptyset$. Suppose $a \in R$, then there exist non-negative numbers t_{ij}, where $1 \geq j \geq k$, $i \in R_j$ such that for all $1 \geq j \geq k$, $\sum_{i \in R_j} t_{ij}\langle Ax_i, x_i\rangle = a$. Let $y_j = \sum_{i\in R_j} \sqrt{t_{ij}} x_j$ for all $1 \geq j \geq k$. Then $\{y_1, \ldots, y_k\}$ is orthonormal and $\langle Ay_j, y_j\rangle = a$ for all $1 \geq j \geq k$. □

Remark 68. Let $\mathcal{H}$ be a n-dimensional Hilbert space with Hermitian operators $\{A_1, \ldots, A_p\}$ on $\mathcal{H}$. If $\mathbf{A} = \{A_i^*A_j\}$ is a tuple for all $1 \leq i,\ j \leq p$ and $\Lambda_k(\mathbf{A}) \neq \emptyset$ then the cardinality of $\Lambda_k(\mathbf{A})$ is $\frac{1}{2}(p^2+p)$. This is because $\{A_1, \ldots A_p\}$ are Hermitian matrices, so if $E_i^*E_j = \alpha_{ij}$ the resulting matrix entries will each be $\alpha_{ij} = \bar{\alpha}_{ji}$. We must take this into account when counting the elements of $\mathbf{A}$ as repetition is not needed. This is the same as $\binom{p}{2}$, except we must add p to account for the $A_i^*E_i$ for all $1 \leq i \leq p$. Thus,

$$\binom{p}{2} + p = \frac{p(p-1)}{2} + p = \frac{p^2+p}{2}.$$

Chapter 4

Isoclinic Subspaces

In this chapter we introduce isoclinic subspaces and discuss some of their properties to familiarize readers.

4.1 Principal Angles

The definitions in this section, extend the concept of the angular relationship between lines and planes. It is generalized first by Jordan in [24]. In [24], Jordan extended ideas and results about 2 and 3 dimensions to further understand what they might mean geometrically in arbitrary n-dimensions.

Jordan began with the more abstract angles between flats. A flat is similar to a linear subspace of a Euclidean space however, it need not pass through the origin. Thus, flats are generalizations of a linear subspace. It was Jordan's formulations within [24] that lead to the current intuition of angles between linear subspaces (all within a large vector space). Angles between subspaces are referred to as principal angles or as canonical angles.

To try and gain some intuition, we begin looking at the angle between two lines both in vector space $\mathcal{V}$ of dimension n. Firstly, let $a, b \in \mathcal{V}$, and label the angle between these two lines θ. This is pictorially described in figure 4.1. The angle between any two lines in $\mathcal{V}$ is the inverse cosine of the dot product of the two normalized vectors or $\langle a, b \rangle = \|a\|\|b\|\cos(\theta)$.

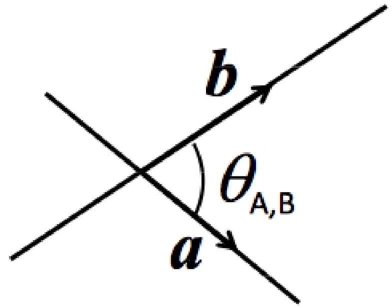

Figure 4.1: Angles Between Lines

This is the case for two lines in a vector space. Now, let's consider the simple case of two subspaces in n-dimensional vector space $\mathcal{V}$. Let S and $\mathcal{T}$ be two k-dimensional subspaces in $\mathcal{V}$, where $k \leq n$.

Let each subspace be spanned by k linearly independent vectors, $S = \text{span}\{s_1, \ldots, s_k\} \subset \mathcal{V}$ and $\mathcal{T} = \text{span}\{t_1, \ldots, t_k\} \subset \mathcal{V}$. Thus each of these spans has a basis defining their respective subspace. If we take a vector from the basis of $\mathcal{S}$, then project it onto $\mathcal{T}$, by the definition of a projection we get can define an angle between these two vectors. If we continue to do that for each linearly independent vector in their bases, we will have k angles in between subspace $\mathcal{S}$ and $\mathcal{T}$, called θ_i. These angles are called the principal angles between subspaces $\mathcal{S}$ and $\mathcal{T}$ and are in the range $0 \leq \theta_1 \leq \cdots \leq \theta_i \leq \cdots \leq \theta_k \leq \frac{\pi}{2}$, arranged in increasing order. In some literature they are called canonical angles as well. These terms will be used interchangeably throughout this work. It can be easily verified that if $\mathcal{S} = \mathcal{T}$, then the principal angles between $\mathcal{S}$ and $\mathcal{T}$ will all equal to 0. If $\mathcal{T} = \mathcal{S}^{\perp}$, then the principal angles between $\mathcal{S}$ and $\mathcal{T}$ will all equal to $\frac{\pi}{2}$ as these to subspaces are orthogonal to each other. A visualization of this concept is found below from [22].

Now that it is clear what a principal angle is, a full formalized definition is given below.

Definition 69. Let $\mathcal{V}$ and $\mathcal{W}$ be finite dimensional subspaces of a Hilbert space $\mathcal{H}$ and let $l = \min\{\dim(\mathcal{V}), \dim(\mathcal{W})\}$. Then the *canonical angles* $\{\theta_1, \ldots, \theta_l\}$ between $\mathcal{V}$ and $\mathcal{W}$ are defined as follows: the first canonical angle is the unique number $\theta_1 \in [0, \frac{\pi}{2}]$ such that

$$\cos(\theta_1) = \max\{|\langle x, y\rangle| : x \in \mathcal{V}, y \in \mathcal{W}, \|x\| = \|y\| = 1\}.$$

Let x_1 and y_1 be unit vectors in $\mathcal{V}$ and $\mathcal{W}$ for which the previous maximum is attained. Then we define the second canonical angle as the unique number $\theta_2 \in [0, \frac{\pi}{2}]$ such that

$$\cos(\theta_2) = \max\{|\langle x, y\rangle| : x \in \mathcal{V} \cap \{x_1\}^\perp, y \in \mathcal{W} \cap \{y_1\}^\perp, \|x\| = \|y\| = 1\}.$$

For each $k \leq l$, similarly choose unit vectors $x_2, \ldots x_{k-1}$ and $y_2, \ldots y_{k-1}$ in $\mathcal{V}$ and $\mathcal{W}$ respectively, in each case where the previous maximum is attained. Then θ_k is taken to be the unique number such that $\cos(\theta_k)$ is equal to the maximum of $|\langle x, y\rangle|$ with unit vectors $x \in \mathcal{V} \cap \{x_1, \ldots, x_{k-1}\}^\perp$ and $y \in \mathcal{W} \cap \{y_1, \ldots, y_{k-1}\}^\perp$.

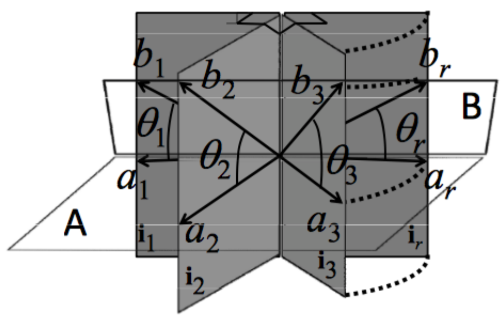

Figure 4.2: Principal Angles Between Subspaces [22]

Example 70. This example will be set in the 4-dimensional vector space $\mathbb{C}^4$. Let $u_1, u_2 \in \mathcal{S}$ and $v_1, v_2 \in \mathcal{T}$, where $\mathcal{S}$ and $\mathcal{T}$ are 2-dimensional linear subspaces of $\mathbb{C}^4$. Let u_1, u_2 span subspace $\mathcal{S}$ and v_1, v_2 span subspace $\mathcal{T}$. Suppose that $|a| < |b|$. Explicitly, these vectors defined as $u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ and $v_1 = \frac{1}{\sqrt{1+a^2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ a \end{pmatrix}$, $v_2 = \frac{1}{\sqrt{1+b^2}} \begin{pmatrix} 0 \\ 1 \\ b \\ 0 \end{pmatrix}$.

The smallest principal vector will be calculated first so since $|a| < |b|$, the principal vectors u_2 and v_2 will be used to calculate θ_1.

$$\cos(\theta_1) = \frac{\langle u_2, v_2 \rangle}{\|u_2\|\|v_2\|} = \frac{1}{\sqrt{1+b^2}} \implies \theta_1 = \cos^{-1}(\frac{1}{\sqrt{1+b^2}}).$$

For θ_2,

$$\cos(\theta_2) = \frac{\langle u_1, v_1 \rangle}{\|u_1\|\|v_1\|} = \frac{1}{\sqrt{1+a^2}} \implies \theta_2 = \cos^{-1}(\frac{1}{\sqrt{1+a^2}}).$$

Next, is an equation used to solve for principal angles efficiently by way of singular value decomposition, a well known numerical method. This result is originally stated in [7]. This is especially useful for subspaces in which $m \times n$ matrices are used to represent it. Computationally, it is a very useful theorem. Here it will be stated for $\mathbb{C}^n$; it was originally stated for $\mathbb{R}^n$. The result and proofs are essentially identical in both cases.

Theorem 71. *Let $\mathcal{S}$ and $\mathcal{T}$ be subspaces of a Hilbert Space $\mathcal{H}$ with dimensions m and n respectively, and let $Q_{\mathcal{S}}$ and $Q_{\mathcal{T}}$ be matrices whose column vectors are the elements of orthonormal bases of $\mathcal{S}$ and $\mathcal{T}$ respectively. Then the cosine of the canonical angles are the singular values of the matrix $Q_{\mathcal{S}}^* Q_{\mathcal{T}}$:*

$$\cos(\theta_k) = \sigma_k^{\downarrow}(Q_{\mathcal{S}}^* Q_{\mathcal{T}}),$$

for all $k = 1, ..., l = \min\{m, n\}$.

Example 72. Using the same subspaces of $\mathbb{C}^n$ from the example above, Theorem 71 shows that the principal angles can also be calculated in this way. Firstly, according to Theorem 71, $Q_{\mathcal{A}}$ and $Q_{\mathcal{B}}$ must be matrices whose column vectors are elements of the bases of subspace $\mathcal{A}$ and $\mathcal{B}$ respectively. So, $Q_{\mathcal{A}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $Q_{\mathcal{B}} = \begin{pmatrix} \frac{1}{\sqrt{1+a^2}} & 0 \\ 0 & \frac{1}{\sqrt{1+b^2}} \\ 0 & \frac{b}{\sqrt{1+b^2}} \\ \frac{a}{\sqrt{1+a^2}} & 0 \end{pmatrix}$.

Now,

$$Q_{\mathcal{A}}^{*}Q_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{1+a^2}} & 0 \\ 0 & \frac{1}{\sqrt{1+b^2}} \\ 0 & \frac{b}{\sqrt{1+b^2}} \\ \frac{a}{\sqrt{1+a^2}} & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1+a^2}} & 0 \\ 0 & \frac{1}{\sqrt{1+b^2}} \end{pmatrix}.$$

Taking the singular value decomposition of the 2×2 matrix, the singular values are arranged in decreasing order in a column vector,

$$\sigma^{\downarrow} = \begin{pmatrix} \frac{1}{\sqrt{1+a^2}} \\ \frac{1}{\sqrt{1+b^2}} \end{pmatrix}.$$

Now that the singular values are in hand, the principal angles are shown to be the same as the above example,

$$\cos(\theta_2) = \frac{1}{\sqrt{1+a^2}} \implies \theta_1 = \cos^{-1}(\frac{1}{\sqrt{1+a^2}})$$

$$\cos(\theta_1) = \frac{1}{\sqrt{1+b^2}} \implies \theta_2 = \cos^{-1}(\frac{1}{\sqrt{1+b^2}}).$$

4.2 Isoclinic Subspaces

Now that the notion of principal angles are formalized, it is time to discuss the main topic of this chapter, isoclinic subspaces. As a note to avoid confusion, a word that will be used interchangeably with subspace in this work is "n-plane". The number "n" is used to denote the dimension of the plane (subspace). To stay consistent with the works that are referenced here, many definitions and statements will use n-planes as the preferred terminology instead of subspace.

The following definition of two n-planes being isoclinic is from [53].

Definition 73. Suppose $\mathcal{A}$ and $\mathcal{B}$ are two n-planes in $\mathbb{F}^{2n}$ (where $\mathbb{F}$ is some field $\mathbb{R}$, $\mathbb{C}$). $\mathcal{A}$ and $\mathcal{B}$ are said to be *isoclinic* (with each other) if the angle between any non-zero vector in $\mathcal{A}$ and its projection on $\mathcal{B}$ is the same for every non-zero vector in $\mathcal{B}$. If that angle is θ, then it is said that $\mathcal{A}$ and $\mathcal{B}$ are isoclinic at angle θ.

This definition is of course consistent with above because the angle θ is actually the principal angles between subspaces $\mathcal{A}$ and $\mathcal{B}$. Comparing Definition 69 and 73 this becomes more apparent.

Proposition 74. *Suppose $\mathcal{A}$ and $\mathcal{B}$ are two n-planes in $\mathbb{F}^{2n}$. The two following statements hold:*

i) The relation of two n-planes being isoclinic is reflexive and symmetric (but not transitive) and

ii) if $\mathcal{A}$ and $\mathcal{B}$ are isoclinic, so are $\mathcal{A}, \mathcal{B}^{\perp}$; $\mathcal{A}^{\perp}, \mathcal{B}$; $\mathcal{A}^{\perp}, \mathcal{B}^{\perp}$.

Proof. $i)$: Firstly, $\sim$, in this proof, will be used to denote the relation of two n-planes being isoclinic. Showing $\sim$ is reflexive is not too difficult, since $\mathcal{A} \cap \mathcal{A} = \mathcal{A}$, all the principal angles between $\mathcal{A}$ and $\mathcal{A}$ will be equal to zero, thus $\mathcal{A}$ is isoclinic to $\mathcal{A}$ and $\sim$ is reflexive.

Now suppose $\mathcal{A} \sim \mathcal{B}$, then $\mathcal{A}$ is isoclinic to $\mathcal{B}$ at θ. Projecting an arbitrary non-zero vector in $\mathcal{B}$ onto $\mathcal{A}$, this will give the same angles as projecting an arbitrary non-zero vector in $\mathcal{A}$ onto $\mathcal{B}$. Thus, $\mathcal{B}$ is isoclinic to $\mathcal{A}$ at θ and $\sim$ is symmetric.

Lastly, suppose $\mathcal{A}$ is isoclinic to $\mathcal{B}$ at θ and $\mathcal{B}$ is isoclinic to $\mathcal{C}$ at γ. Then $\mathcal{A}$ is not necessarily isoclinic to $\mathcal{C}$. To illustrate this, suppose in $\mathbb{C}^5$, subspace $\mathcal{A} = \text{span}\{e_1, e_2\}$, $\mathcal{B} = \text{span}\{e_3, e_4\}$ and $\mathcal{C} = \text{span}\{e_1, e_5\}$. It can easily be shown that subspace $\mathcal{A}$ and $\mathcal{C}$ are not isoclinic, thus $\sim$ is not a transitive relation.

$ii)$: Suppose $\mathcal{A}$ is isoclinic to $\mathcal{B}$ at θ, it must be shown that $\mathcal{A}, \mathcal{B}^{\perp}$; $A^{\perp}, \mathcal{B}$; $\mathcal{A}^{\perp}, \mathcal{B}^{\perp}$

are all isoclinic pairs.

For $\mathcal{A}$ and $\mathcal{B}^{\perp}$, rotate every vector in $\mathcal{B}$ by $\frac{\pi}{2}$ to obtain $\mathcal{B}^{\perp}$. Now, every principal angle between $\mathcal{A}$ and $\mathcal{B}^{\perp}$ are $\frac{\pi}{2} - \theta$. Thus, $\mathcal{A}$ and $\mathcal{B}^{\perp}$ are isoclinic at $\frac{\pi}{2} - \theta$.

This is similarly the case for $\mathcal{A}^{\perp}$ and $\mathcal{B}$, except rotating the vectors in $\mathcal{A}$ by $\frac{\pi}{2}$ to obtain $\mathcal{A}^{\perp}$.

For $\mathcal{A}^{\perp}$ and $\mathcal{B}^{\perp}$, all the vectors in $\mathcal{A}$ and $\mathcal{B}$ are rotated to obtain their orthogonal complements. Since they were rotated by the same angle, they will still be isoclinic at the same value of θ. Thus, $\mathcal{A}^{\perp}$ and $\mathcal{B}^{\perp}$ are isoclinic.

□

Moving on to some set characterization of isoclinic subspaces,

Definition 75. In $\mathbb{F}^{2n}$, a set of n-planes is called an *isoclinic set of n-planes* if every two n-planes of the set are isoclinic. An isoclinic set is called a maximal set of mutually isoclinic n-planes (or maximal isoclinic set) if it is not contained in any larger isoclinic set.

The next theorem will not be proved but is an interesting result about the dimension of maximal sets of mutually isoclinic n-planes from Wong, Radon, and Adams [53, 42, 1].

Definition 76. If n is any positive integer, we can write it, uniquely, as

$$n = m2^{c+4d},$$

where m is an odd integer, and c and d are non-negative integers with $0 \leq c \leq 3$. We define the function $p : \mathbb{N} \to \mathbb{N}$

$$p(n) = 2^c + 8d.$$

Theorem 77. *The largest maximal isoclinic set (or sets) in each $\mathbb{F}^{2n}$ are of dimension $p(n)$.*

4.3 Constructing Isoclinic Subspaces

In order to construct isoclinic subspaces we will employ some results from separate papers. Primarily, [50, 53, 23] will be used. The exact results will be specified in greater detail below.

It may be useful for other researchers for all these results to be in one place in this section as opposed to the separate papers where they can be found. Wong uses an uncommon method of notation to describe n-planes in his works. In this thesis it will be changed to a traditional notation method along with proving this for the complex plane $\mathbb{C}^n$ to stay consistent with later results. Wong's outlines a proof in [50] and [53] of how to construct maximal sets of mutually isoclinic n-planes in $\mathbb{R}^{2n}$. Here however, it is more relevant to explain the $\mathbb{C}^{2n}$ analogue, which is close to identical. To begin, the n-planes that will be used will be defined. An n-plane through the origin is defined as $x_1 = Ax$ where $x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$ and $x_1 = \begin{pmatrix} x^{n+1} \\ \vdots \\ x^{2n} \end{pmatrix}$. Note that x and x_1 are $n \times 1$ matrices, while A is a $n \times n$ matrix with elements found in the complex field. Suppose there is another n-plane in the larger space as well, let this n-plane be defined as $x_1 = Bx$. We begin first with the simpler case, n-planes $\mathcal{A}$ and $\mathcal{B}$ in $\mathbb{C}^4$. A and B are the matrices that represent $\mathcal{A}$ and $\mathcal{B}$ respectively in $\mathbb{C}^4$. From what we know of principal angles thus far, it is clear that there is a maximum and minimum angle between these two subspaces.

Now, the rectangular coordinates for an n-plane is (x^1, x^2, x^3, x^4) in $\mathbb{C}^4$. The matrix form is $y = Ax$, where A is a 2×2 matrix that represents a 2-plane in $\mathbb{C}^4$. A lemma from [52] describes what a 2-plane is in relation to the 2-plane $x = 0$.

Lemma 78. *Let $(x, y) = (x^1, x^2, x^3, x^4)$ be a system of rectangular coordinates in $\mathbb{C}^n$, and $\boldsymbol{0}^\perp$, the orthogonal complement of the 0 subspace, be the 2-plane with the equation $x = 0$. Then a 2-plane $\mathcal{A}$ in $\mathbb{C}^4$ can be represented by a matrix equation of the form $y = Ax$ if and only if $\mathcal{A} \cap \boldsymbol{0}^\perp = \{0\}$. This means that $y = Ax$ if and only if $\mathcal{A}$ contains no nonzero vector orthogonal to the 2-plane $\boldsymbol{0} : y = 0$.*

Proof. The forward direction is trivial since the only value n-plane $\mathcal{A}$ and $\mathbf{0}^\perp$ share is 0. However, for the converse, suppose $\mathcal{A} \cap \mathbf{0}^\perp = \{0\}$. Since A represents the 2-plane $\mathcal{A}$ with the 2×1 matrices x and y then we can represent it with the matrix form $Cx + Dy = 0$ where the matrices C and D are 2×2 matrices. The 2-plane $\mathbf{0}^\perp$ has the equation $x = 0$

and our assumption $\mathcal{A} \cap \mathbf{0}^{\perp} = \{0\}$ means that $Cx + Dy = 0$ only has one trivial solution, $(x, y) = (0, 0)$. Thus D is invertible and $y = -xD^{-1}C$. Letting $A = D^{-1}C$, then $y = Ax$. □

A fairly obvious result is the inner product of the vector in A and its projection in an n-plane $\mathcal{B}$. It will not be proved here but details can be found in [52].

Lemma 79. *Let (u, Au) be a vector in the 2-plane $\mathcal{A}$ and (v, Bv) a vector of the 2-plane $\mathcal{B}$, then*

$$< u, v >= u^*(I + A^*B)v.$$

Moving on from 2-planes in $\mathbb{C}^n$, in [50, 53], there is an equation that is used to determine if n-planes $\mathcal{A}$ and $\mathcal{B}$ are actually isoclinic to each other. That equation is:

$$(I + A^*B)(I + B^*B)^{-1}(I + B^*A) = \rho^2(I + A^*A). \tag{4.1}$$

As mentioned before [50] deals with $\mathbb{R}^{2n}$, so the Hermitian adjoint '$*$' on the matrices is just the transpose. Again they are changed to represent the complex case. It is first beneficial to show where this matrix formula stems from. In [23], Hoggar states a proposition that will be used to show this. As a note $\mathbb{H}$ below denotes the Quaternions. The Quaternions form a non-commutative division algebra that extends the complex number system to 4 dimensions. There are many uses of quaternions, especially for representing rotations in 3 dimensions. In quantum mechanics there are some uses for quaternions but often are a difficult algebra to work with. For further information on quaternions these references are provided, [2, 21].

Proposition 80. *Let $\mathbb{F} = \mathbb{R}, \mathbb{C}, \mathbb{H}$,*

a) $M \in \mathbb{F}(r)$ (where $\mathbb{F}(r)$ is the $r \times r$ matrices over $\mathbb{F}$) is the orthogonal projection onto the range if and only if $M^2 = M = M^*$

b) Two n-planes in $\mathbb{F}^n$ with projections P, Q are isoclinic with parameter γ if and only if $PQP = \gamma P$.

Part b) is of most interest here. The reason for this is that it implies the isoclinic matrix equation condition found in [50, 53] as we will see.

Also, a result is used from [53].

Lemma 81. *Let $\mathcal{A} : y = Ax$ and $\mathcal{B} : y = Bx$ be any two n-planes in R^{2n}. Then*

a) the projection of the vector $(u, Au) \in \mathcal{A}$ on $\mathcal{B}$ is the vector (v, Bv), where $v = (I + B^T B)^{-1}(I + A^T B)u$;

b) $\mathcal{A}$ and B are orthogonal if and only if the matrix $I + A^T B = 0$.

Proof. Since the vector $(v - u, Bv - Au)$ is orthogonal to every vector (w, Bw) of $\mathcal{B}$,

$$\begin{aligned}
\langle (w, Bw), (v - u, Bv - Au) \rangle &= 0 \\
w^T(v - u) + w^T B^T (Bv - Au) &= 0 \\
w^T((v - w) + B^T(Bv - Au)) &= 0 \\
v - u + B^T Bv - B^T Au &= 0 \\
(I + B^T B)v &= (I + B^T A)u.
\end{aligned}$$

Since $I + B^T B$ is positive semi definite,

$$v = (I + B^T B)^{-1}(I + A^T B)u. \tag{4.2}$$

□

Now that the projections of these vectors are defined, the angle between a vector in $\mathcal{A}$ and its projection in $\mathcal{B}$ can be written as a continuous function on the set of unit vectors in the n-plane $\mathcal{A}$ [52].

Theorem 82. *The angle θ between any nonzero vector (u, Au) of the n-plane $\mathcal{A}$ and its projection in the n-plane $\mathcal{B}$ is given by the equation*

$$\cos^2(\theta) = \frac{u^*(I + B^*A)(I + B^*B)^{-1}(I + A^*B)u^*}{u(I + A^*A)u},\ 0 \leq \theta \leq \frac{\pi}{2}. \tag{4.3}$$

θ is called the principal angle between the two n-planes $\mathcal{A}$ and $\mathcal{B}$.

Proof.

$$\cos^2(\theta) = \frac{\|v\|^2}{\|u\|^2} \tag{4.4}$$

By Lemma 81 and Lemma 79, $\|u\|^2 = u(I + A^*A)u^*$ and

$$\begin{aligned}
\|v\|^2 &= v^*(I + B^*B)v \\
&= [u^*(I + B^*A)(I + B^*B)^{-1}](I + B^*B)[(I + B^*B)^{-1}(I + A^*B)u] \\
&= u^*(I + B^*A)(I + B^*B)^{-1}(I + A^*B)u.
\end{aligned}$$

By subbing this into (4.4), this theorem is proved. □

This formula is a generalization of the method used in the above examples to calculate the principal angles of two subspaces.

Another definition from [52] follows.

Definition 83. If θ is an angle between two n-planes $\mathcal{A}$ and $\mathcal{B}$ and u is any nonzero vector of $\mathcal{A}$ which makes an angle θ with its projection in $\mathcal{B}$, then we call u an *angle-vector of a $\mathcal{A}$ rel $\mathcal{B}$* associated with the angle θ.

Proposition 84. *Let $y = Ax$ define an n-plane $\mathcal{A}$ and $y = Bx$ define an n-plane $\mathcal{B}$ both in vector space $\mathcal{V}$. $\mathcal{A}$ and $\mathcal{B}$ are isoclinic if and only if*

$$(I + A^*B)(I + B^*B)^{-1}(I + B^*A) = \rho^2(I + A^*A)$$

(Equation 4.1) is satisfied for $\mathcal{A}$ and $\mathcal{B}$.

Proof. Assume $u \in \mathcal{A}$ and $v \in \mathcal{B}$. Let P be a projection onto n-plane $\mathcal{A}$ and Q be a projection onto n-plane $\mathcal{B}$. Now, using $PQP = \gamma P$ from [23], we can derive Equation (4.1). From Lemma 81 we know that the projection of a vector on $\mathcal{A}$ is $u = (I + A^*A)^{-1}(I + A^*B)v$. and similarly the projection of a vector on $\mathcal{B}$ is $v = (I + B^*B)^{-1}(I + B^*A)u$. Thus the projectors P and Q are:

$$P = (I + A^*A)^{-1}(I + A^*B) \text{ and } Q = (I + B^*B)^{-1}(I + B^*A).$$

Multiplying P and Q together we will end up with

$$PQ = (I + A^*A)^{-1}(I + A^*B)(I + B^*B)^{-1}(I + B^*A) = \gamma I$$
$$(I + A^*B)(I + B^*B)^{-1}(I + B^*A) = \gamma(I + A^*A).$$

Now that the we almost have the matrix equation (4.1), we follow the convention from the original paper [50], since γ actually is the same as ρ^2:

$$(I + B^*A)(I + B^*B)^{-1}(I + A^*B) = \rho^2(I + A^*A).$$

For the converse of this proof it is the same steps but in the opposite direction so it will be omitted. □

Now that equation (4.1) has been derived, we are able to continue constructing maximal sets of mutually isoclinic n-planes. In [50, 53], Wong parameterizes the n-plane $\mathcal{A}$ described by $x_1 = Ax$ by defining $A = \lambda_0 I + \lambda_1 B_1 + \cdots + \lambda_q B_q$, where λ_i are real scalar parameters and $\{B_1, \ldots, B_q\}$ is a set of complex $n \times n$ matrices. It is important that λ_i's are real in the complex case otherwise this construction will not work. An important restriction on the set $\{B_1, \ldots, B_q\}$ that they satisfy the following equations:

$$B_i + B_i^* = 0 \tag{4.5}$$

$$B_i^2 = -I \tag{4.6}$$

$$B_iB_j + B_jB_i = 0, i, j = 1, 2, \ldots; i \neq j. \tag{4.7}$$

This matrix arrangement is called the maximal solution of the Hurwitz matrix equations. Wong states in [50] that these matrix equations already have a set of solutions hence, they are ideal when creating isoclinic subspaces.

Now, for n-plane $\mathcal{B}$, we have the equation $x_1 = Bx$ or more specifically $B = \mu_0 I + \mu_1 B_1 + \cdots + \mu_q B_q$, with $\mu_i \in \mathbb{R}$.

Now in this special case with n-planes arranged with the restrictions (4.5), (4.6) and (4.7), it will be shown that equation (4.1) is satisfied. This will be done by showing both sides of the equation are scalar multiples of the identity. This will in turn construct our arbitrary isoclinic n-planes $\mathcal{A}$ and $\mathcal{B}$.

Since we know that the left hand side of (4.1) is a multiple of the identity, we know that each term will commute with each other. So, $(I + BB^*)^{-1}$ is moved to the right hand side of the equation and $(I + B^*A)(I + A^*B) = I + A^*B + B^*A + B^*AA^*B$ will be calculated to be a multiple of the identity. In order to show this A^*A, B^*B, $A^*B + B^*A$ will be calculated and shown to be multiples of the identity.

Firstly,

$$A^*A = (\lambda_0 I + \lambda_1 B_1^* + \cdots + \lambda_q B_q^*)(\lambda_0 I + \lambda_1 B_1 + \cdots + \lambda_q B_q),$$

so, there will be 3 cases;

i) When $\lambda_k \lambda_k$ are the coefficients for all k:

Using (4.5) and (4.6) for the sum attained, this will give us a scalar multiple of the identity matrix,

$$\begin{aligned}\lambda_0^2 I + \lambda_1^2 B_1^* B_1 + \cdots + \lambda_q^2 B_q^* B_q &= \lambda_0^2 I + \lambda_1^2(-B_1)B_1 + \cdots + \lambda_q^2(-B_q)B_q \\ &= \lambda_0^2 I - \lambda_1^2 B_1^2 + \cdots + \lambda_q^2 B_q^2 \\ &= \lambda_0^2 I - \lambda_1^2(-I) + \cdots + \lambda_q^2(-I) \\ &= (\lambda_0^2 + \lambda_1^2 + \cdots + \lambda_q^2) I.\end{aligned}$$

ii) When $\lambda_0\lambda_k$ or $\lambda_k\lambda_0$ are the coefficients for all $k \neq 0$:

Using (4.5) it is easy to see that this sum turns out to be zero,

$$\begin{aligned}&\lambda_0\lambda_1 B_1 + \cdots + \lambda_0\lambda_q B_q + \lambda_1\lambda_0 B_1^* + \cdots + \lambda_q\lambda_0 B_q^* \\ &= \lambda_0\lambda_1(B_1^* + B_1) + \lambda_0\lambda_2(B_2^* + B_2) + \cdots + \lambda_0\lambda_q(B_q^* + B_q) \\ &= \lambda_0\lambda_1((-B_1) + B_1) + \lambda_0\lambda_2((-B_2) + B_2) + \cdots + \lambda_0\lambda_q((-B_q) + B_q) \\ &= 0.\end{aligned}$$

iii) When $\lambda_h\lambda_k$ are the coefficients for all $h, k \neq 0$ and $h \neq k$:

The sum is $\lambda_1\lambda_2(B_2^* B_1 + B_1^* B_2) + \lambda_1\lambda_3(B_3^* B_1 + B_1^* B_3) + \cdots + \lambda_1\lambda_q(B_q^* B_1 + B_1^* B_q) + \cdots + \lambda_{q-1}\lambda_q(B_q^* B_{q-1} + B_{q-1}^* B_q)$. We can show that this sum turns out to be zero by using the general term, i.e$\lambda_k\lambda_h B_k^* B_h + \lambda_h\lambda_k B_h^* B_k$. Using (4.5) and (4.7) we can see these terms will cancel out,

$$\begin{aligned}\lambda_k\lambda_h B_k^* B_h + \lambda_h\lambda_k B_h^* B_k &= \lambda_k\lambda_h(-B_k)B_h + \lambda_h\lambda_k(-B_h)B_k \\ &= -[\lambda_k\lambda_h(B_k B_h + B_h B_k)] \\ &= -[\lambda_k\lambda_h(0)] \\ &= 0, \text{ for all } h, k \neq 0 \text{ and } h \neq k.\end{aligned}$$

Thus, $A^*A = (\lambda_0^2 + \lambda_1^2 + \cdots + \lambda_q^2)I)$, which is a multiple of the identity. This is similarly done for B^*B however, $B^*B = (\mu_0^2 + \mu_1^2 + \cdots + \mu_q^2)I$. Of course since these expressions are multiples of the identity, then AA^* and BB^* will also be equal to multiples of the identity.

Next,

$$
\begin{aligned}
A^*B + B^*A =& (\lambda_0 I + \lambda_1 B_1^* + \cdots + \lambda_q B_q^*)(\mu_0 I + \mu_1 B_1 + \cdots + \mu_q B_q) \\
&+ (\mu_0 I + \mu_1 B_1^* + \cdots + \mu_q B_q^*)(\lambda_0 I + \lambda_1 B_1 + \cdots + \lambda_q B_q).
\end{aligned}
$$

There will also be 3 cases for this calculation;

i) When $\mu_k \lambda_k$ are the coefficients for all k:

We use (4.5) and (4.6) to reduce this equation to a scalar multiple of the identity matrix,

$$
\begin{aligned}
&\lambda_0\mu_0 I + \lambda_1\mu_1 B_1^* B_1 + \cdots + \lambda_q\mu_q B_q^* B_q + \mu_0\lambda_0 I + \mu_1\lambda_1 B_1^* B_1 + \cdots + \mu_q\lambda_q B_q^* B_q \\
&\quad = 2\lambda_0\mu_0 I - 2\lambda_1\mu_1 B_1^2 - \cdots - 2\lambda_q\mu_q B_q^2 \\
&\quad = 2\lambda_0\mu_0 I + 2\lambda_1\mu_1 I + \cdots + 2\lambda_q\mu_q I \\
&\quad = (2\lambda_0\mu_0 + 2\lambda_1\mu_1 + \cdots + 2\lambda_q\mu_q) I.
\end{aligned}
$$

ii) When $\mu_0 \lambda_k$ or $\mu_k \lambda_0$ are the coefficients for all $k \neq 0$:

The sum is $\lambda_0 u_1 B_1 + \cdots + \lambda_0\mu_q B_q + \lambda_1\mu_0 B_1^* + \cdots + \lambda_q\mu_0 B_q^* + \mu_0\lambda_1 B_1 + \cdots + \mu_0\lambda_q B_q + \mu_1\lambda_0 B_1^* + \cdots + \mu_q\lambda_0 B_q^*$. Using (4.5), we are able the cancel all terms of this expression and the sum will equal to zero. This will be done on the general term,

$$\begin{aligned}\lambda_0\mu_k B_k + \lambda_k\mu_0 B_k^* + \mu_0\lambda_k B_k + \mu_k\lambda_0 B_k^* \\ &= \lambda_0\mu_k(B_k + B_k^*) + \lambda_k\mu_0(B_k + B_k^*) \\ &= \lambda_0\mu_k(B_k - B_k) + \lambda_k\mu_0(B_k - B_k) \\ &= \lambda_0\mu_k(0) + \lambda_k\mu_0(0) \\ &= 0, \text{ for all } k \neq 0.\end{aligned}$$

iii) When λ_{hk} or $\lambda_k\mu_h$ are the coefficients for all $h, k \neq 0$ and $h \neq k$:

For simplicity of this calculation the general term will be looked at. With (4.5) and (4.7) the summation will be shown to be zero.

$$\begin{aligned}\lambda_h\mu_k A_h^* A_k + \lambda_k\mu_h A_k^* A_h + \mu_h\lambda_k A_h^* A_k \mu_k\lambda_h A_k^* A_h \\ &= \lambda_h\mu_k(-A_h)A_k + \lambda_k\mu_h(-A_k)A_h + \mu_h\lambda_k(-A_h)A_k + \mu_k\lambda_h(-A_k)A_h \\ &= -[\lambda_h\mu_k(A_hA_k + A_kA_h) + \lambda_k\mu_h(A_kA_h + A_hA_k)] \\ &= -[\lambda_h\mu_k(0) + \lambda_k\mu_h(0)] \\ &= 0, \text{ for all } h, k \neq 0 \text{ and } h \neq k.\end{aligned}$$

Thus, $A^*B + B^*A$ is also a multiple of the identity. This shows $I + A^*B + B^*A + B^*AA^*B$ is also a multiple of the identity. Hence, both the right hand and left hand side of equation (4.1) are multiples of the identity matrix and arbitrary isoclinic n-planes are able to be constructed.

Some simple examples are now shown to illustrate this construction.

Example 85. Suppose $A = I$ and $B = B_k$, then

$$\begin{aligned}
&(I + B_k^* I)(I + B_k^* B_k)^{-1}(I + IB_k) = \rho^2(I + I^* I)^{-1} \\
&= (I - B_k)(I - B_k^2)^{-1}(I + B_k) \\
&= 2(I - B_k + B_k - B_k^2) \\
&= 2(I - (-I)) = \rho^2(2I) \\
&\implies 4I = \rho^2(2I) \\
&\implies \rho = \frac{1}{2}
\end{aligned}$$

Thus, these two n-planes are isoclinic at an angle of $\cos^{-1}(\frac{1}{2})$.

Example 86. Suppose $A = B_h$ and $B = B_k$, then entering these into equation 4.1:

$$\begin{aligned}
&(I + B_k^* B_h)(I + B_k^* B_k)^{-1}(I + B_h^* B_k) = \rho^2(I + B_h^* B_h)^{-1} \\
&(I - B_k B_h)(2I)^{-1}(I - B_h B_k) = \rho^2(2I) \\
&\frac{1}{2}(I - B_k B_h)(I - B_h B_k) = \rho^2(2I) \\
&\frac{1}{2}(I - B_k B_h - B_h B_k + B_k B_h B_h B_k) = \rho^2(2I) \\
&\frac{1}{2}(I - (B_k B_h + B_h B_k) + B_k (B_h)^2 B_k) = \rho^2(2I) \\
&\frac{1}{2}(I - 0 - B_k I B_k) = \rho^2(2I) \\
&\frac{1}{2}(I + I) = \rho^2(2I) \\
&\frac{1}{2}(2I) = \rho^2(2I) \\
&\frac{1}{2}I = \rho^2 I
\end{aligned}$$

This implies $\rho = \frac{1}{\sqrt{2}}$. Thus, these two n-planes are isoclinic at an angle of $cos^{-1}(\frac{1}{\sqrt{2}})$.

Chapter 5

The Relationship Between Isoclinic Subspaces and Quantum Error Correction

5.1 Introduction

The classical notions of canonical angles and isoclinic subspaces have played a role in Euclidean geometry, and in matrix and operator theory and beyond for over a century [24, 3, 7, 23, 50, 52]. On the other hand, quantum information theory is relatively new, with roots going back several decades but only emerging as a formal field of study over the past quarter century or so [40]. Quantum error correction is a fundamental subfield with aspects touching on all parts of quantum information, from theory to experiment [46, 47, 17, 5, 27, 30].

This chapter is based on the paper [31]. In it, we bring together equivalent conditions for isoclinic subspaces, including a new description based on canonical angles. We establish connections with the theory of quantum error correction, showing how quantum error correcting codes are associated with families of isoclinic subspaces. We also show how higher rank numerical ranges of matrices, originally introduced for quantum error correction purposes [14, 12, 33, 49, 38, 11, 35, 34, 32, 15], arise in the study of isoclinic subspaces.

This chapter is organized as follows. The next section includes a review of the classical notions of canonical angles and isoclinic subspaces, and we give equivalent conditions for families of subspaces to be isoclinic. In the following section we show how every quantum error correcting code and error model determines a family of isoclinic subspaces and we prove a converse for pairs of such subspaces. In the final section we show how the canonical angles for isoclinic subspaces are embedded in the structure of the higher rank numerical ranges for the corresponding orthogonal projections. We also include a pair of illustrative examples.

5.2 Canonical Angles and Isoclinic Subspaces

In Chapter 3 the classical notion of canonical angles between pairs of subspaces with Definition 69 was introduced. These are sometimes referred to as principal angles and were first formulated by Jordan [24]. Following from Definition 69, Bjorck and Golub [7] showed that the canonical angles can be characterized in terms of the singular values of the product of two matrices that encode their respective subspace. A characterization of this was stated in the previous chapter as Theorem 71. We can view the matrix $Q_{\mathcal{V}}$ in operator theoretic terms as well. If $\mathcal{V}$ is an m-dimensional subspace of $\mathcal{H}$, then $Q_{\mathcal{V}}$ is an isometry from $\mathbb{C}^m$ into $\mathcal{H}$ with range equal to $\mathcal{V}$. A consequence of this is that $Q_{\mathcal{V}}Q_{\mathcal{V}}^*$ is a matrix representation of the orthogonal projection from $\mathcal{H}$ onto $\mathcal{V}$, whereas on the other hand $Q_{\mathcal{V}}^*Q_{\mathcal{V}} = I_m$.

Definition 87. Let $\mathcal{V}$ and $\mathcal{W}$ be two m-dimensional subspaces of a Hilbert space $\mathcal{H}$, where $1 \leq m \leq \dim(\mathcal{H})$. Then $\mathcal{V}$ and $\mathcal{W}$ are said to be *isoclinic* if all m canonical angles between $\mathcal{V}$ and $\mathcal{W}$ are equal. If that angle is θ, then the subspaces are said to be *isoclinic at angle* θ. A collection of m-dimensional subspaces of a Hilbert space are said to be isoclinic if all pairs of distinct subspaces from the collection are pairwise isoclinic.

Of course, any family of mutually orthogonal subspaces are isoclinic at angle $\frac{\pi}{2}$, but there are other possibilities as well. There are a variety of useful equivalent characterizations of isoclinic subspaces, as shown in the following result.

Theorem 88. *Let $\mathcal{V}$ and $\mathcal{W}$ be two m-dimensional subspaces of a Hilbert space $\mathcal{H}$, with $m \geq 1$ and $d = \dim \mathcal{H}$. Let $P_\mathcal{V}$ and $P_\mathcal{W}$ denote the orthogonal projections onto the subspaces $\mathcal{V}$ and $\mathcal{W}$ respectively. Let $Q_\mathcal{V}$ and $Q_\mathcal{W}$ be $d \times m$ matrices whose column vectors are elements of the orthonormal bases of the subspaces $\mathcal{V}$ and $\mathcal{W}$ respectively, represented in any orthonormal basis for $\mathcal{H}$. Then the following conditions are equivalent:*

(i) *$\mathcal{V}$ and $\mathcal{W}$ are isoclinic subspaces.*

(ii) *$Q_\mathcal{V}^* Q_\mathcal{W}$ is a scalar multiple of a unitary on $\mathbb{C}^m$.*

(iii) *There exists $\lambda \geq 0$ such that*

$$P_\mathcal{V} P_\mathcal{W} P_\mathcal{V} = \lambda P_\mathcal{V} \quad \text{and} \quad P_\mathcal{W} P_\mathcal{V} P_\mathcal{W} = \lambda P_\mathcal{W}. \tag{5.1}$$

Here, $\lambda = \cos(\theta)$ where $\mathcal{V}$, $\mathcal{W}$ are isoclinic at angle θ.

(iv) *The angle between any non-zero vector in $\mathcal{V}$ and its projection on $\mathcal{W}$ is constant; in other words, $\|P_\mathcal{W} x\| \|x\|^{-1}$ is constant for $0 \neq x \in \mathcal{V}$. And the same holds true with the roles of $\mathcal{V}, \mathcal{W}$ reversed.*

Proof. The equivalence of (i) and (ii) follows from Theorem 71 above, as all the singular values of a unitary matrix are equal to one.

For $(ii) \implies (iii)$, assume $Q_\mathcal{V}^* Q_\mathcal{W}$ is a multiple of a unitary on $\mathbb{C}^m$. Then the same is true of $Q_\mathcal{W}^* Q_\mathcal{V} = (Q_\mathcal{V}^* Q_\mathcal{W})^*$. Recall from the discussion just after Theorem 71, the projections onto subspace $\mathcal{V}$ and $\mathcal{W}$ respectively have matrix representations $P_\mathcal{V} = Q_\mathcal{V} Q_\mathcal{V}^*$ and $P_\mathcal{W} = Q_\mathcal{W} Q_\mathcal{W}^*$. Since $Q_\mathcal{V}^* Q_\mathcal{W}$ is a multiple of a unitary on $\mathbb{C}^m$, we have for some $0 \leq \lambda \leq 1$, $Q_\mathcal{V}^* Q_\mathcal{W} Q_\mathcal{W}^* Q_\mathcal{V} = \lambda I_m$. Hence,

$$P_\mathcal{V} P_\mathcal{W} P_\mathcal{V} = Q_\mathcal{V} Q_\mathcal{V}^* Q_\mathcal{W} Q_\mathcal{W}^* Q_\mathcal{V} Q_\mathcal{V}^* = Q_\mathcal{V} (\lambda I) Q_\mathcal{V}^* = \lambda Q_\mathcal{V} Q_\mathcal{V}^* = \lambda P_\mathcal{V}.$$

This is similarly done for $P_\mathcal{W} P_\mathcal{V} P_\mathcal{W} = \lambda P_\mathcal{W}$. (Note that the λ obtained for $Q_\mathcal{W}^* Q_\mathcal{V}$ is the same as that for $Q_\mathcal{V}^* Q_\mathcal{W} = (Q_\mathcal{W}^* Q_\mathcal{V})^*$.)

For $(iii) \implies (ii)$, assume there exists a scalar λ such that $P_{\mathcal{V}}P_{\mathcal{W}}P_{\mathcal{V}} = \lambda P_{\mathcal{V}}$ and $P_{\mathcal{W}}P_{\mathcal{V}}P_{\mathcal{W}} = \lambda P_{\mathcal{W}}$ (necessarily $0 \leq \lambda \leq 1$). Recall $Q_{\mathcal{V}}^* Q_{\mathcal{V}} = I_m = Q_{\mathcal{W}}^* Q_{\mathcal{W}}$. Together this implies that:

$$\begin{aligned} P_{\mathcal{V}}P_{\mathcal{W}}P_{\mathcal{V}} &= \lambda P_{\mathcal{V}} \\ Q_{\mathcal{V}}Q_{\mathcal{V}}^* Q_{\mathcal{W}}Q_{\mathcal{W}}^* Q_{\mathcal{V}}Q_{\mathcal{V}}^* &= \lambda Q_{\mathcal{V}}Q_{\mathcal{V}}^* \\ (Q_{\mathcal{V}}^*)Q_{\mathcal{V}}Q_{\mathcal{V}}^* Q_{\mathcal{W}}Q_{\mathcal{W}}^* Q_{\mathcal{V}}Q_{\mathcal{V}}^*(Q_{\mathcal{V}}) &= (Q_{\mathcal{V}}^*)\lambda Q_{\mathcal{V}}Q_{\mathcal{V}}^*(Q_{\mathcal{V}}) \\ (I)Q_{\mathcal{V}}^* Q_{\mathcal{W}}Q_{\mathcal{W}}^* Q_{\mathcal{V}}(I) &= \lambda (I)(I) \\ (Q_{\mathcal{W}}^* Q_{\mathcal{V}})^* Q_{\mathcal{W}}^* Q_{\mathcal{V}} &= \lambda I. \end{aligned}$$

Thus, $Q_{\mathcal{W}}^* Q_{\mathcal{V}}$ is a multiple of a unitary on $\mathbb{C}^m$. This is similarly true for $P_{\mathcal{W}}P_{\mathcal{V}}P_{\mathcal{W}} = \lambda P_{\mathcal{W}}$ and $Q_{\mathcal{V}}^* Q_{\mathcal{W}}$.

To see $(iii) \implies (iv)$, assume there exists $0 \leq \lambda \leq 1$ such that $P_{\mathcal{V}}P_{\mathcal{W}}P_{\mathcal{V}} = \lambda P_{\mathcal{V}}$ and $P_{\mathcal{W}}P_{\mathcal{V}}P_{\mathcal{W}} = \lambda P_{\mathcal{W}}$. Let $0 \neq x = P_{\mathcal{V}}x \in \mathcal{V}$. Then as $P_{\mathcal{V}}P_{\mathcal{W}}P_{\mathcal{V}} = \lambda P_{\mathcal{V}}$, we have,

$$\lambda \|x\|^2 = \lambda \langle P_{\mathcal{V}}x, x\rangle = \langle P_{\mathcal{V}}P_{\mathcal{W}}P_{\mathcal{V}}x, x\rangle = \langle P_{\mathcal{W}}x, x\rangle = \|P_{\mathcal{W}}x\|^2.$$

Thus, $\sqrt{\lambda} = \|P_{\mathcal{W}}x\|\|x\|^{-1}$ for all $0 \neq x \in \mathcal{V}$. Similarly, from $P_{\mathcal{W}}P_{\mathcal{V}}P_{\mathcal{W}} = \lambda P_{\mathcal{W}}$, we obtain $\sqrt{\lambda} = \|P_{\mathcal{V}}x\|\|x\|^{-1}$ for all $0 \neq x \in \mathcal{W}$.

Finally for $(iv) \implies (iii)$, if $r = \|P_{\mathcal{W}}x\|\|x\|^{-1}$ for all $0 \neq x \in \mathcal{V}$, then one can follow a similar argument to that above to show $r^2 P_{\mathcal{V}} = P_{\mathcal{V}}P_{\mathcal{W}}P_{\mathcal{V}}$. □

Remark 89. We note that condition (iv) was taken as the definition of isoclinic subspaces in [23, 52], with the equivalence of (iii) and (iv) being noted without proof in [23]. The connection with canonical angles given by the equivalence of (ii) and (iii) appears to be new.

5.3 Connection with Quantum Error Correction

Error models in quantum information are described by sets of operators $\{E_i\}$ on a Hilbert space $\mathcal{H}$ associated with a given quantum system. We know from Chapter 2 that these operators must satisfy the completeness relation ($\sum_i E_i^* E_i \leq I$) to ensure $\mathcal{E}(\rho) = \sum_i E_i \rho E_i^*$ is a quantum channel. In Chapter 3 we define a *correctable* code for $\mathcal{E}$ by Definition 59. Further, the Knill-Laflamme theorem [26] frames the correctability of a code in terms of these error operators, this can again be referred to in Chapter 3, Theorem 61. These are important definitions to refer back to as they are the main results to keep in mind for this section.

We establish a correspondence between isoclinic subspaces and quantum error correcting codes in the following result. Without loss of generality we will assume the code is non-degenerate in the sense that the set of restricted error operators $\{E_i|_{\mathcal{C}}\}$ is minimal in size.

Theorem 90. *Suppose $\mathcal{C}$ is a subspace of a Hilbert space $\mathcal{H}$ that is correctable for a non-degenerate error model $\{E_i\}$. For each i, let $\mathcal{V}_i = \mathrm{Range}\,(E_i|_{\mathcal{C}})$ be the range subspace of the restriction of E_i to $\mathcal{C}$. Then $\{\mathcal{V}_i\}$ is a set of isoclinic subspaces of $\mathcal{H}$.*

Proof. We have Eqs. (3.7) satisfied for the E_i and $P_{\mathcal{C}}$. Let U_i be the partial isometries obtained through the polar decompositions of the operators $E_i P_{\mathcal{C}}$:

$$E_i P_{\mathcal{C}} = U_i |E_i P_{\mathcal{C}}| = U_i \sqrt{P_{\mathcal{C}} E_i^* E_i P_{\mathcal{C}}} = \sqrt{\alpha_{ii}} U_i P_{\mathcal{C}}.$$

Note that each $\alpha_{ii} \neq 0$ by non-degeneracy. We can thus reformulate the error correction conditions in terms of the U_i as follows:

$$P_{\mathcal{C}} U_i^* U_j P_{\mathcal{C}} = \frac{1}{\sqrt{\alpha_{ii}}}(P_{\mathcal{C}} E_i^*) \frac{1}{\sqrt{\alpha_{jj}}}(E_j P_{\mathcal{C}}) = \left(\frac{\alpha_{ij}}{\sqrt{\alpha_{ii}\alpha_{jj}}} \right) P_{\mathcal{C}}.$$

Also observe that for each i, by construction we have $P_i := U_i P_{\mathcal{C}} U_i^*$ is the projection onto the range $\mathcal{V}_i$ of $E_i P_{\mathcal{C}}$ and $P_{\mathcal{C}} = P_{\mathcal{C}} U_i^* U_i P_{\mathcal{C}}$.

Now for each pair i, j, let $\lambda_{ij} = \alpha_{ij}(\sqrt{\alpha_{ii}\alpha_{jj}})^{-1}$ and note that $\overline{\lambda_{ij}} = \lambda_{ji}$. Then we have:

$$\begin{aligned} P_i P_j P_i &= P_i U_j (P_{\mathcal{C}} U_j^* U_i P_{\mathcal{C}}) U_i^* \\ &= \lambda_{ji} P_i U_j P_{\mathcal{C}} U_i^* \\ &= \lambda_{ji} U_i (P_{\mathcal{C}} U_i^* U_j P_{\mathcal{C}}) U_i^* \\ &= \lambda_{ji} \lambda_{ij} U_i P_{\mathcal{C}} U_i^* \\ &= |\lambda_{ij}|^2 P_i. \end{aligned}$$

Similarly, $P_j P_i P_j = |\lambda_{ij}|^2 P_j$. As $\mathcal{V}_i = P_i \mathcal{H}$, it follows from Theorem 88 that the subspaces $\{\mathcal{V}_i\}$ are isoclinic. □

We present the following example of a simple error model to illustrate this result.

Example 91. Consider a two-qubit error model describing a bit flip on the first qubit with the probability of some fixed $0 < p < 1$. We can formulate this mathematically by taking $|ij\rangle = |i\rangle \otimes |j\rangle$, $i, j = 0, 1$, as a fixed orthonormal basis for $\mathbb{C}^4 = \mathbb{C}^2 \otimes \mathbb{C}^2$. Then if we let X be the Pauli bit flip operator ($X|0\rangle = |1\rangle$, $X|1\rangle = |0\rangle$), we can define $X_1 = X \otimes I_2$ and the error model as a map on two-qubit density operators is given by:

$$\mathcal{E}(\rho) = (1-p)\rho + p\, X_1 \rho X_1^*.$$

Here the error operators are $E_1 = \sqrt{1-p}\, I_4$ and $E_2 = \sqrt{p} X_1$.

Now define two subspaces of $\mathbb{C}^4$ as follows: $\mathcal{C}_1 = \text{span}\{|00\rangle, |11\rangle\}$ and $\mathcal{C}_2 = \text{span}\{|10\rangle, |01\rangle\}$. Let P_1, P_2 be the corresponding projections. Then $\mathcal{C}_1$ (and similarly $\mathcal{C}_2$) is a correctable code for $\mathcal{E}$, with $\mathcal{C}_1$, $\mathcal{C}_2$ the relevant family of subspaces as in the theorem, and in this case the matrix $\alpha = (\alpha_{ij})$ satisfies $\alpha_{11} = 1 - p$, $\alpha_{22} = p$, $\alpha_{12} = \alpha_{21} = 0$. So here the canonical angles are both equal to $\theta = \frac{\pi}{2}$ (indeed we have $P_1 P_2 = 0 = P_2 P_1$), and the subspaces are isoclinic.

We can complicate things slightly and obtain more interesting isoclinic subspace structure. Suppose the system is exposed to noise that induces a rotation of angle $0 < \phi < 2\pi$ to the original error model; that is, the original error operators are replaced by

$$F_1 = (\cos\phi)E_1 + (\sin\phi)E_2 \quad \text{and} \quad F_2 = (-\sin\phi)E_1 + (\cos\phi)E_2,$$

which can also be seen through the matrix relation $[F_1\ F_2] = [E_1\ E_2]U$ where U is the rotation matrix $U = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$.

The Knill-Laflamme conditions show that correctable codes are the same for error models whose operators are linear combinations of each other, hence $\mathcal{C}_1$ is correctable for $\{F_1, F_2\}$. Indeed, here we have, with $c = \cos\phi$, $s = \sin\phi$,

$$\begin{aligned} P_{\mathcal{C}} F_1^* F_1 P_{\mathcal{C}} &= (c^2(1-p) + s^2 p)P_{\mathcal{C}} \\ P_{\mathcal{C}} F_2^* F_2 P_{\mathcal{C}} &= (s^2(1-p) + c^2 p)P_{\mathcal{C}} \end{aligned}$$

and

$$P_{\mathcal{C}} F_1^* F_2 P_{\mathcal{C}} = \big(cs(2p-1)\big)P_{\mathcal{C}} = P_{\mathcal{C}} F_2^* F_1 P_{\mathcal{C}}.$$

One can check that the unitary U factors through to give the new error correction coefficient matrix as $\alpha' = U^*\alpha U$. Moreover, the isoclinic angle θ is computed from the proof of Theorem 90 in terms of the rotation ϕ and probability p as follows:

$$\theta = \cos^{-1}\Big(\frac{|cs(2p-1)|^2}{(c^2(1-p)+s^2p)(s^2(1-p)+c^2p)}\Big).$$

See the figure below for a 3-space depiction of $\theta \in [0, \frac{\pi}{2}]$ as it depends on $0 \leq p \leq 1$ and $0 \leq \phi \leq 2\pi$.

There is at least a partial converse of the above theorem given as follows.

Proposition 92. *Let $\mathcal{H}$ be a Hilbert space and let $\{P_1, P_2\}$ be a pair of projections on $\mathcal{H}$ associated with two m-dimensional isoclinic subspaces. Then each of the subspaces $P_i\mathcal{H}$ is correctable for the error model $\{\frac{1}{\sqrt{2}}P_i\}_{i=1}^2$.*

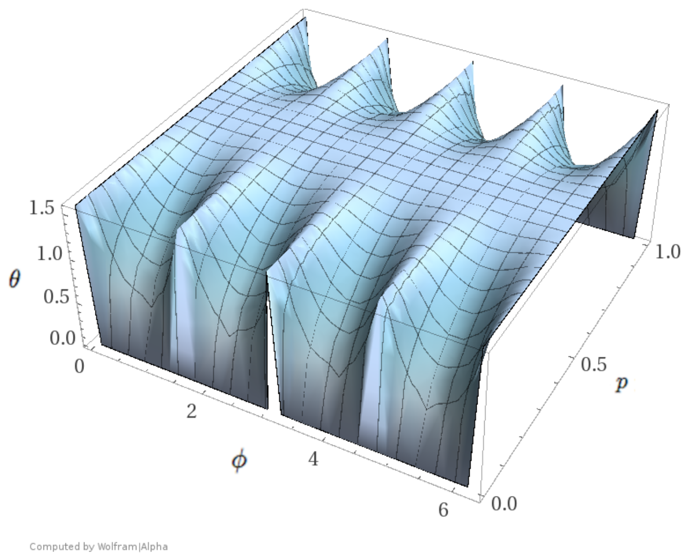

Figure 5.1: The dependence of θ on p, ϕ derived in Example 91.

Proof. The projections P_1, P_2 satisfy the isoclinic identities Eq. (5.1), with say $P_iP_jP_i = \lambda P_i$ for $i \neq j \in \{1, 2\}$. Hence we have

$$P_1 P_1^* P_2 P_1 = P_1 P_2 P_1 = \lambda P_1.$$

Similar identities hold for each product $P_i P_j^* P_k P_i$, $i, j, k = 1, 2$, and the result follows from the quantum error correction conditions of Eq. (3.7). □

Motivated by this result, we finish this section by presenting an example of a pair of isoclinic subspaces that arise in matrix theory and Euclidean geometry, found in Wong's original monograph [52].

Before the example, an alternate derivation of the isoclinic subspace, Equation 4.1, will be presented to use in this example.

In general, we focus on the following set:

$$S = \{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^4 \text{ such that } y = Mx \}, \text{ where}$$

$\{\begin{pmatrix} x \\ Mx \end{pmatrix} : x \in \mathbb{C}^2\}$ and the orthogonal complement is$\{\begin{pmatrix} -M^*x \\ z \end{pmatrix} : z \in \mathbb{C}^2\}$.

Firstly, to show these subspaces are isoclinic and determine the angle at which they are isoclinic to each other. We will have to derive an equation that will accomplish this. Adding the orthogonal complement component this is possible:

$$\begin{pmatrix} y^* & z^* \end{pmatrix} \begin{pmatrix} x \\ Mx \end{pmatrix} = 0$$

$$y^*x + z^*Mx = 0$$

$$(y^*M^*z)^*x = 0$$

$$y + M^*z = 0$$

$$y = -M^*z$$

To find the projection onto each of these subspaces:

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} x \\ Mx \end{pmatrix} + \begin{pmatrix} -M^*z \\ z \end{pmatrix}$$

Turning the vector form into equation form we get:

$$v_1 = x - M^*z \tag{5.2}$$

$$v_2 = Mx + z \tag{5.3}$$

Multiply equation (5.3) by M^* from the left,

$$M^*v_2 = M^*Mx + M^*z, \tag{5.4}$$

now, add equation (5.2) and (5.4),

$$\begin{aligned} v_1 + M^*v_2 &= x - M^*z + M^*Mx + M^*z \\ v_1 + M^*v_2 &= x + M^*Mx \\ x &= (I + M^*M)^{-1}(v_1 + M^*v_2). \end{aligned}$$

Note that $(I + M^*M)$ because M^*M is positive semi-definite, $(I + M^*M)$ will also not have any zero eigenvalues and thus will have an inverse.

Finally, if we map from $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ to $\begin{pmatrix} x \\ Ax \end{pmatrix}$ then

$$\begin{pmatrix} (I + M^*M)^{-1} & (I + M^*M)^{-1}M^* \\ M(I + M^*M)^{-1} & M(I + M^*M)^{-1}M^* \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} x \\ Mx \end{pmatrix}.$$

Observe we are also able to write this in a different form,

$$\begin{pmatrix} I \\ M \end{pmatrix} (I + M^*M) \begin{pmatrix} I & M^* \end{pmatrix}. \tag{5.5}$$

This is the general form of these matrices' projections onto their respective subspace. It is easy to see that these are projections by checking their idempotency. This will be shown in the following example using numerical matrices.

Continuing on, if the matrices describing the subspaces are inputted into equation (5.5) and multiplied by each other then we obtain,

$$\begin{pmatrix} I & A^* \end{pmatrix} \begin{pmatrix} I \\ B \end{pmatrix} (I + B^*B)^{-1} \begin{pmatrix} I & B^* \end{pmatrix} \begin{pmatrix} I \\ A \end{pmatrix} (I + A^*A)^{-1} = \lambda, \text{ for some } \lambda.$$

With some simple vector multiplication we get the isoclinic equation from [51],

$$(I + A^*B)(I + B^*B)^{-1}(I + B^*A) = \sigma(I + A^*A) \tag{5.6}$$

where $\sigma = \cos^2\theta$, the angle of isoclinicity. It is easy to see this is equivalent to Hoggar's equation of projections onto isoclinic subspaces from [23]

$$PQP = \lambda P. \tag{5.7}$$

Example 93. Given a 2×2 complex matrix M, one can consider the *graph of* M which is the subspace of $\mathbb{C}^4$ given by:

$$\mathcal{V}_M := \left\{ \begin{pmatrix} x \\ Mx \end{pmatrix} : x \in \mathbb{C}^2 \right\}.$$

The orthogonal complement of $\mathcal{V}_M$ inside $\mathbb{C}^4$ is given as follows:

$$\mathcal{V}_M^{\perp} := \left\{ \begin{pmatrix} -M^*x \\ x \end{pmatrix} : x \in \mathbb{C}^2 \right\}.$$

The following subspaces in $\mathbb{C}^4$ are described by the matrix equations:

$$y = Ax \text{ such that } A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and} \tag{5.8}$$

$$y = Bx \text{ such that } B = \begin{bmatrix} \frac{\sqrt{3}+1}{\sqrt{3}-1} & 0 \\ 0 & 0 \end{bmatrix}. \tag{5.9}$$

Now using equation (5.6) derived above, we are able to check if these subspaces corresponding to matrices A and B are isoclinic. So inputting A and B into equation (5.6):

$$(\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}+\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}\frac{\sqrt{3}+1}{\sqrt{3}-1} & 0\\0 & 0\end{bmatrix})(\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}+\begin{bmatrix}\frac{\sqrt{3}+1}{\sqrt{3}-1} & 0\\0 & 0\end{bmatrix}^2)^{-1}(\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}+\begin{bmatrix}\frac{\sqrt{3}+1}{\sqrt{3}-1} & 0\\0 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix})$$

$$=(\begin{bmatrix}\frac{2\sqrt{3}}{\sqrt{3}-1} & 0\\0 & 1\end{bmatrix})(\begin{bmatrix}\frac{8}{(\sqrt{3}-1)^2} & 0\\0 & 0\end{bmatrix})^{-1}(\begin{bmatrix}\frac{\sqrt{3}+1}{\sqrt{3}-1} & 0\\0 & 1\end{bmatrix})=\sigma(I+A^*A)^{-1}$$

$$\frac{1}{8}\begin{bmatrix}2\sqrt{3}(\sqrt{3}+1) & 0\\0 & 8\end{bmatrix}=\sigma(\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}+\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}^2)$$

$$\frac{1}{8}\begin{bmatrix}8 & 0\\0 & 8\end{bmatrix}=\sigma(2I)\implies I=2\sigma I\implies \sigma=\frac{1}{2}$$

Since $\sigma=\cos^2\theta$,

$$\cos^2\theta=\frac{1}{2}\implies\theta=\frac{\pi}{4}$$

Thus, these two subspaces are isoclinic at the angle of $\theta=\frac{\pi}{4}$. Now that we have equation (5.5) we can create projection matrices and show that they also satisfy the Quantum Error Correction Conditions.

For $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$:

$$\begin{pmatrix} I \\ A \end{pmatrix} (I + A^*A)^{-1} \begin{pmatrix} I & A^* \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

$$= P_A$$

For $B = \begin{bmatrix} \frac{\sqrt{3}+1}{\sqrt{3}-1} & 0 \\ 0 & 0 \end{bmatrix}$:

$$\begin{pmatrix} I \\ B \end{pmatrix} (I + B^*B)^{-1} \begin{pmatrix} I & B^* \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\sqrt{3}+1}{\sqrt{3}-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2-\sqrt{3}}{4} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \frac{\sqrt{3}+1}{\sqrt{3}-1} & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} \frac{2-\sqrt{3}}{4} & 0 & \frac{1}{4} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{2-\sqrt{3}}{4} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= P_B$$

Simple matrix multiplication will show that P_A, P_B are in fact projection matrices as $P_A^2 = P_A$ and $P_B^2 = P_B$. It is also worth noting that these projections are partial isometries.

Now that we have P_A, P_B and that A and B are isoclinic to each other equation (5.1) is satisfied for some $\lambda = \cos\theta$, i.e. $P_A P_B P_A = \lambda P_A$.

We must show that these projectors satisfy Equation (3.7), the quantum error correction conditions, but first we need a projection onto the code subspace, so let $P_{\mathcal{C}} = P_A$ then our error operators are $E_1 = P_A$ and $E_2 = P_B$.

Checking if Equation (3.7) is satisfied for all i, j:

$$P_{\mathcal{C}} E_1^* E_1 P_{\mathcal{C}} = P_A P_A^* P_A P_A = (P_A P_A)(P_A P_A) = P_A P_A = (1) P_A$$
$$P_{\mathcal{C}} E_1^* E_2 P_{\mathcal{C}} = P_A P_A^* P_B P_A = P_A P_B P_A = \lambda P_A = \cos\theta P_A$$
$$P_{\mathcal{C}} E_2^* E_1 P_{\mathcal{C}} = P_A P_B^* P_A P_A = P_A P_B P_A = \lambda P_A = \cos\theta P_A$$
$$P_{\mathcal{C}} E_2^* E_2 P_{\mathcal{C}} = P_A P_B^* P_B P_A = P_A P_B P_A = \lambda P_A = \cos\theta P_A$$

Thus the quantum error correction conditions are satisfied for the two subspaces corresponding to matrices A and B for the Hermitian matrix

$$\alpha = \begin{bmatrix} 1 & cos\theta \\ cos\theta & cos\theta \end{bmatrix}.$$

Since we have found that these subspaces are isoclinic at angle $\theta = \frac{\pi}{4}$ we get:

$$\alpha = \begin{bmatrix} 1 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Thus, it follows from Theorem 88 that $\mathcal{V}_A$ and $\mathcal{V}_B$ are isoclinic at angle $\theta = \frac{\pi}{4}$.

5.4 Higher Rank Numerical Ranges and Isoclinic Subspaces

We can also derive a connection with the higher rank numerical range of a matrix or operator. Originally considered in the setting of quantum error correction [14, 12], these numerical ranges have been intensely investigated for over a decade now in matrix theory and beyond [33, 49, 38, 11, 35, 34, 32, 15].

Given an operator or matrix A on $\mathbb{C}^n$ and $1 \le k \le n$, the *rank-k numerical range of A* is the subset of the complex plane given by:

$$\Lambda_k(A) = \left\{ \lambda \in \mathbb{C} \,\middle|\, PAP = \lambda P \, \text{for some rank}{-}k \, \text{projection} \, P \, \text{on} \, \mathbb{C}^n \right\}.$$

The following statement is Theorem 2.4 from [13]. It will be used in this section to provide a direct correlation to isoclinic subspaces from the higher rank numerical range. This theorem is of interest because if rank k of the projectors and rank l of operator A are known then $\Lambda_k(A)$ is also known for an Hermitian operator A.

Theorem 94. *Let A be a $n \times n$ Hermitian matrix with eigenvalues (counting multiplicities) given by $a_1 \le \cdots \le a_n$ and let $k \ge 1$ be a fixed integer with $1 \le k \le n$. Then the higher rank-k numerical range $\Lambda_k(A)$ coincides with $[a_k, a_{n-k+1}]$ which is:*

i) a non-degenerate closed interval if $a_k < a_{n-k+1}$,

ii) a singleton set if $a_k = a_{n-k+1}$,

iii) an empty set if $a_k > a_{n-k+1}$.

Moreover, $\Lambda_k(A)$ coincides with the intersection of numerical ranges $\Lambda_1(V^*AV)$, where V runs through all isometries $V : \mathbb{C}^{n-k+1} \to \mathbb{C}^n$.

Here we are interested in the case of higher rank numerical ranges of projections, which can be viewed as a special case of Hermitian operators considered in [13]. If Q is a non-zero projection with $\text{rank}(Q) = l < n$, then through an application of Theorem 94 from [13], it follows that $\Lambda_k(Q) = [0, 1]$ whenever $k \leq \min\{l, n - l\}$. Furthermore, if k is small then λ can range from any value in the interval $[0, 1]$, hence producing interesting examples of isoclinic subspaces. Contrarily, if k becomes considerably larger, λ is forced to the ends of the above interval causing λ to be either 0 or 1.

Proposition 95. *Let P and Q be nonzero projections on $\mathbb{C}^n$ of the same rank $1 \leq k \leq n$. Then $P\mathbb{C}^n$ and $Q\mathbb{C}^n$ are isoclinic subspaces at angle θ if and only if $PQP = \cos(\theta)P$ if and only if $QPQ = \cos(\theta)Q$.*

Proof. Firstly, the case that $\theta = \frac{\pi}{2}$ and $\cos(\theta) = 0$ corresponds to orthogonality of the two subspaces and $PQ = 0 = QP$. So let us assume $\cos(\theta) \neq 0$ for the rest of the proof.

Suppose $PQP = \cos(\theta)P$, and so

$$(QPQ)(QPQ) = QP(QQ)PQ = Q(PQP)Q = \cos(\theta)QPQ.$$

Next, dividing both sides by $\cos^2(\theta)$ we get,

$$\frac{1}{\cos^2(\theta)}(QPQ)(QPQ) = \frac{1}{\cos(\theta)}QPQ.$$

Hence $\cos^{-1}(\theta)QPQ$ is a projection that is evidently supported on $Q\mathbb{C}^n$. However, we also have, with $(\cdot)$ the trace functional,

$$\begin{aligned}(\frac{1}{\cos(\theta)}QPQ) = \frac{1}{\cos(\theta)}(QPQ) &= \frac{1}{\cos(\theta)}(QP) \\ &= \frac{1}{\cos(\theta)}(PQP) = (P) = (Q).\end{aligned}$$

As the rank of a projection is equal to its trace, it follows that in fact $QPQ = \cos(\theta)Q$.

Thus we have shown that $PQP = \cos(\theta)P$ if and only if $QPQ = \cos(\theta)Q$. The equivalence of these conditions with $P\mathbb{C}^n$ and $Q\mathbb{C}^n$ being isoclinic follows from Theorem 88.

□

Remark 96. In particular, for the projections P, Q corresponding to a pair of isoclinic subspaces, each of the projections is encoded into the structure of the other projection's higher rank numerical ranges in the sense that: P (respectively Q) is a projection corresponding to $\cos(\theta) \in \Lambda_k(Q)$ (respectively $\Lambda_k(P)$).

Returning to a $\text{rank}(Q) = l < n$ Hermitian projection Q in a n-dimensional space we can provide a complete characterization of what isoclinic subspaces will arise from the rank-k numerical range.

The next corollary will depend on the kth largest and kth smallest eigenvalues of the rank-r projection. Note that the eigenvalues of a projection are always either 0 or 1. The kth largest eigenvalue will be denoted $\lambda_\uparrow$ and the kth smallest eigenvalue will be denoted $\lambda_\downarrow$.

Corollary 97. *Let Q be a* rank *l Hermitian projection, and $\Lambda_k(Q)$ be the* rank *l-numerical range, then there are four cases that depend on the kth largest eigenvalue and kth smallest eigenvalue of the projection Q:*

i) If $\lambda_\downarrow = 0 = \lambda_\uparrow$ then $\Lambda_k(Q) = 0$.

ii) If $\lambda_\downarrow = 1 = \lambda_\uparrow$ then $\Lambda_k(Q) = 1$.

iii) If $\lambda_\downarrow = 1$ and $\lambda_\uparrow = 0$ then $\Lambda_k(Q) = \emptyset$.

iv) If $\lambda_\downarrow = 0$ $\lambda_\uparrow = 1$ then $\Lambda_k(Q) = [0, 1]$.

Remark 98. This corollary provides an interesting observation in relation to isoclinic subspaces. Each of the cases from the above corollary yield different subspaces.

1. For case i) and ii), the isoclinic subspaces that the rank l Hermitian projector produces are trivial that is, the subspaces are only isoclinic to themselves.

2. In case iii) where $\Lambda_k(Q) = \emptyset$, no isoclinic subspaces arise.

3. Lastly, case iv) will yield non-trivial subspaces that is, many different subspaces that are isoclinic to each other will arise depending on the different values of $\lambda \in \Lambda_k(Q) = [0, 1]$.

5.5 Correlation Matrices and QEC Conditions

In this section a different approach is taken to the existence of isoclinic subspaces. The question of interest in this section is: Do there exist isoclinic subspaces of $\mathbb{C}^m$, $\mathcal{V}_1$, $\mathcal{V}_2$,..., $\mathcal{V}_n$ such that the dimension of all subspaces are k and the canonical angles between $\mathcal{V}_i$ and $\mathcal{V}_j$ are given angles θ_{ij}? What are the necessary and sufficient conditions on the θ_{ij}s so that these subspaces exist? Some intuition on this problem is given in terms of correlation matrices.

Firstly, we must define what a correlation matrix is.

Definition 99. Let B be a $n \times n$ matrix and let b_{ij} be an entry in B. B is called a correlation matrix if B is a positive semi-definite matrix and all entries $b_{ii} = 1$ and all entries $b_{ij} = b_{ji}$.

Recall the Quantum error correction conditions, equation (3.7),

$$P_{\mathcal{C}} E_i^* E_j P_{\mathcal{C}} = \alpha_{ij} P_{\mathcal{C}},$$

for some Hermitian matrix α. From the quantum error correction conditions a correlation matrix can be obtained by letting $A_i = \frac{1}{\sqrt{\alpha_i i}} E_i P_{\mathcal{C}}$ and $\beta_{ij} = \frac{\alpha_{ij}}{\sqrt{\alpha_{ii}}\sqrt{\alpha_{jj}}}$. From this substitution we get:

$$\begin{aligned} A_i^* A_j &= \frac{1}{\sqrt{\alpha_{ii}}\sqrt{\alpha_{jj}}} P_{\mathcal{C}} E_i^* E_j P_{\mathcal{C}} \\ &= \frac{\alpha_{ij}}{\sqrt{\alpha_{ii}}\sqrt{\alpha_{jj}}} P_{\mathcal{C}} \\ &= \beta_{ij} P_{\mathcal{C}}, \end{aligned}$$

where β is a real correlation matrix for all entries β_{ij}.

Relating this back to isoclinic subspaces, a proposition will now be stated to attempt to answer the question posed at the beginning of this section.

Proposition 100. *Let* $\mathcal{V}_1, \mathcal{V}_2, ..., \mathcal{V}_n$ *be* k *dimensional subspaces of* $\mathbb{C}^m$*. If* $M = [\cos(\theta_{ij})]_{i,j=1}^n$ *is a correlation matrix satisfying the inequality* $\operatorname{rank}(M) \leq \lfloor \frac{m}{k} \rfloor$ *where* θ_{ij} *is the canonical angle between subspace* $\mathcal{V}_i$ *and* $\mathcal{V}_j$ *then* $\mathcal{V}_i$ *and* $\mathcal{V}_j$ *are pairwise isoclinic at angle* θ_{ij} *for all* i, j*.*

Proof. Let $d = \lfloor \frac{m}{k} \rfloor$. Then there exists $\{v_i\}_{i=1}^n \subseteq \mathbb{C}^d$ all unit vectors such that $\langle v_i, v_j \rangle = \cos(\theta_{ij})$ for all i, j. Now consider $\mathbb{C}^d \otimes \mathbb{C}^k$ to be a subspace of $\mathbb{C}^m$. Let $\mathcal{V}_i = (\operatorname{span}\{v_i\}) \otimes \mathbb{C}^k$ for all $1 \leq i \leq n$. Then $\mathcal{V}_i$ and $\mathcal{V}_j$ are pairwise isoclinic at angle θ_{ij} for all i, j. □

The converse of this proposition will be left as an open problem. The converse can in fact be shown for $k = 1$, this is a straightforward case as the rank of the vectors making up the matrix M will always be less then or equal to 1. However, the general case of the converse has not been able to be shown true yet. There is potential in this correlation matrix route to study isoclinic subspaces but this proposition must be proofed fully. This will be discussed more in the conclusions and further work section.

Definition 101. Let $\{v_i\}_{i=1}^m$ be a set of vectors, where $v_i \in \mathbb{C}^n$. The Gram matrix G is of all inner products of the set $\{v_i\}_{i=1}^m$, that is, the entries of G are

$$g_{ij} = g_i^* g_j.$$

Remark 102. If you have a set of unit vectors then the resultant Gram matrix M is a correlation matrix of $\operatorname{rank}(M) \leq \lfloor \frac{m}{k} \rfloor$

Chapter 6

Conclusions and Further Work

This thesis has been a mathematical exploration of the relationship between quantum error correction and isoclinic subspaces. It outlines a new connection in quantum error correction and serves as the preliminary framework to introduce canonical angles to the mathematics of quantum information. Quantum mechanics and information theory are intricate fields of study. The elaborate mathematics that arise in response to the complexity of quantum studies is a way of breaking the field down into subjects already well understood in the hopes of further understanding quantum. Providing an easier approach to quantum information, is the primary purpose of this work. While the connection made here is only a start, the hope is that other researchers can take the approach that was delineated in this work and use it as a way to further understand and simplify quantum information theory.

6.0.1 Open Problems

There were two main problems that we were unable to be solved fully, the first being Theorem 90. The reason that Theorem 90 is only an "if" statement is because a full converse was unable to be shown. Instead, a partial converse was only able to be completed, namely Proposition 92. Theorem 90 is the more general case which states that for a correctable error model the subspaces that arise from the corresponding error operators will be isoclinic. Proposition 92 was shown only for a pair of projections on Hilbert space $\mathcal{H}$. The reason that an arbitrary correctable error model is possible for a pair of projectors associated with isoclinic subspaces is because we are able to choose the subspaces accordingly such that they are arranged isoclinic to each other already. One subspace can be adjusted so that Eq. (5.1) of Theorem 88 is satisfied as the proof for Proposition 92 shows. To go further for three or more subspaces this becomes more difficult. We cannot necessarily adjust all the subspaces so that they are all pairwise isoclinic to each other, we can only be sure that at most two subspaces will be isoclinic to each other. This is left as an open problem for now. Perhaps more examples may fuel some intuition towards solving this problem. This problem is of interest because in theory it is a way to find new correctable quantum error models from isoclinic subspaces.

The next problem was originally forefront of this work but changed directions once Proposition 100 could not be proved fully. Sufficiency was shown to be true, but for necessity, only special cases emerged in which it will hold. If we have subspaces $\{\mathcal{V}_i\}_{i=1}^n$ that are pairwise isoclinic then we will get a correlation matrix M such that the entries will equal $\cos(\theta_{ij})$ where θ_{ij} is the angle of isoclinicity between each subspace. The problem lies in the rank(M). It is not immediately clear that this rank inequality, namely, rank$(M) \leq \lfloor \frac{m}{k} \rfloor$ will hold in general. The most promising method of proof was after assuming the subspaces $\{\mathcal{V}_i\}_{i=1}^n$ of dimension k to be isoclinic, we can select k vectors in each subspace such that the inner product of these any of these vectors with any other of k vectors from another isoclinic subspace will be equal to $\cos(\theta_{ij})$. This method is clear for 2 isoclinic subspaces and can be shown using the joint higher rank numerical range definition with condition $iii)$ of Theorem 88. However, for 3 and greater isoclinic subspaces there is not an apparent way of getting a single set of vectors for matrix M that will give rise to this rank condition.

These open problems seem to be similar to each other. Perhaps solving one could lead to solving the other if they are indeed feasible problems to be solved.

6.0.2 Canonical Angles and Other Connections in Quantum Information

This work introduces canonical angles to quantum error correction. It is a natural entry point because of the use of subspaces in the mathematical theory of quantum error correction. Canonical angles are perhaps another avenue in which quantum error correction can be studied from in quantum information. It is important to find alternative ways to study error correction because of its importance in the field. Error correction and its short comings can be traced back to virtually every area of quantum information, so it may be the case that working with canonical angles and isoclinic subspaces can also be extended to other places within quantum information. Further, canonical angles are an easy theory to work with. Due to Bjorck and Golub (Theorem 71), canonical angles can be calculated using singular values. Hence, if canonical angles are found to be useful in another part of quantum information these methods would be quite accessible.

6.0.3 Infinite Dimensional case

This thesis focuses on finite dimensional spaces where further works could consider the infinite dimensional Hilbert space case. There would be two considerations, the first would be for an infinite dimensional Hilbert space and finite dimensional isoclinic subspaces. In this case, one of the main points of the work, Theorem 88 would still hold to be true. The reason for this is the dimensions of the columns of the projections onto these subspaces. Since these subspaces are finite dimensional, we would assume the columns of these projections to also be finite dimensional, in turn this would allow Theorem 88 to carry through.

However, in the second case things become more confusing. This case would deal with an infinite dimensional Hilbert space and subspaces that are infinite dimensional as well. This is a much more interesting case and where future work could be centered around. We must be careful about these subspaces though because in infinite dimensions an extra condition is added, that the subspaces must be closed (subspaces in finite dimensions are always closed). Now, to build some intuition on this problem, the most obvious way to define these infinite dimensional subspaces is in terms of projections. There will always be a unique orthogonal projection onto these subspaces. Using these projections we could define the subspaces to be isoclinic at some angle and proceed to explore their properties using the projection relation in *iii*) of Theorem 88. A clear next step to defining these infinite subspaces more precisely would be defining them more generally using C^*-algebras and von Neumann algebras (for background on these subjects [54, 8]). Using these interpretations we can see what could be said about these infinite dimensional subspaces and isoclinicity.

Further, there are some works, [6, 29] to name some, that already discuss infinite dimensional quantum error correction in terms of operator algebras. These may be quite helpful in future works when trying to merge infinite dimensional quantum error correction and isoclinic subspaces. Specifically, by attempting to explore the isoclinic projection relation from *iii*) of Theorem 88 for infinite dimensions.

Bibliography

[1] Adams, J. F. Vector fields on spheres. *Annals of Mathematics* (1962), 603–632.

[2] Altmann, S. L. *Rotations, quaternions, and double groups.* Courier Corporation, 2005.

[3] Balla, I., and Sudakov, B. Equiangular subspaces in Euclidean spaces. *Discrete & Computational Geometry 61*, 1 (2019), 81–90.

[4] Bennett, C. H., DiVincenzo, D. P., Smolin, J. A., and Wootters, W. K. Mixed-state entanglement and quantum error correction. *Physical Review A 54*, 5 (1996), 3824.

[5] Bennett, C. H., DiVincenzo, D. P., Smolin, J. A., and Wootters, W. K. Mixed state entanglement and quantum error correction. *Physical Review A 54* (1996), 3824.

[6] Bény, C., Kempf, A., and Kribs, D. W. Quantum error correction on infinite-dimensional Hilbert spaces. *Journal of Mathematical Physics 50*, 6 (2009), 062108.

[7] Björck, Å., and Golub, G. H. Numerical methods for computing angles between linear subspaces. *Mathematics of Computation 27*, 123 (1973), 579–594.

[8] Blackadar, B. *Operator algebras: theory of C*-algebras and von Neumann algebras*, vol. 122. Springer Science & Business Media, 2006.

[9] Blahut, R. E. *Theory and practice of error control codes.* Addison-Wesley, 1983.

[10] Choi, M.-D. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications 10*, 3 (1975), 285–290.

[11] Choi, M.-D., Giesinger, M., Holbrook, J. A., and Kribs, D. W. Geometry of higher-rank numerical ranges. *Linear and Multilinear Algebra 56*, 1-2 (2008), 53–64.

[12] Choi, M.-D., Kribs, D. W., and Życzkowski, K. Higher-rank numerical ranges and compression problems. *Linear Algebra and its Applications 418*, 2-3 (2006), 828–839.

[13] Choi, M.-D., Kribs, D. W., and Życzkowski, K. Higher-rank numerical ranges and compression problems. *Linear algebra and its applications 418*, 2-3 (2006), 828–839.

[14] Choi, M.-D., Kribs, D. W., and Zyczkowski, K. Quantum error correcting codes from the compression formalism. *Reports on Mathematical Physics 58*, 1 (2006), 77–91.

[15] Gau, H.-L., Li, C.-K., Poon, Y.-T., and Sze, N.-S. Higher rank numerical ranges of normal matrices. *SIAM Journal of Matrix Analysis and its Applications 32* (2011), 23–43.

[16] Gentle, J., and Gentle, J. Cholesky factorization. *Numerical Linear Algebra for Applications in Statistics* (1998), 93–95.

[17] Gottesman, D. Class of quantum error-correcting codes saturating the quantum Hamming bound. *Physical Review A 54* (1996), 1862.

[18] Gottesman, D. An introduction to quantum error correction and fault-tolerant quantum computation. In *Quantum information science and its contributions to mathematics, Proceedings of Symposia in Applied Mathematics* (2010), vol. 68, pp. 13–58.

[19] Griffel, D. H. *Applied functional analysis.* Courier Corporation, 2002.

[20] Hamming, R. W. Error detecting and error correcting codes. *The Bell System Technical Journal 29*, 2 (1950), 147–160.

[21] Hardy, A. S. *Elements of quaternions.* Ginn, Heath, & Company, 1881.

[22] Hitzer, E. Angles between subspaces. *arXiv preprint arXiv:1306.1629* (2013).

[23] Hoggar, S. New sets of equi-isoclinic n-planes from old. *Proceedings of the Edinburgh Mathematical Society 20*, 4 (1977), 287–291.

[24] Jordan, C. Essai sur la géométrie à n dimensions. *Bulletin de la Société mathématique de France 3* (1875), 103–174.

[25] Kaye, P., Laflamme, R., Mosca, M., et al. *An introduction to quantum computing.* Oxford University Press, 2007.

[26] Knill, E., and Laflamme, R. A theory of quantum error-correcting codes. *Physical Review A* (1996).

[27] Knill, E., Laflamme, R., and Viola, L. Theory of quantum error correction for general noise. *Physical Review Letters 84*, 11 (2000), 2525.

[28] Kreyszig, E. *Introductory functional analysis with applications*, vol. 1. wiley New York, 1978.

[29] Kribs, D., Laflamme, R., and Poulin, D. Unified and generalized approach to quantum error correction. *Physical Review Letters 94*, 18 (2005), 180501.

[30] Kribs, D. W. A quantum computing primer for operator theorists. *Linear Algebra and its Applications 400* (2005), 147–167.

[31] Kribs, D. W., Mammarella, D., and Pereira, R. Isoclinic subspaces and quantum error correction. *arXiv preprint arXiv:1912.10100* (2019).

[32] Li, C.-K., and Poon, Y.-T. Generalized numerical ranges and quantum error correction. *Journal of Operator Theory* (2011), 335–351.

[33] Li, C.-K., Poon, Y.-T., and Sze, N.-S. Higher rank numerical ranges and low rank perturbations of quantum channels. *Journal of Mathematical Analysis and Applications 348*, 2 (2008), 843–855.

[34] Li, C.-K., Poon, Y.-T., and Sze, N.-S. Condition for the higher rank numerical range to be non-empty. *Linear and Multilinear Algebra 57*, 4 (2009), 365–368.

[35] Li, C.-K., and Sze, N.-S. Canonical forms, higher rank numerical ranges, totally isotropic subspaces, and matrix equations. *Proceedings of the American Mathematical Society 136*, 9 (2008), 3013–3023.

[36] MacWilliams, F. J., and Sloane, N. J. A. *The theory of error-correcting codes*, vol. 16. Elsevier, 1977.

[37] Martin, W. J. A physics-free introduction to quantum error correcting codes. *Utilitas Mathematica 65* (2004), 133–158.

[38] Martinez-Avendano, R. A. Higher-rank numerical range in infinite-dimensional Hilbert space. *Operators and Matrices 2*, 2 (2008), 249–264.

[39] Mermin, N. D. *Quantum computer science: an introduction.* Cambridge University Press, 2007.

[40] Nielsen, M. A., and Chuang, I. Quantum computation and quantum information, 2002.

[41] Preskill, J. Reliable quantum computers. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 454*, 1969 (1998), 385–410.

[42] Radon, J. Lineare scharen orthogonaler matrizen. In *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* (1922), vol. 1, Springer, pp. 1–14.

[43] Rynne, B., and Youngson, M. A. *Linear functional analysis.* Springer Science & Business Media, 2013.

[44] Shannon, C. A mathematical theory of communication. *Bell System Technical Journal, vol. 27, pp. 379-423 & 623-656, July & October* (1948).

[45] Shor, P. W. Scheme for reducing decoherence in quantum computer memory. *Physical Review A 52*, 4 (1995), R2493.

[46] Shor, P. W. Scheme for reducing decoherence in quantum computing memory. *Physical Review A 52* (1995), R2493.

[47] Steane, A. M. Error correcting codes in quantum theory. *Physical Review Letters 77*, 5 (1996), 793.

[48] Tverberg, H. A generalization of Radon's theorem. *Journal of the London Mathematical Society 1*, 1 (1966), 123–128.

[49] Woerdeman, H. J. The higher rank numerical range is convex. *Linear and Multilinear Algebra 56*, 1-2 (2008), 65–67.

[50] Wong, Y.-C. Clifford parallels in elliptic (2n - 1)-spaces and isoclinic n-planes in euclidean 2n-space. *Bulletin of the American Mathematical Society 66*, 4 (1960), 289–293.

[51] Wong, Y.-C. *Linear geometry in Euclidean 4-space.* SEAMS, 1977.

[52] Wong, Y.-C. *Linear geometry in Euclidean 4-space.* No. 1. Southeast Asian Mathematical Society, 1977.

[53] Wong, Y.-C., and Mok, K.-P. Normally related n-planes and isoclinic n-planes in $\mathbb{R}^{2n}$ and strongly linearly independent matrices of order n. *Linear Algebra and its Applications 139* (1990), 31–52.

[54] Zhu, K. *An introduction to operator algebras.* CRC press, 2018.