The Structure of the Real Numerical Range and the Surface Area Quantum Entanglement Measure

by

Matthew Kazakov

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Master of Science
in
Mathematics and Statistics

Guelph, Ontario, Canada

© Matthew Kazakov, December, 2018

ABSTRACT

THE STRUCTURE OF THE REAL NUMERICAL RANGE AND THE SURFACE AREA QUANTUM ENTANGLEMENT MEASURE

Matthew Kazakov
University of Guelph, 2018

Advisors:
Dr. Rajesh Pereira
Dr. David Kribs

An extensive analysis has been done on the numerical range of an operator, however, little research has been done on its real analogue. In this thesis we give a number of results and properties regarding the real numerical range, and real higher rank numerical range. We motivate this study by providing the reader with an application of how the real higher rank numerical range may be used in the study of conic sections. Finally, we end the paper with a short introduction into the field of quantum information theory, eventually building up to introduce a new measure of entanglement for pure symmetric states.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank both of my parents, Pat and Dragi Kazakov. They have been on this journey with me from day one through till the very end. All of those long nights in the office solving problems, deriving new expressions and ideas are dedicated to you both. This would all have been a dream were it not for you both to make it a reality. I am forever grateful for all the love and support you both have given me. Thanks mom and dad!

Next, I would like to take this opportunity to thank each of my advisors individually. To Dr. David Kribs, thank you for all your time and advice on all things operator and quantum information theory related. These fields can, at times, seem to be very complicated and confusing in nature. However, with your knowledge and experience, you were able to simplify many concepts for me so that I was able to better my overall understanding and thinking abilities, not just in these fields, but in mathematics as a whole. Thank you David, it has been a pleasure!

To Dr. Rajesh Pereira, I cannot say enough good things! So many times I have been blown away at the immense knowledge you have in this ever-expanding universe that is mathematics. You are one of the most intelligent people I have had the pleasure of knowing and working with. I thank you for all of our insightful conversations. I have walked away with so much more knowledge about so many areas of mathematics and a better abstract problem solver. Thank you Rajesh.

Lastly I would like to thank the University of Guelph- Mathematics and Statistics

department for funding my Masters research.

Contents

Abstract **ii**

Acknowledgements **iii**

List of Figures **vii**

1 Introduction **1**

2 Background **3**
2.0.1 Some basic definitions and notation 3
2.0.2 The numerical range $W(A)$ 11
2.0.3 The real numerical range $R(A)$ 13
2.0.4 The complex higher rank numerical range $\Lambda_k(A)$ 18
2.0.5 The real higher rank numerical range $R_k(A)$ 20
2.0.6 The joint numerical range (real and complex) 20
2.0.7 Hyperboloids and ellipsoids 21

3 Conic sections and the real k- rank numerical range **29**
3.0.1 Results on the real numerical range 29
3.0.2 Hyperspherical cross sections of ellipsoids and hyperboloids . . 44

4 Connections to quantum information and the surface area entanglement measure **47**
4.1 Quantum information preliminaries 47
4.1.1 An introduction to basic quantum information theory 49
4.1.2 Quantum error correction and the (complex) k- rank numerical range 53
4.1.3 A synopsis of entanglement 57
4.2 Introducing the new measure 63
4.2.1 Regular polygons and special case polyhedra 67
4.2.2 The surface area entanglement measure 74

4.2.3 A possible extension of the measure 80
4.2.4 Known results for the n point problem 81
4.2.5 A link to the Thomson problem in physics 86

5 Conclusions and future work **88**
5.0.1 The real higher rank numerical range 88
5.0.2 The surface area entanglement measure 89

List of Figures

2.1 Image of a one-sheeted hyperbola in three dimensional space. Taken from [35]. 24
2.2 Image of a two sheeted hyperbola in three dimensional space. Taken from [36]. 24

4.1 Depiction of a maze. Item taken from Google images. 48
4.2 Image of the Bloch sphere taken from [1]. 50
4.3 A regular polygon decomposed into n triangles. 68
4.4 The inscribed QSP for $|\psi\rangle$ 77
4.5 The inscribed QSP for the GHZ state in 3 particles as depicted by the red triangle. 79
4.6 Figure taken form [29] depicting the maximal volume polyhedra for the cases of 28, 29 and 30 vertices. 85

Chapter 1

Introduction

In this thesis we examine both the higher rank numerical range (what will be referred to as the *complex* higher rank numerical range throughout this thesis) as well as its real analogue. The terms higher rank numerical range and rank-k numerical range are used interchangeably.

In chapter 2 we provide an extensive amount of prerequisite information to outline the results and conclusions given in the latter sections. This chapter includes the classical numerical range which we denote $W(A)$, the real numerical range, denoted $R(A)$, the complex higher rank numerical range, denoted $\Lambda_k(A)$, the real higher rank numerical range, denoted $R_k(A)$, the joint complex higher rank numerical range and the joint real higher rank numerical range. A natural ordering presents itself here based on the timeline in which these objects were constructed.

In chapter 3 we give the main results concerning the real higher rank numerical range. One of the main theorems we derive is the real elliptical range theorem. This result and others are presented here.

In chapter 4, our focus is primarily on quantum information theory. The beginning remarks here are designed to introduce the topic and notation used further on in the section. A brief connection is defined, stating how the complex higher rank numerical range appears in quantum error correction. We then give a detailed introduce to entanglement theory which serves as a stepping stone for the entanglement measure we create. Finally, we introduce the surface area entanglement measure and a possible extension.

In chapter 5 we summarize the important results and key concepts of the paper. A short discussion regarding future work is also discussed here.

Chapter 2

Background

We introduce the notation to be used throughout the course of the paper, along with a detailed overview of the different numerical ranges that will be seen. We end off with a discussion of conic sections as they do manifest themselves naturally in this context.

2.0.1 Some basic definitions and notation

We provide here a complete set of definitions and relevant facts needed throughout the remainder of the paper. Much of what follows here is a discussion about matrices and some definitions from analysis.

Definition 2.1. *An* ***eigenvalue*** *of a matrix A is a scalar λ that satisfies the equality $Ax = \lambda x$ for some non-zero vector x. Here x is said to be an* ***eigenvector****.*

Definition 2.2. *The* ***singular value*** *of a matrix A, denoted s_i (for the ith singular*

value) are the square roots of the eigenvalues of the matrix A^*A. They are usually written in descending order, i.e.

$$s_1 \geq s_2 \geq \cdots \geq s_n$$

Throughout this paper, we will typically use λ to denote eigenvalues of operators. We also will use the letters $a, b, c, d, \alpha, \beta$ to denote complex numbers, though this should be clear based on the context they come up in. It should also be stated that all operators mentioned in this paper are finite dimensional, hence we will use this interchangeably with the term, matrix. In the definitions that follow, we will take A^* to mean the conjugate transpose of the matrix A.

We now introduce some notions and remarks regarding types of matrices, properties and two well known decompositions.

Definition 2.3. *A matrix H is said to be **Hermitian** (or **self-adjoint**) if $H = H^*$.*

Lemma 2.1. *Hermitian matrices have an entirely real spectrum.*

Proof. Suppose $Hx = \lambda x$ for some norm 1 vector x. Then we see that

$$\lambda = x^*Hx = (x^*H^*x)^*$$

but by the hermiticity of H we get that

$$(x^*H^*x)^* = (x^*Hx)^* = \bar{\lambda},$$

and hence we have shown $\lambda = \bar{\lambda}$ which implies λ must have imaginary part equal to zero, thus making it entirely real. □

Definition 2.4. *A matrix U is said to be* ***unitary*** *if it satisfies $U^*U = UU^* = I$.*

Definition 2.5. *A matrix P is said to be* ***positive definite*** *if it is symmetric and all of its eigenvalues are strictly greater than zero. Equivalently, P is positive definite if for all non-zero $x \in \mathbb{R}^n$ we have $x^*Px > 0$. A similar definition holds for positive semidefinite operators, allowing there to be zero eigenvalues (and with $x^*Px \geq 0$).*

Lemma 2.2. *Let P be a positive semidefinite matrix. Then there exists a matrix B such that $P = B^*B$.*

Proof. If P is positive definite, there exists an eigen-decomposition (see statement immediately after proof) where

$$P = U^*\Lambda U$$

for some unitary matrix U and diagonal matrix Λ. Because Λ is diagonal and itself positive definite we may assert $\Lambda = \Lambda^{1/2}\Lambda^{1/2}$. And thus,

$$P = U^*\Lambda U = U^*\Lambda^{1/2}\Lambda^{1/2}U = (\Lambda^{1/2}U)^*(\Lambda^{1/2}U)$$

here we can choose our matrix B to simply be $\Lambda^{1/2}U$ and the result follows. □

Eigen-decomposition: Given any normal matrix A, there exists square matrices V and diagonal matrix D such that

$$A = VDV^{-1}$$

In fact, the columns of V will be the eigenvectors of A and similarly, the diagonal entries of D will be the eigenvalues of A.

Remark: This is equivalently the spectral theorem for normal matrices.

Remark: For any continuous complex function f, we have $f(A) = Vf(D)V^{-1}$.

These two definitions will be useful in the quantum information setting, talked about later on.

Definition 2.6. *Let V be a real vector space, then the mapping*

$$\langle\ ,\rangle : V \times V \to R$$

*is an **inner product** if it satisfies;*

1. $\langle x, cy\rangle = c\langle x, y\rangle$
2. $\langle x + z, y\rangle = \langle x, y\rangle + \langle z, y\rangle$
3. $\langle x, y\rangle = \langle y, x\rangle$
4. $\langle x, x\rangle \geq 0$

for all $x, y, z \in V$ and $c \in \mathbb{R}$.

The above definition can be modified to the complex case. Here the only points that would change would be (1), where we will take the inner product to be conjugate linear in the second argument, i.e. $\langle x, cy\rangle = \overline{c}\langle x, y\rangle$ and (3) would also be modified to read $\langle x, y\rangle = \overline{\langle y, x\rangle}$. It should also be noted that in point (4), the equality is held if and only if x is the zero vector.

Definition 2.7. *A **Hilbert space**, denoted by $\mathcal{H}$ is a (real or complex), inner product space that is complete with respect to the metric induced by the inner product. I.e. for some $a \in \mathcal{H}$ we have*

$$||a|| = \sqrt{\langle a, a \rangle}.$$

Definition 2.8. *We write $B(\mathcal{H})$ to denote the set of bounded linear operators on the Hilbert space $\mathcal{H}$.*

Polar decomposition: Given a complex matrix A of size n, there exists a unitary matrix U and positive semidefinite matrix P such that $A = UP$.

Remark: The statement above is also valid for rectangular matrices, i.e. supposing if A was $m \times n$ in size with $m \geq n$, the same decomposition holds true, the only difference being that U will be of size $m \times n$ as well. This statement given is (formally) known as the *right* polar decomposition of A. There is also a left polar decomposition phrased in a similar way. The right and left decompositions will be the same provided A is square.

Definition 2.9. *A set $\mathcal{S}$ is **convex** if for any points u and $v \in \mathcal{S}$ the point $tu+(1-t)v$ is also contained in $\mathcal{S}$ for any $t \in [0, 1]$.*

Definition 2.10. *The **convex hull** of a set $\mathcal{S}$, denoted $conv(\mathcal{S})$ is the smallest convex set containing $\mathcal{S}$.*

Related to this definition we have the following definition,

Definition 2.11. *A set $\mathcal{C}$ is said to be* ***connected*** *if it cannot be decomposed into two non-empty, disjoint, open subsets. Similarly, a* ***simply connected*** *set is one such that any closed loop formed within the set can be shrunk to a single point (without passing outside the set).*

Example 2.1. *Any disjoint union of two closed intervals is not connected. By contrast, the entire real line is said to be connected, as well as any open or closed interval of $\mathbb{R}$.*

Having now stated some terms used in analysis, we briefly go back to the subject of matrices; talking about the Pauli and Gell-Mann matrices respectively. These are not only useful in the pure-mathematical context we are diving into, but also the quantum information setting discussed later on.

Definition 2.12. *We denote the $n+1$ dimensional unit hypersphere $\mathcal{S}^n$ to be the set*

$$\{x \in \mathbb{R}^{n+1} \mid ||x|| = 1\}$$

Definition 2.13. *The* ***Pauli matrices*** *are four complex 2×2 matrices that form a basis for all 2×2 complex matrices. They are given as*

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The Paulis have many application in quantum mechanics and quantum information respectively. Particularly in the quantum information setting, they represent different rotations of the Bloch sphere. Furthermore, in the quantum computation

setting, they can be strung together in quantum circuits for designing certain algorithms. A very nice illustration of the Pauli matrices and Bloch sphere are given in [30], which has come to be the standard reference for students and researchers in quantum information.

Remark: The Pauli matrices have several important properties. A few of them include that they are unitary, i.e. following the property $U^*U = UU^* = I$ and that they are each trace zero.

In 2 dimensions the Pauli matrices are quite nice to work with, however, in higher dimensions we no longer have as nice of a decomposition for arbitrary matrices. Although, one such generalization of the Pauli matrices, the Gell-Mann matrices, does give us something to work with.

Definition 2.14. *The* ***Gell-Mann matrices*** *(defined here in 3 dimensions) are a collection of 9 matrices that span the space of complex* 3×3 *matrices. They are given as*

$$g_1^a = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix} \quad g_2^a = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad g_3^a = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}$$

$$g_4^s = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad g_5^s = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad g_6^s = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$g_7^d = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad g_8^d = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \qquad g_9^d = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The notation we use to denote the Gell-Mann matrices comes from [4]. These matrices have their primary application in particle physics, where they span the Lie algebra of the special unitary group of 3×3 matrices, which is used to study the strong interactions between particles. Should the reader be curious about these objects, [15] is a good, informative source on the matter.

Remark: These matrices, unlike the Paulis, are not unitary. However, they are still Hermitian. They too are also trace zero and either have all entries purely real or purely imaginary (as in the case of the Pauli matrices).

Remark: A generalization of the Gell-Mann matrices called the **generalized Gell-Mann matrices** exist for $n \geq 4$. A basis for any real symmetric matrix of size n may then be given by

$$g_{j,k}^{sym} = E_{j,k} + E_{k,j}$$

and

$$g_l^{diag} = \sqrt{\frac{2}{l(l+1)}} \Big(\sum_{j=1}^{l} E_{j,j} - l E_{l+1,l+1} \Big)$$

(ignoring all the antisymmetric Gell-Mann matrices). Here the notation E_i, j defines the matrix with only a 1 in the (i, j) entry, and zeros everywhere else.

Remark: In dimension n, there are $n^2 - 1$ Gell-Mann matrices. Of these, $\frac{n(n-1)}{2}$ are symmetric and similarly the same number of antisymmetric matrices and $n - 1$ diagonal Gell-Mann matrices of size n.

2.0.2 The numerical range $W(A)$

The ordinary numerical range, or field of values of an operator (one of the earlier names used for it) has been a subject of much study over the past century. Since then many generalizations have been made, and applied to other research fields.

Definition 2.15. *Given a finite dimensional vector space $V = \mathbb{C}^n$ and some operator $A \in \mathcal{M}_n(\mathbb{C})$, we define the* ***numerical range*** *of A as the set*

$$W(A) = \{\lambda \in \mathbb{C} | \lambda = x^* A x \text{ with } x^* x = 1\}.$$

This set has a number of nice features, of which we present five here. Four of them are listed as a corollary below and the final is reserved as a theorem, originally published 100 years ago, independently by Toeplitz and Hausdorff.

Corollary 2.3. *Given any operator $A \in \mathcal{M}_n(\mathbb{C})$ for $n < \infty$ with numerical range $W(A)$, the following are always true.*

1. *For any choice of $x, y \in \mathbb{C}$,*

$$W(xA + yI_n) = xW(A) + y$$

2. *For any choice of* $A_1, A_2 \in \mathcal{M}_n(\mathbb{C})$,

$$W(A_1 + A_2) \subseteq W(A_1) + W(A_2)$$

3. *If* A *is Hermitian, then* $W(A) \subset \mathbb{R}$ *(in fact this is an if and only if statement). Also,* $\partial W(A)$ *will consist of only the minimum and maximum eigenvalues of* A.

4. *For any choice of* A, $W(A) = W(A^T)$ *and* $W(A^*) = \overline{W(A)}$ *(A consequence of this property is that if* A *is taken to be real, then* $W(A)$ *will be symmetric to with respect to the real axis).*

Theorem 2.4. *(Toeplitz-Hausdorff)* $W(A)$ *is convex.*

The proof of this theorem can be found in either of [11] or [17]. Each of these references have a nice construction of this theorem.

Although much research has been done on the numerical range, open problems still remain concerning certain properties of it. Two of them are listed here. Each of these can be found in [33].

Open problem 1: Define what is called the **inner numerical radius** by

$$r_{inn}(A) = \min\{|\lambda| : \lambda \in \partial W(A)\}$$

and let $s_{min}(A)$ denote the smallest singular value of A. Then is it always true that $r_{inn}(A) \leq s_{min}(A)$?

Open Problem 2: What are the necessary and sufficient conditions to have the origin be an element of $W(A)$?

Finally, we present one final theorem concerning the numerical range of a two dimensional operator. This theorem will be important as we present a real version of it in the following chapter. For proof of this theorem we direct the reader to [23] for a very nice construction.

Theorem 2.5. *(Elliptical range) Let T be a 2 dimensional, complex operator with eigenvalues λ_1 and λ_2. The numerical range, $W(T)$ is then an ellipse centred at $\frac{1}{2}Tr(T^*T)$ and with minor axis length $\sqrt{Tr(T^*T) - |\lambda_1|^2 - |\lambda_2|^2}$ and foci equal to the two respective eigenvalues.*

Example 2.2. *Consider the two dimensional operator given by, $T = \begin{pmatrix} i & 1 \\ 1 & 0 \end{pmatrix}$. Here, $T^*T = \begin{pmatrix} 2 & -i \\ i & 1 \end{pmatrix}$ and so $W(T)$ would trace out an ellipse centred at $\frac{3}{2}$ with minor axis of length 1 and foci of $\lambda_1 = \frac{1}{\sqrt{2}}(\sqrt{3} + i)$ and $\lambda_2 = \frac{1}{\sqrt{2}}(-\sqrt{3} + i)$.*

2.0.3 The real numerical range $R(A)$

The real numerical range, similar in spirit to that of the classical numerical range, $W(A)$ is discussed in detail here. The goal of this section is to provide an insight into the overlapping similarities to $W(A)$, as well as some differences. While doing so, we effectively explore the history of the real numerical range, making summary of the two primary papers on the subject. In what follows, a synopsis of [6] and [27] is given.

Definition 2.16. *The **real numerical range** of an operator $A \in \mathcal{M}_n(\mathbb{R})$, denoted by $R(A)$ is given by*

$$\left\{x^T Ax \mid x \in \mathbb{R}^n \text{ with } x^T x = 1\right\}$$

Note that this definition only differs in the slightest from that of the ordinary numerical range. The only change made was in the vector x. Here, they must be real.

Remark: One obvious assertion one can immediately make is that $R(A) \subset W(A)$.

We now proceed in giving a summary of the results of [27].

Initially stated without a proof, the first major result stated in this paper pertains to the study of partial differential equations. We opt to phrase this result as a theorem.

Theorem 2.6. *Given the partial differential operator*

$$A = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij}\frac{\partial^2}{\partial x_i \partial_j} + \sum_{i}^{n} \beta_i \frac{\partial}{\partial x_i} + C$$

where $\alpha_{i,j}, \beta_i$ and C are constants over $\mathbb{C}$. Then this operator is elliptic if zero is not contained in $R(A)$ and strongly elliptic if the zero element is not contained in $Re(R(A))$.

This one theorem highlights a nice connection between different areas of mathematics. Here, the term elliptic is used in the context of partial differential equations. Let S be some subset of $\mathbb{R}^n$. Then a differential operator A is said to be elliptic if $A(s, x)$ is non-zero for every $s \in S$ and non-zero $x \in \mathbb{R}^n$. For more information on strongly elliptic operators, we direct the reader to [7].

Theorem 2.7. *Let A be a complex 2×2 matrix. Then for the case of $n = 2$, $R(A)$ forms an ellipse (possibly degenerate- meaning one of its major axes have been flattened) centred at $\frac{1}{2}Tr(A)$.*

This theorem is a key component of a result in chapter 3. We give a generalization of this result, to the joint real higher rank numerical range of n, 2×2 symmetric matrices. McIntosh also gives the following parametrization for theorem 2.7;

$$R(A) = \Big\{ \frac{1}{2}(a_{11} - a_{22})cos2\theta + \frac{1}{2}(a_{12} + a_{21})sin2\theta + \frac{1}{2}(a_{11} + a_{22}) \mid 0 \leq \theta \leq \pi \Big\}.$$

The next result is likely to be the most important as it is essentially the Toeplitz-Hausdorff theorem for the real numerical range.

Theorem 2.8. *For dimension $n \geq 3$, $R(A)$ is a convex subset of $\mathbb{R}$.*

Proof. First define the function $f : \mathcal{S}^{n-1} \to \mathbb{C}$ by $f(x) = \langle Ax, x \rangle$. Recall that $\mathcal{S}^{n-1}$ denotes the unit sphere in dimension n. Then $R(A) = f(\mathcal{S}^{n-1})$. Since $\mathcal{S}^{n-1}$ is connected and compact, if f is a continuous mapping, then the image of $\mathcal{S}^{n-1}$ under f will also be connected and compact. Hence f will actually map $\mathcal{S}^{n-1}$ to a closed and bounded interval. Our result follows. □

Following this we have two other results. These both assume A is a complex, symmetric matrix of dimension n.

Theorem 2.9. $R(A) = R(\frac{A+A^T}{2})$.

Proof. The proof of this is quite simple, and follows from the definition of the real numerical range. Let $\lambda = x^T Ax$. $R(\frac{A+A^T}{2})$ is the set of elements of the form $x^T(\frac{A+A^T}{2})x = \frac{1}{2}x^T Ax + \frac{1}{2}x^T Ax = \frac{1}{2}x^T Ax + \frac{1}{2}(x^T Ax)^T = \frac{1}{2}\lambda + \frac{1}{2}\lambda = \lambda$, which is precisely the definition of $R(A)$. □

Theorem 2.10. *For any $A \in \mathcal{M}_n(\mathbb{R})$, $W(A) = conv(R(A))$.*

Proof. We give the identical construction McIntosh provides in his paper. Here we can first note that

$$W(A) = R\left(\begin{pmatrix} A & iA \\ -iA & A \end{pmatrix}\right) = R\left(\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}\right).$$

This is now equivalent to the set

$$\{||v||^2(A(||v||^{-1}v), ||v||^{-1}v) + ||u||^2(A(||u||^{-1}u), ||u||^{-1}u) \mid u, v \in \mathbb{R}^n \wedge ||v||^2 + ||u||^2 = 1\}$$

which is exactly the convex hull of $R(A)$. □

This essentially summarizes the results of McIntosh's paper. We now give a brief synopsis of [6].

We will start off by stating a theorem that, is much in the same spirit as the convexity result stated by McIntosh.

Theorem 2.11. *Let A_1 and A_2 be real symmetric matrices of size n. Then $R(A_1, A_2)$ (that is the joint real numerical range of A_1 and A_2) is convex for $n \geq 3$.*

McIntosh provided the result for a single matrix A, showing that the real numerical range is convex for $n > 2$. Here, we have a kind of joint-analogue of his result, specifically for two real symmetric matrices.

Theorem 2.12. *Let A_1 and A_2 be real symmetric matrices of size n. $R(A_1, A_2) = W(A_1, A_2)$ for $n > 2$.*

Proof. We again give the reader an idea of the proof. Clearly $R(A_1, A_2) \subset W(A_1, A_2)$ so we need only show that $R(A_1, A_2) \supset W(A_1, A_2)$. Consider any point $(a_1, a_2) \in W(A_1, A_2)$ then for some norm one vector x we have

$$a_1 = x^T A_1 x$$

and

$$a_2 = x^T A_2 x$$

Now consider a decomposition of x in the form $x_1 + ix_2$ where x_1 and x_2 are entirely real. Then one may re-write a_1 and a_2 in the form

$$x_1^T A_1 x_1 + x_2^T A_1 x_2$$

and

$$x_1^T A_2 x_1 + x_2^T A_2 x_2$$

respectively. Our result now follows from the convexity of $R(A_1, A_2)$. □

Among these two results, there are others involving the joint numerical range of three Hermitian matrices. For more information we encourage the reader to look through Brickman's paper.

2.0.4 The complex higher rank numerical range $\Lambda_k(A)$

The higher rank numerical range is an extension of the ordinary numerical range. Pertinent information regarding this object is discussed here.

Definition 2.17. *Let $\mathcal{P}_k$ denote the set of all orthogonal projectors of rank k. The $k-$rank numerical range of a finite dimensional operator A is then given by the set*

$$\Lambda_k(A) = \{\lambda \in \mathbb{C} | PAP = \lambda P \ with \ P \in \mathcal{P}_k\}$$

Remark: When $k = 1$ we simply get back the (ordinary) numerical range.

In similar spirit to what we provided for the ordinary numerical range previously, we list a number of properties associated with the complex higher rank numerical range. Initially these appeared in [10].

Proposition 2.1. *Let $k \leq n$ and M and N any complex matrices of size n. Also let c_1 and c_2 be some constants. Then the following all hold true.*

1. $\Lambda_k(c_1 M + c_2 I) = c_1 \Lambda_k(M) + c_2$
2. $\Lambda_k(M \oplus N) \subseteq \Lambda_k(M) \cup \Lambda_k(N)$
3. $\Lambda_k(M) \subseteq \Lambda_k(Re(M)) + i\Lambda_k(Im(M))$
4. $\Lambda_k(M^*) = \overline{\Lambda_k(M)}$

In particular, we have a nice result regarding Hermitian matrices in particular.

Proposition 2.2. *The higher rank numerical ranges are nested for any operator A. I.e. for any $k \in [1, n]$ we have $\lambda_{k+1}(A) \subset \lambda_k(A)$. This can also be seen as*

$$\Lambda_n(A) \subseteq \Lambda_{n-1}(A) \subseteq \cdots \subseteq \Lambda_2(A) \subseteq \Lambda_1(A) = W(A).$$

Proof. It is not difficult to prove these inlcusions. It is sufficient to prove that for any arbitrary $k < n$, one has $\Lambda_k(E) \subseteq \Lambda_{k-1}(E)$. If an element $\lambda \in \Lambda_k(E)$ then there exist projections P_k such that $P_k E P_k = \lambda P_k$. Next note that we have the following relationship; $P_{k-1}P_k = P_k P_{k-1} = P_{k-1}$. Thus we can use this projection sending elements to some $k-1$ dimensional subspace of $\mathcal{H}$ to write

$$P_{k-1}P_k E P_k P_{k-1} = P_{k-1}\lambda P_k P_{k-1}$$

We note that the left hand side is simply just $P_{k-1}EP_{k-1}$. The right hand side also simplifies nicely;

$$P_{k-1}\lambda P_k P_{k-1} = \lambda P_{k-1}P_k P_{k-1} = \lambda P_{k-1}(P_{k-1}) = \lambda P_{k-1}^2 = \lambda P_{k-1}$$

Here we have just applied basic facts regarding projections (i.e. $P^2 = P$). Hence we come to $P_{k-1}EP_{k-1} = \lambda P_{k-1}$ and thus $\lambda \in \Lambda_{k-1}(E)$. □

Not much more needs to be said here. For more information regarding how the higher rank numerical range manifests itself in quantum information theory, we point the reader to the section contained in chapter four titled *Quantum error correction and the k- rank numerical range.*

2.0.5 The real higher rank numerical range $R_k(A)$

The main results of this thesis pertain to the real higher rank numerical range. Formally, this can be defined in a very similar way to that of the complex higher rank numerical range.

Definition 2.18. *Let $A \in \mathcal{M}_n(\mathbb{R})$. Then the* **real rank-$k$ numerical range of** *A, denoted $R_k(A)$ is the set*

$$\{\lambda \in \mathbb{R} : P_k A P_k = \lambda P_k\}$$

where P_k ranges over all $k-$rank, real orthogonal projection. We put $R(A) := R_1(A)$.

There is not a consummate amount of research done into this, as there were no known application of it, until this paper (from the perspective of a pure mathematical point of view). The problem of asking what cross sections running through the origin of an ellipsoid form hyperspherical cross sections, can be reformulated as a problem involving the real higher rank numerical range. This discussion is reserved for chapter 3.

2.0.6 The joint numerical range (real and complex)

One can look specifically at the numerical ranges of a single operator, however, it is also possible to look at the numerical range for a collection of operators, simultaneously. This is what is meant by the joint numerical range. Formally we have,

Definition 2.19. *Let $A = (A_1, A_2, \ldots, A_n)$ be a collection of n complex square matrices. Then the* ***joint numerical range*** *of A is the set of n dimensional vectors,*

$$\Big\{(x^*A_1x, x^*A_2x, \ldots, x^*A_nx)\Big\}$$

subject to $x^*x = 1$.

One can also modify this definition to get the real joint numerical range. This is done by taking x to be entirely real. The joint real numerical range has been studied far less in comparison to its complex counterpart. However, in the subsequent chapter, we provide a result for the joint case of n symmetric matrices of size 2.

Finally we remark that there is yet another extension to the joint case, that is the joint higher rank numerical range. The definition of the complex case is given below. Only a subtle change in this definition will give the real case, i.e. making the vector x real.

Definition 2.20. *Given a collection of m square matrices over $\mathbb{C}$, denoted with $A = (A_1, \ldots, A_m)$, the joint rank k numerical range is the set*

$$\Lambda_k(A) = \{(\lambda_1, \lambda_2, \ldots, \lambda_m)\}$$

with a rank k projection P_k such that $P_kA_iP_k = \lambda_iP_k$ for any $1 \leq i \leq m$.

2.0.7 Hyperboloids and ellipsoids

The geometry of these objects outlined previously are interesting to visualize and think about in higher dimensions. The purpose of this section is to introduce some

of the notions regarding conic sections, i.e. what notation will be used, naming conventions, various properties of hyperbolas and ellipses, and so forth. More of a connection is given in the subsequent chapter regarding the appearance of these objects in the context of numerical ranges.

It should be worth iterating again that one of the initial questions posed beginning this research was- what particular cross sections of an ellipse in n dimensions gives a hypersphere of dimension k, where $k < n$? We now pave the way for some remarks on these objects in arbitrary dimensions.

Definition 2.21. *An **ellipsoid** in n dimensional Euclidean space is given as the set*

$$\mathcal{E}_n = \left\{ x \in \mathbb{R}^n \mid \sum_{k=1}^{n} \frac{x_k^2}{c_k^2} = 1 \right\} = \left\{ x \in \mathbb{R}^n \mid x^T E x = 1, E > 0 \right\}$$

Here we have two equivalent ways of imagining the ellipsoid. The first is a standard way of defining it based on its semi-axes and the latter (which we will be interested in) is to write the same summation, $\sum_{k=1}^{n} \frac{x_k^2}{c_k^2} = 1$ as an inner product, i.e. $x^T E x$ for some positive definite matrix E. Note that the eigenvalues of E are precisely each of the reciprocal squares of the c_k's in the first definition.

Definition 2.22. *In dimension 2, a **hyperbola** is given by the expression*

$$\frac{x^2}{c^2} - \frac{y^2}{d^2} = 1$$

for positive real numbers c and d. One can also think of this object as plotting the curve

$$y = \frac{d}{c}\sqrt{x^2 - c^2}$$

and its reflection over the x axis.

We can now define what a hyperboloid is, in much the same way we did for the ellipsoid.

Definition 2.23. *A **hyperboloid** in dimension $n > 3$ is expressed by*

$$\left\{x \in \mathbb{R}^n |\ x^T H x = 1\right\} = \left\{x \in \mathbb{R}^n |\ \sum_{i=1}^{k} \frac{x_i^2}{c_i^2} - \sum_{i=k+1}^{n} \frac{x_i^2}{d_i^2} = 1\right\}$$

with $n > k$ and $c_i, d_i > 0,\ \forall i \in [1, n]$.

The only real change we have to make to move to this conic section, is to change the number of positive eigenvalues the matrix H carries in the definition above.

Example 2.3. *In dimension 3, one can have two different types of hyperboloids which are respectively given as*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

and

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

The first of which is known as the one sheeted hyperboloid which appears something of the shape

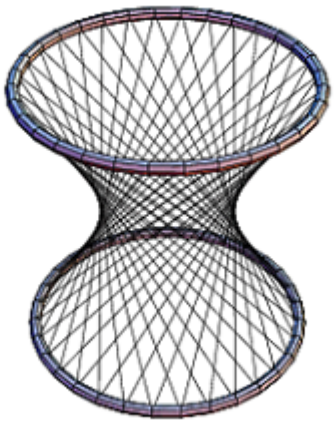

Figure 2.1: Image of a one-sheeted hyperbola in three dimensional space. Taken from [35].

and the latter known as the two sheeted hyperboloid

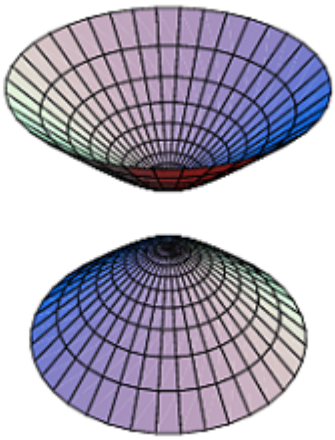

Figure 2.2: Image of a two sheeted hyperbola in three dimensional space. Taken from [36].

Example 2.4. *Although appearing in several different contexts in mathematics, hyperboloids also arise in areas of theoretical physics. In natural units (taking* $c = 1$*)*

the spacetime interval

$$ds^2 = dt^2 - (dx^2 + dy^2 + dz^2)$$

can be regarded as a 4 dimensional hyperboloid. See [19] for more on the physics behind this.

As mentioned, the primary difference between an ellipse and a hyperbola (and different types of hyperbolas for that matter) is the number of positive and negative eigenvalues its defining matrix has. Explicitly, given a conic section

$$x^T Ax = 1$$

if A is positive definite, then $x^T Ax = 1$ parametrizes an ellipse with semi-axes equal to the reciprocal square roots of the eigenvalues of A, and likewise, if A contains anywhere between 1 and $n-1$ negative eigenvalues, it will parametrize a hyperboloid. For this reason, it appears important to keep track of the number of positive and negative eigenvalues there are, hence we can look to the inertia of the matrix.

Remark: In order for H to describe a hyperboloid, it must have at least one positive eigenvalue. Having all negative eigenvalues will mean that the expression $x^T Hx = 1$ is never satisfied.

Transitioning into some matrix terminology, we discuss the inertia of matrix. This will be helpful as we develop a naming convention for conic sections in any dimension.

Definition 2.24. *Let A be an $n \times n$ matrix with real eigenvalues. The inertia of A denoted $In(A)$ is the triple*

$$(n_+, n_-, n_0)$$

where n_+, n_-, n_0 are the number of positive, negative and zero eigenvalues respectively.

Theorem 2.13. *(Sylvester's law of inertia) Let A and B be Hermitian matrices of size n. Then there exists a non-singular matrix S such that $A = S^*BS$ if and only if both A and B have the same inertia.*

Remark: Sylvester's law of inertia essentially gives a way of mapping 'equivalent' conic sections to and from one another (about the origin). Since Sylvester's law states that two congruent matrices must have the same inertia, they therefore have the same number of positive eigenvalues (and negative) which are the defining features of a certain conic section type in n dimensions. This implies that the transformation being applied to some ellipsoid, or hyperboloid will preserve its type and thus only stretch and/or rotate the conic section.

With this in mind, we can now develop a naming convention for these objects. Our convention will be be broken down into two classes- one for ellipses and the other for hyperbolas.

Concerning the latter of these conic sections, we develop names for hyperboloids in any dimension based on the inertia of their defining matrix (i.e. where $x^THx = 1$, we look to the inertia of H). We then prefix the term *hyperboloid* with this triplet to

look something like-

$$(n_+, n_-, n_0) - \textbf{hyperboloid}$$

This tells the reader that it is of $(n_+) + (n_-) + (n_0)$ dimensions with n_+ positive eigenvalues and n_- negative respectively.

Example: In three dimensions, the one sheeted hyperbola can be equivalently named the $(2, 1, 0) - \textit{hyperbola}$. Similarly, the two sheeted hyperbola can be named the $(1, 2, 0) - \textit{hyperbola}$.

The former is quite simple and straightforward. We will call an ellipse in 2 dimensions, simply just that. In dimensions greater than 2, we will use the terms *ellipsoid* or *n dimensional ellipse* or even *hyperellipse* interchangeably. It will be clear based on the context what dimension we are working in. Because any 2 ellipses in some arbitrary dimension will have the same inertia, there is no need to prefix their name with the inertia. There is only one type of ellipse. Concordantly, if one wanted to be consistent with the rules above, we may also refer to any ellipse in n dimensions as the $(n, 0, 0)-$ ellipse.

Proposition 2.3. *(Geometry of inertias) Let $x^T Ax = 1$ be an n dimensional conic section with $In(A) = (n_+, n_-, n_0)$ and $n_+ \neq n$. Then if $1 < n_+ \leq n-1$ the hyperboloid will be connected.*

Proof. Let us prove the contrapositive, that is if $n_+ = 1$, then the $(n_+, n_-, n_0)-$ hyperboloid will not be connected. Suppose first we cut through this hyperboloid with a hyperplane of dimension $n - 1$. This will change the inertia of A from $(1, n_-, n_0)$

to $(0, n_-, n_0)$. In what follows, we use A' to denote the new matrix after we cut through A with an $n-1$ dimensional hyperplane. Hence, because the spectrum of A' is entirely nonpositive, A' no longer defines a hyperboloid as the inner product $x^T A' x$ can never equal 1 (as there are no positive eigenvalues to attain this). It follows that since $x^T A' x$ is not a hyperboloid, $x^T A x$ (or equivalently the $(1, n_-, n_0)-$ hyperboloid) cannot be connected. □

Chapter 3

Conic sections and the real k- rank numerical range

We now introduce the main results of the paper. We have divided this chapter into two sections, the first discussing new results regarding the real numerical range and the real higher rank numerical range. We then discuss the original idea of how investigation into the real numerical range came about, with some preliminary findings.

3.0.1 Results on the real numerical range

Before proceeding, we give a list of elementary properties regarding the real higher rank numerical range. A proof of each is omitted as they are fairly intuitive and clear from the definition of R_k. In what follows, let $a, b \in \mathbb{R}$, $A, B \in \mathcal{M}_n(\mathbb{R})$ and I the

identity.

1. $R_k(aA + bI) = aR_k(A) + b$

2. $R_k(A^T) = R_k(A)$

3. $R_k(A) \cup R_k(B) \subseteq R_k(A \oplus B)$

4. $R_{k_1}(A) \cap R_{k_2}(B) \subseteq R_{k_1+k_2}(A \oplus B)$

5. $R_n(A) \subseteq R_{n-1}(A) \subseteq \cdots \subseteq R_2(A) \subseteq R_1(A) = R(A)$

We state this next result as a proposition, as it is not as intuitive as the list above, however, still important.

Proposition 3.4. *For any $A \in \mathcal{M}_n(\mathbb{R})$, $R_k(A)$ forms a compact set.*

Proof. Since A is finite dimensional, it suffices to show that R_k is closed and bounded. For boundedness, if we recall the definition of R_k, we have $PAP = \lambda P$ and we can consider the norm of PAP. We find that $||PAP|| \leq ||P||||AP|| \leq ||P||(||A||)||P|| \leq ||A||$. (Here the final bound follows from $||P|| \leq 1$). And so we conclude that $||A|| \geq ||\lambda P|| \geq |\lambda|$. Thus we see that the sequence of λ is bounded by the norm of A.

To prove closedness, we can look again at the sequence $\{\lambda_n\}$. From our proof above, we know that $\{\lambda_n\}$ is bounded and because this sequence is also real, we can apply the Bolzano-Weierstrass theorem to assert that $\{\lambda_n\}$ has a convergent subsequence. Let λ' denote the limit point of this subsequence. Then there exists a

sequence of projections $\{P_n\}$ such that $AP_n = \lambda_n P_n$. Because the sequence of $\{P_n\}$ is bounded, it has a convergent subsequence which we denote $\{P_{n_k}\}$ whose limit is P. This implies that $\{P_{n_k} A P_{n_k}\}$ converges to PAP which is equivalent to $\lambda' P$. □

Proposition 3.5. *For any $A \in \mathcal{M}_n(\mathbb{R})$ we have $Re(\lambda) \in R(A)$ where λ is an eigenvalue of A.*

Proof. If λ is in the spectrum of A then there exists a non-zero vector x such that $Ax = \lambda x$. Since λ could be complex, we can write $\lambda = a + ib$ where $a = \mathrm{Re}(\lambda)$ and $b = \mathrm{Im}(\lambda)$. Similarly, $x = u + iv$. Then,

$$\begin{aligned} Ax &= \lambda x \\ A(u+iv) &= (a+ib)(u+iv) \\ Au + iAv &= (au - bv) + i(av + bu) \end{aligned}$$

This implies that $Au = au - bv$ and $Av = (av + bu)$ (simply equating real and imaginary parts on either side). Now,

$$\begin{aligned} u^T(Au) = u^T(au - bv) &= a||u||^2 - bu^T v \\ v^T(Av) = v^T(av + bu) &= a||v||^2 + bv^T u \end{aligned}$$

Here, summing these two equations will give us

$$u^T Au + v^T Av = (||u||^2 + ||v||^2)a \tag{3.1}$$

since $bu^T v = bv^T u$. Now, if we make the following substitutions;

$$\begin{aligned} u &\rightarrow ||u||\tilde{u} \\ v &\rightarrow ||v||\tilde{v} \end{aligned}$$

3.1 becomes

$$a = ||u||^2(\tilde{u}^T A\tilde{u}) + ||v||^2(\tilde{y}^T A\tilde{y})$$

And we see that the real part of λ, is a convex combination of elements in the real numerical range, and since $R(A)$ is convex, it follows that $a \in R(A)$. □

Theorem 3.14. *Let $A \in \mathcal{M}_n(\mathbb{R})$ be a real symmetric matrix with eigenvalues $\{\lambda_k(A)\}_{k=1}^n$ listed in descending order and fix a positive integer $k \geq 1$. Then we have*

$$R_k(A) = \Big[\lambda_k(A), \lambda_{n-k+1}(A)\Big].$$

Proof. We note the following proof is identical to that of theorem 1 in [10] for real projections $\mathcal{P}_k$. For completeness, we do state it here however.

Let $\lambda \in R_k(A)$ and let P_k be a rank-k projection with $P_k A P_k = \lambda P_k$. If $V : \mathbb{R}^{N-k+1} \rightarrow \mathbb{R}^N$ is an isometry, then the subspace $P_k\mathbb{R}^N$ and the range space $VV^T(\mathbb{R}^N)$ have non-zero intersection. Thus, there exists a unit vector $|\psi\rangle \in \mathbb{R}^N$ such that

$|\psi\rangle = P_k|\psi\rangle = VV^T|\psi\rangle$. Let $|\psi'\rangle$ be the unit vector in $\mathbb{R}^{N-k+1}$ given by $|\psi'\rangle = V^T|\psi\rangle$. Then we have

$$\begin{aligned}\langle V^TAV\psi'||\psi'\rangle &= \langle A\psi||\psi\rangle \\ &= \langle P_kAP_k\psi||\psi\rangle = \lambda\langle P_k\psi||\psi\rangle = \lambda.\end{aligned}$$

Hence we have shown that λ belongs to $R(V^TAV)$. As $V : \mathbb{R}^{N-k+1} \to \mathbb{R}^N$ was an arbitrary isometry, it follows that $R_k(A)$ is contained in the intersection of all such numerical ranges $R(V^TAV)$.

Next, let $\{|i\rangle : 1 \leq i \leq N-k+1\}$ be a fixed orthonormal basis for $\mathbb{R}^{N-k+1}$ and let $\{|\psi_i\rangle\}$ be an orthonormal basis for $\mathbb{R}^N$ of eigenvectors for A corresponding to the eigenvalues $a_1, \ldots, a_N$. Consider two linear isometries $V_1, V_2 : \mathbb{R}^{N-k+1} \to \mathbb{R}^N$ defined by $V_1(|i\rangle) = |\psi_i\rangle, V_2(|i\rangle) = |\psi_{N-i+1}\rangle$.

Then $V_1^TAV_1$ and $V_2^TAV_2$ are operators on $\mathbb{R}^{N-k+1}$ that are diagonal with respect to the basis $\{|i\rangle\}$, and we have $R(V_1^TAV_1) = [a_1, a_{N-k+1}]$ and $R(V_2^TAV_2) = [a_k, a_N]$. It follows that

$$\begin{aligned}R(A) \subseteq \bigcap_V R(V^TAV) &\subseteq R(V_1^TAV_1)\bigcap R(V_2^TAV_2) \\ &= [a_k, a_{N-k+1}].\end{aligned}$$

We complete the proof by showing $R(A)$ contains the set $[a_k, a_{N-k+1}]$ when $a_k \leq a_{N-k+1}$. Suppose first that $a_{N+1-k} > a_k$ (and so $2k \leq N$). Fix λ in the interval $[a_k, a_{N+1-k}]$. We shall directly construct a rank-k projection P_k such that $P_kAP_k = \lambda P_k$. Consider the set of k pairs $\{a_{k+1-j}, a_{N-k+j}\}$, $1 \leq j \leq k$. As a notational

convenience we shall write $\{b_j, b'_j\}$ for the ordered pair $\{a_{k+1-j}, a_{N-k+j}\}$, and so $b_j > b'_j$. (The following construction may be easily modified for any joint partition of the sets $\{a_N, \ldots, a_{N-k+1}\}$ and $\{a_k, \ldots, a_1\}$ into ordered pairs.)

We may write A, up to unitary equivalence, as a direct sum

$$A = \left(\oplus_j A_j \right) \oplus B,$$

where each A_j is a diagonal 2×2 matrix with spectrum $\{b_j, b'_j\}$, and B is either vacuous, or is the diagonal matrix with diagonal entries $\{a_{k+1}, \ldots, a_{N-k}\}$. As λ satisfies,

$$\lambda \in [a_k, a_{N-k+1}] \subseteq [b_j, b'_j] = R(A_j) \quad \forall\, 1 \leq j \leq k,$$

we may find angles θ_j such that

$$\lambda = b_j \cos^2 \theta_j + b'_j \sin^2 \theta_j \quad \forall\, 1 \leq j \leq k.$$

Now define an orthonormal set of k vectors by

$$|\phi_j\rangle = \cos\theta_j |\psi_{N-k+j}\rangle + \sin\theta_j |\psi_{k-j+1}\rangle \quad \forall\, 1 \leq j \leq k,$$

and the rank-k projection P_k onto the subspace spanned by these vectors;

$$P = |\phi_1\rangle\langle\phi_1| + |\phi_2\rangle\langle\phi_2| + \ldots + |\phi_k\rangle\langle\phi_k|.$$

It follows that $P_k A P_k = \lambda P_k$. Indeed, observe that for $1 \leq j \leq k$ we have

$$\begin{aligned}
\langle A\phi_1 || \phi_j \rangle &= \cos\theta_1 \langle A\psi_N || \phi_j \rangle + \sin\theta_1 \langle A\psi_1 || \phi_j \rangle \\
&= a_N \cos\theta_1 \langle \psi_N || \phi_j \rangle + a_1 \sin\theta_1 \langle \psi_1 || \phi_j \rangle \\
&= b_1 \cos\theta_1 \cos\theta_j \delta_{N,N-1+j} + b_1' \sin\theta_1 \sin\theta_j \delta_{j,1} \\
&= \lambda \delta_{j,1}.
\end{aligned}$$

Similarly, $\langle A\phi_i || \phi_j \rangle = \lambda \delta_{ij}$ for $1 \leq i, j \leq k$.

The remaining case is characterized by the constraint $\lambda := a_k = a_{N-k+1}$. If, in addition, $a_{N-k+2} > a_{k-1}$, then we may split the sets $\{a_N, \ldots, a_{N-k+2}\}$ and $\{a_{k-2}, \ldots, a_1\}$ into pairs as above, and similarly define $k-1$ vectors $|\phi_1\rangle, \ldots, |\phi_{k-1}\rangle$. As the final vector we can take $|\phi_k\rangle := |\psi_k\rangle$, and define $P_k = \sum_{j=1}^{k} |\phi_j\rangle\langle\phi_j|$. If $a_{N-k+2} = a_{k-1}$, but $a_{N-k+3} > a_{k-2}$, then we will use $|\psi_k\rangle$ and $|\psi_{k-1}\rangle$ as two of the vectors. This process may be continued, if required, to account for degeneracies in the spectrum of A around the eigenvalue a_k, and construct a rank-k projection which yields $\lambda \in R_k(A)$. The result now follows. □

We can state an obvious result as a corollary of this theorem.

Corollary 3.15. *For any real symmetric matrix A, $R_k(A)$ is convex.*

Proof. By theorem 3.16, we know $R_k(A)$ is just some closed interval of $\mathbb{R}$ which is always convex. □

Open Problem: Given an arbitrary $A \in \mathcal{M}_n(\mathbb{R})$, is $R_k(A)$ convex for $k > 1$?

Proposition 3.6. *Let $A \in \mathcal{M}_n(\mathbb{R})$ then*

$$R(A) = R\Big(\frac{A + A^T}{2}\Big)$$

Proof. $R(A)$ is the set of all elements of the form $x^T Ax$ where $x \in \mathbb{R}^n$, hence $R(\frac{A+A^T}{2})$ is the set consisting of all elements of the form $x^T(\frac{A+A^T}{2})x$. We see that

$$x^T(\frac{A + A^T}{2})x = \frac{1}{2}(x^T Ax + x^T A^T x) = \frac{1}{2}(x^T Ax + (x^T Ax)^T)$$

and thus, since $x^T Ax \in \mathbb{R}$, $(x^T Ax)^T = x^T Ax$ and our result follows. □

We can state a similar result regarding the higher rank version of proposition 2. The equality that appeared in the previous proposition no longer holds as seen below.

Proposition 3.7. *Let $A \in \mathcal{M}_n(\mathbb{R})$, then we have*

$$R_k(A) \subseteq R_k\Big(\frac{A + A^T}{2}\Big) = \Big[\lambda_k\Big(\frac{A + A^T}{2}\Big), \lambda_{n-k+1}\Big(\frac{A + A^T}{2}\Big)\Big]$$

Proof. Let $\lambda \in R_k(A)$, then there must exist a real k- rank projection P_k such that $P_k A P_k = \lambda P_k$. Taking the transposition of both sides yields $P_k A^T P_k = \lambda P_k$ which implies that $P_k \frac{A+A^T}{2} P_k = \lambda P_k$ from which our result follows. □

The opposite containment however is not true in general. To show this we prove the following lemma.

Lemma 3.16. *Let $A \in \mathcal{M}_n(\mathbb{R})$. If there exists $k > \frac{n}{2}$ such that $0 \in R_k(A)$ then $rank(A) \leq 2n - 2k$.*

Proof. If A is a matrix and B is any submatrix formed by deleting m rows and m columns, then $rank(B) \geq rank(A) - 2m$. So if $B = 0$ and $m = n - k$, we get $rank(A) \leq 2n - 2k$. □

Now A be an n by n real skew-symmetric matrix. Then $\frac{A+A^T}{2} = 0$ and $R_k(0) = \{0\}$ for all k. We can choose A to be rank n if n is even and to be rank $n-1$ if n is odd. In either case $0 \notin R_k(A)$ if $k > \frac{n}{2}$ by the previous lemma.

We do however have the following result:

Theorem 3.17. *Let A be an n by n real matrix and $1 \leq k \leq n$. Then*

$$R_k\Big(\frac{A+A^T}{2}\Big) \subseteq R_{\lceil \frac{k}{2} \rceil}(A).$$

Proof. Let $\lambda \in R_k\Big(\frac{A+A^T}{2}\Big)$. Then we know by definition that the matrix $\frac{A+A^T}{2}$ must be unitarily equivalent to the matrix $\begin{pmatrix} \lambda I & * \\ * & * \end{pmatrix}$ This forces A to have the form $\begin{pmatrix} \lambda I + B & * \\ * & * \end{pmatrix}$ for some skew-symmetric matrix B. Since B is a skew symmetric matrix, we know its eigenvalues must of the form $\pm i\mu$, where $i = \sqrt{-1}$. Hence, the this $k \times k$ block will consist of nonzero eigenvalues in conjugate pairs $\lambda \pm i\mu_j$ together with an arbitrary number of zero eigenvalues. Now any eigenvectors of B corresponding to the nonzero eigenvalues in conjugate pairs are themselves of the form $u_j = v_j \pm iw_j$. Then we have

$$Bu_j = B(v_j \pm iw_j) = \pm i\mu_j(v_j \pm iw_j).$$

This implies that

$$\begin{aligned} Bv_j &= -\mu_j w_j \\ Bw_j &= \mu_j v_j \end{aligned}$$

Therefore, if we pick an ordered basis consisting of all of the v_js, then the eigenvectors corresponding to the eigenvalue zero and finally all of the w_js, we get

$$B = \begin{bmatrix} 0 & D \\ -D & 0 \end{bmatrix}$$

and thus,

$$\lambda I + B = \begin{bmatrix} \lambda I & D \\ -D & \lambda I \end{bmatrix}$$

Note that the upper left block must be at least $\lceil \frac{k}{2} \rceil$ in size, from which the result follows. □

Theorem 3.18. *Let $A = (A_1, A_2, ..., A_m)$ be a set of 2 by 2 real symmetric matrices. Then $R(A)$ is the boundary of an ellipse in $\mathbb{R}^m$. For sets of 2 by 2 real symmetric matrices the joint real numerical range is the boundary of the numerical range: i.e. $R(A) = \partial W(A)$.*

Proof. Along the same lines as the proof given by Chandler Davis (for the ordinary numerical range) in [4], we give a similar construction here for the result of the theorem.

We can write any A_i as a linear combination of Pauli matrices. For real scalars $\{\omega_0, \omega_x, \omega_y, \omega_z\}$ with $\omega_0^2 + \omega_x^2 + \omega_y^2 + \omega_z^2 = 1$ we have the decomposition

$$A_i = \omega_0^i I + \omega_x^i \sigma_x + \omega_y^i \sigma_y + \omega_z^i \sigma_z$$

Since A is real, we may take $\omega_y^i = 0$. Furthermore, because the vector x in the definition of the joint real numerical range of A must be real and have norm one, we may express x in the form $\begin{pmatrix} cos\theta \\ \\ sin\theta \end{pmatrix}$ for some $\theta \in [0, \pi]$. Now, for any $A_i \in A$, we have

$$\begin{aligned} x^T A_i x &= (cos\theta \ sin\theta) \begin{bmatrix} \omega_0^i + \omega_z^i & \omega_x^i \\ \\ \omega_x^i & \omega_0^i - \omega_z^i \end{bmatrix} \begin{bmatrix} cos\theta \\ \\ sin\theta \end{bmatrix} \\ &= (\omega_0^i + \omega_z^i) cos^2\theta + 2\omega_x^i sin\theta cos\theta + (\omega_0^i - \omega_z^i) sin^2\theta \\ &= \omega_0^i + \omega_x^i sin2\theta + \omega_z^i cos2\theta \end{aligned}$$

and hence for the m vector $(x^T A_1 x, x^T A_2 x, \ldots, x^T A_m x)$ we obtain

$$(\omega_0^1 + \omega_x^1 sin2\theta + \omega_z^1 cos2\theta, \ \omega_0^2 + \omega_x^2 sin2\theta + \omega_z^2 cos2\theta, \ \ldots, \ \omega_0^m + \omega_x^m sin2\theta + \omega_z^m cos2\theta)$$

$$\begin{aligned} &= \Big(\sum_{i=1}^{m} \omega_0^i, \ \sum_{i=1}^{m} \omega_x^i sin2\theta, \ \sum_{i=1}^{m} \omega_z^i cos2\theta \Big) \\ &= \vec{\omega_0} + \vec{\omega_x} sin2\theta + \vec{\omega_z} cos2\theta \end{aligned}$$

This parametrizes an ellipse centred at $\vec{\omega_0}$ with major and minor axes of $\vec{\omega_x}$ and $\vec{\omega_z}$ respectively. □

Lemma 3.19. *Let $A = (A_1, \ldots, A_n)$ be a collection of real symmetric matrices of dimension* 3. *Then any oblique cross section through the origin of $R(A)$ will be an ellipse.*

Proof. This proof follows the general structure as the proof of the real elliptical range theorem. It is simply an extension to matrices of size 3. We first note that any real unit norm vector can be expressed as

$$x = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}$$

Then we consider a basis for each of the real symmetric A_i. We will use five of the eight, size 3 Gell-mann matrices. In particular, we will use all the symmetric and diagonal ones. Hence, we may write for any given A_i a decomposition in this basis as

$$\begin{aligned} A_i &= \omega_1^i g_5^s + \omega_2^i g_7^d + \omega_3^i g_4^s + \omega_4^i g_6^s + \omega_5^i g_8^d \\ &= \begin{pmatrix} \omega_2^i + \frac{1}{\sqrt{3}}\omega_5^i & \omega_1^i & \omega_3^i \\ \omega_1^i & -\omega_2^i + \frac{1}{\sqrt{3}}\omega_5^i & \omega_4^i \\ \omega_3^i & \omega_4^i & -\frac{2}{\sqrt{3}}\omega_5^i \end{pmatrix} \end{aligned}$$

Thus, the matrix described in the expression above is for any given A_i and we are now able to compute $x^T A_i x$.

$$
\begin{aligned}
x^T A_i x &= \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}^T \begin{pmatrix} \omega_2^i + \frac{1}{\sqrt{3}}\omega_5^i & \omega_1^i & \omega_3^i \\ \omega_1^i & -\omega_2^i + \frac{1}{\sqrt{3}}\omega_5^i & \omega_4^i \\ \omega_3^i & \omega_4^i & -\frac{2}{\sqrt{3}}\omega_5^i \end{pmatrix} \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix} \\
&= \Big(\omega_2^i + \frac{1}{\sqrt{3}}\omega_5^i\Big)\sin^2\theta\cos^2\phi + 2\omega_1^i\sin^2\theta\sin\phi\cos\phi + 2\omega_3^i\sin\theta\cos\theta\cos\phi \\
&+ 2\omega_4^i\sin\theta\sin\phi\cos\theta - \frac{2}{\sqrt{3}}\omega_5^i\cos^2\theta \\
&= \Big(\omega_2^i + \frac{1}{\sqrt{3}}\omega_5^i\Big)\sin^2\theta\cos^2\phi + \omega_1^i\sin^2\theta\sin2\theta \\
&+ \Big(\omega_3^i\cos\phi + \omega_4^i\sin\phi\Big)\sin2\theta - \frac{2}{\sqrt{3}}\omega_5^i\cos^2\theta
\end{aligned}
$$

Using basic trigonometric identities we may simplify (and rearrange) this expression to read

$$
\begin{aligned}
x^T A_i x &= \Big(\omega_3^i\cos\phi + \omega_4^i\sin\phi\Big)\sin2\theta \\
&+ \Big(-\frac{1}{2}\Big((\omega_2^i + \frac{1}{\sqrt{3}}\omega_5^i)\cos^2\phi + \omega_1^i\sin2\phi\Big) - \frac{1}{\sqrt{3}}\omega_5^i\Big)\cos2\theta \\
&+ \Big(\frac{1}{2}\Big((\omega_2^i + \frac{1}{\sqrt{3}}\omega_5^i)\cos^2\phi + \omega_1^i\sin2\phi\Big) - \frac{1}{\sqrt{3}}\omega_5^i\Big) \\
&= f(\phi)\sin 2\theta + g(\phi)\cos 2\theta + h(\phi)
\end{aligned}
$$

which parametrizes an ellipse if ϕ is held constant and θ runs through from 0 to π. Therefore, although this object is likely not to be an ellipse, it has elliptic nature inherent in it. □

This next result is a generalization of the previous two results. It encapsulates the pattern seen in two and three dimensions and extends them for higher dimensions.

Theorem 3.20. *(Elliptical path theorem) Let $A = (A_1, \ldots, A_m)$ be a string of m real symmetric matrices of dimension n. Then for any two elements r_1 and r_2 in the joint real numerical range of A, there always exists an elliptical path joining r_1 and r_2 that is entirely contained in $R(A)$.*

Proposition 3.8. *Let $A = (A_1, A_2, ..., A_m)$ be a set of n by n real symmetric matrices. Then $R(A)$ lies in an affine subspace of $\mathbb{R}^m$ which has dimension $\frac{n(n+1)}{2} - 1$.*

Proof. In what follows, we provide a counting argument for the dimension $R(A)$ lies in. We can us the generalized Gell-Mann matrices as a basis for each A_j. If we recall, the generalized Gell-Mann matrices can be divided into three classes; a symmetric class given by

$$\Lambda_{j,k}^{sym} = E_{j,k} + E_{k,j}$$

an antisymmetric class

$$\Lambda_{j,k}^{anti} = i(E_{j,k} - E_{k,j})$$

and finally a class of diagonal elements given by

$$\Lambda_l^{diag} = \sqrt{\frac{2}{l(l+1)}}\Big(\sum_{j=1}^{l} E_{j,j} - lE_{l+1,l+1}\Big)$$

Here both the symmetric and antisymmetric parts contain $\frac{n(n-1)}{2}$ elements, while the diagonal matrices contain $n-1$ elements. Hence in total there are

$$\frac{n(n-1)}{2} + \frac{n(n-1)}{2} + n - 1 = n^2 - 1$$

Gell-Mann matrices of size n.

However, because each A_j is symmetric, we only require our basis to include the symmetric and diagonal Gell-Mann matrices, hence the joint real numerical range of A will lie in a subspace of dimension

$$\frac{n(n-1)}{2} + n - 1 = \frac{n(n+1)}{2} - 1$$

□

We can prove a similar estimate for the ordinary (complex) joint numerical range in a similar way.

Proposition 3.9. *Let $A = (A_1, A_2, ..., A_m)$ be a set of n by n Hermitian matrices. Then $W(A)$ lies in an affine subspace of $\mathbb{R}^m$ which has dimension $n^2 - 1$.*

Proof. Similar to the proof of proposition 4, we may use the generalized Gell-Mann matrices as a basis for each A_j. Since each A_j is Hermitian now (and not necessarily real) we require all $n^2 - 1$ Gell-Mann matrices to be in the basis for A_j. Hence, $dim(W(A)) = n^2 - 1$ □

The following is a consequence of proposition 2.4 in [25].

Proposition 3.10. *Let $A \in S(\mathcal{H})^m$ such that each A_i are linearly independent, then $R_k(A)$ is non-empty provided*

$$dim(\mathcal{H}) \geq (k-1)(m+1)^2$$

where $S(\mathcal{H})^m$ is the real m dimensional linear space of self-adjoint operators in $B(\mathcal{H})$.

Proof. We provide a sketch of the proof here. Start by assuming the dimension of the Hilbert space to be $(m+1)^2(k-1)$ and also set $p = (m+1)(k-1)+1$. Then choose an eigenvector x_1 of unit length correpsonding to A_1. If we then consider a second norm one vector x_2 such that it is orthogonal to each of the vectors x_1, A_2x_1, ..., A_mx_1 then we may look at the set of p vectors- $\{x_1, \ldots, x_p\}$. Constructing this set will help in developing orthogonality conditions on each of the x_i's. We may eventually apply Tverberg's theorem and the result follows shortly after. □

3.0.2 Hyperspherical cross sections of ellipsoids and hyperboloids

Let us call the n by n matrix above E. Therefore this ellipsoid will have a hyperspherical cross-section of radius R through the origin if and only if there exists a unitary matrix U such that

$$UEU^{-1} = \left(\begin{array}{c|c} \frac{1}{R^2} I_{k\times k} & * \\ \hline * & * \end{array}\right) \tag{3.2}$$

Remark: So if we are able to find a subset S of n- dimensional, real vectors that give us this feature, we will be able to find such spheres, recognized from the ellipse.

Proposition 3.11. *Let $k, n \in \mathbb{N}$ with $1 \leq k \leq n$. Then an n- dimensional ellipsoid*

centred at the origin with n principal semiaxes of lengths $c_1 \geq c_2 \geq c_3 \geq ... \geq c_n$ *will have a k-dimensional hyperspherical cross-section through the origin of radius* R *if and only if* $c_k \geq R \geq c_{n-k+1}$.

Proof. Follows from theorem 3.16. □

Example 3.5. *Suppose we take the ellipse*

$$\frac{x_1^2}{2} + \frac{x_2^2}{2} + \frac{x_3^2}{5} + \frac{x_4^2}{2} = 1$$

then the cross section of $\mathcal{E}_4$ *defined at* $x_3 = 0$ *will generate a sphere of radius* $\sqrt{2}$. *One could also choose to look in the* 2 *dimensional subspaces spanned by each of* $\{x_1, x_2\}$, $\{x_2, x_4\}$ *and* $\{x_1, x_4\}$ *where in which they would see a circle of radius* $\sqrt{2}$. *However, looking at this projection is just looking at cross sections of the sphere of radius* $\sqrt{2}$, *so for this reason we wish to find the largest possible hypersphere of* $\mathcal{E}_n$.

Proof. One can take the change of basis matrix U to be the operator $I_2 \oplus \sigma_x$ and in doing so we obtain

$$E \longrightarrow UEU^{-1} = (I_2 \oplus \sigma_x)E(I_2 \oplus \sigma_x)^{-1} = \text{diag}\{1/2, 1/2, 1/2, 1/5\}$$

This new matrix is an ellipse viewed in the $\{x_1, x_2, x_4, x_3\}$ ordered basis. And the subspace spanned by $\{x_1, x_2, x_4\}$ defines the sphere already mentioned. Equivalently we can look at the real higher rank numerical range, i.e.

$$\Lambda_2((I_2 \oplus \sigma_x)E(I_2 \oplus \sigma_x)^{-1}) = [1/2]$$

and we arrive at the conclusion proposition 3.12 is referring. □

Example 3.6. *For a slightly harder example, one that is less obvious to see, consider the ellipse given by*

$$\frac{x^2}{4} + \frac{y^2}{10} + \frac{z^2}{16} = 1$$

Viewed in the standard basis, this is equivalent to

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 16 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 1 \tag{3.3}$$

And here we would require some unitary U such that the diagonal matrix seen in 3.3 matches the form of 3.2.

If we take

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

then

$$UEU^{-1} = \begin{bmatrix} 10 & 0 & -6 \\ 0 & 10 & 0 \\ -6 & 0 & 10 \end{bmatrix} = \left(\begin{array}{c|c} 10I_{2\times 2} & * \\ \hline * & * \end{array}\right)$$

and we deduce that certain cross sections of this ellipse in 3 dimensions will give circles of radius $\frac{1}{\sqrt{10}}$ about the origin.

Chapter 4

Connections to quantum information and the surface area entanglement measure

4.1 Quantum information preliminaries

Quantum information theory is the study of how information can be processed at the subatomic level. Here, it may be beneficial to compare quantum computation to its classical counterpart, namely, classical computing. In ordinary classical information theory, information itself is represented by bits that can only ever assume one of two possible states at any given time. These states are typically given as the 0-state or the 1- state, which make up a binary set. Equivalently, the off or on state

respectively. Classical computers process information by reading input in the form of a string of binary digits. For example, each letter of the alphabet will have its own unique binary representation to distinguish it from the others. More relevant for our discussion however, is the computing ability associated to any classical machine. These machines are really only limited to checking one independent case of a problem at a time. For example, if a classical computer was given the following maze seen in the figure below, one may ask what is the (optimal) solution to run from the start to finish? Classical computers may only ever check one possible route at a time, and here in lies the distinction between computing information classically versus at the quantum scale.

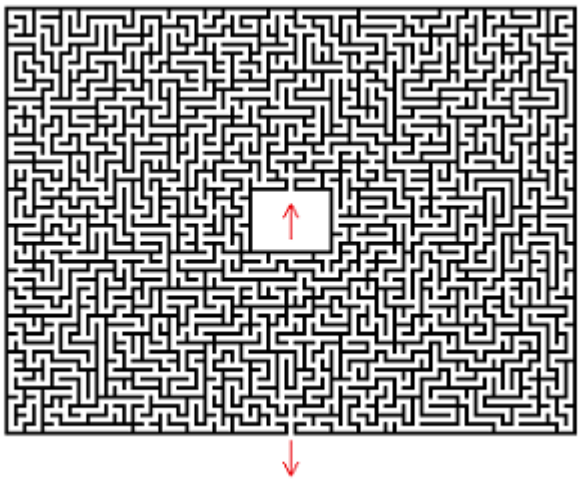

Figure 4.1: Depiction of a maze. Item taken from Google images.

Quantum computers (currently still in their preliminary stages of development) can offer a far quicker way of checking solutions to problems (and even solving them

for that matter!). In the same problem looked at in the classical sense with the maze, quantum machines may be able to check multiple cases of the problem simultaneously, thus cutting down on both memory space and computing time, two features that are quite important in the efficiency of these machines. Hence, if we come full circle, quantum bits of information can not only be in the 0 or 1 states, but any superposition of these two. It is this fundamental distinction between the two that offer an entire new realm of computation. As recently as in [5], it is beginning to become very evident of the promise these futuristic machines hold. On a final note, for more intricate examples of quantum computing, one may be interested in various algorithms such as Shor's algorithm, Deutsch's algorithm or other well known quantum algorithms.

4.1.1 An introduction to basic quantum information theory

As briefly introduced in the opening remarks of the section, quantum information is represented by the superposition of 0 and 1 states (in the case of a dimension 2 basis), mathematically, any unit norm vector $|\psi\rangle$ is expressed as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

subject to the constraint $|\alpha|^2+|\beta|^2 = 1$. One may naively dismiss this subtly, however, it cannot be ignored. The explanation for this has to do with probabilities- in other words the squares of α and β correspond to probabilities. The likelihood you would observe the particle in each of the states independently are exactly these. The physics essentially stems from quantum mechanics, where a particle, unseen by an observer, is

expressed as a probability distribution, philosophically existing in all points in space, until measured.

It is also, this very constraint that gives a parametrization for the complex, unit 2-sphere, more formally known in quantum information literature as the Bloch sphere (figure below).

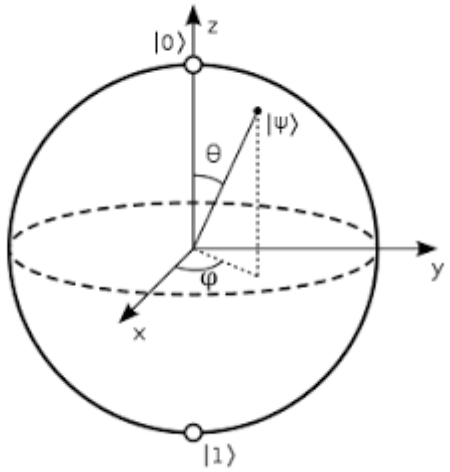

Figure 4.2: Image of the Bloch sphere taken from [1].

The Bloch sphere is a closed, convex subset of $\mathbb{C}^3$, which implies that it has some nice properties. For one, all extreme points of the Bloch sphere make up exactly the boundary of it. The physical representation of this (and further discussed below) is that all pure states are found exactly here. Any point in the interior of the sphere is what is referred to as a mixed state. With this in mind, we have motivated the following two definitions.

Definition 4.25. *A **pure quantum state** is any rank one projection. For instance,*

$|\psi\rangle$ is pure if the operator $|\psi\rangle\langle\psi|$ is rank one. Examples include states such as $|0\rangle$ and $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.

Definition 4.26. *Concordantly, a **mixed quantum state** is any state that can be written as a convex combination of extreme (pure) states.*

We also have another crafty way of characterizing pure versus mixed states. Given any state $|\psi\rangle$, we may define its corresponding **density matrix**, which is given by

$$\rho = |\psi\rangle\langle\psi|$$

or equivalently in the case of mixed states,

$$\rho = \sum_{i\in[1,n]} p_i|\psi_i\rangle\langle\psi_i|$$

We may remark that, whenever there exists an $i \in [1, n]$ such that $p_i = 1$, we have a pure state. Also note that any pure state will always yield a rank one density matrix. For our general knowledge we quickly state a the following lemma regarding density matrices.

Lemma 4.21. *Let ρ be a density matrix representing a mixed state. Then $\rho \geq 0$ and $Tr(\rho) = 1$.*

Proof. Based on our definition, we can let $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$ where $0 < p_i < 1$ for all i (without loss of generality we may ignore the equality at each of the boundaries here. The case of when $p_i = 1$ for some i follows from this result) and $\sum_i p_i = 1$. Then, since $p \geq 0$ implies that for any $j \in [1, n]$ the inner product $\langle\psi_j|\rho|\psi_j\rangle \geq 0$ we see,

$$\begin{aligned}
\langle\psi_j|\rho|\psi_j\rangle &= \langle\psi_j|\Big(\sum_{i\in[1,n]} p_i|\psi_i\rangle\langle\psi_i|\Big)|\psi_j\rangle && (4.1)\\
&= \sum_i p_i\langle\psi_j|\psi_i\rangle\langle\psi_i|\psi_j\rangle && (4.2)\\
&= \sum_i p_i\delta_{ij}\delta ji && (4.3)\\
&\geq 0 && (4.4)
\end{aligned}$$

As for the trace condition, first note that for any n dimensional vector v,

$$Tr(vv^*) = v^*v$$

and that for any $|\psi_i\rangle = \sum_i \alpha_i|i\rangle$ we get

$$Tr(|\psi_i\rangle\langle\psi_i|) = \langle\psi_i|\psi_i\rangle = \sum_i \alpha_i^2 = 1$$

Therefore,

$$\begin{aligned}
Tr(\rho) &= Tr(\sum_i p_i|\psi_i\rangle\langle\psi_i|) && (4.5)\\
&= \sum_i p_i Tr(|\psi_i\rangle\langle\psi_i|) && (4.6)\\
&= \sum_i p_i\langle\psi_i|\psi_i\rangle = \sum_i p_i = 1 && (4.7)
\end{aligned}$$

□

This also provides us with the following fact- a matrix M is a density matrix if and only if $M \geq 0$ and $Tr(M) = 1$. Density matrices are also always Hermitian. The proof is trivial.

Remark: The states described above reside in what is called a **Hilbert space**, typically denoted with $\mathcal{H}$. One may often see an additional subscript as in $\mathcal{H}_A$ or $\mathcal{H}_B$ to denote the Hilbert space associated to the famous experimentalists, Alice and Bob. One may also have the tensor product space of n separate Hilbert spaces,

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_n$$

4.1.2 Quantum error correction and the (complex) k- rank numerical range

A closely related to field to quantum information is that of quantum error correction (QEC). QEC aims to provide a way to accurately send and receive quantum bits of information, while keeping a certain level of security. By this notion alone, one can see that this area of study will necessarily be of great value in the engineering of quantum computers. In what follows, a brief synopsis of QEC is given along with a connection to the complex higher rank numerical range.

Classical error correcting methods have been around for some time and are quite well studied at this point. This is beneficial to the study of error correcting codes in QEC as we can aim to model these quantum codes after their classical analogues. However, due to the already complex nature of quantum mechanics, this is not as straight forward as it seems. Consider the following example: suppose you and your close friend are talking over the phone to one another in the midst of an unforgiving

thunderstorm. The line adjoining you both is quite choppy and unclear so to more accurately communicate information across the line, each person may repeat what they say (for instance) three times. Repeating oneself thrice gives the individual on the other end a greater chance of making sense of the information you are passing them. This motivates our example. Classically, we may, instead of simply sending a single bit of information in the state 0 or 1, we may send the information in threes, thus using,

$$0 \to 000$$

$$1 \to 111$$

The conclusion to take from this is that, if the receiver on the other line receives a state of 101, it is likely that the sending was attempting to send information in the 1 state, as 101 is close to 111, which represents the 1 state. Although this method of three-fold repetition may take longer to send and may use more memory space, it does give a nice solution to the problem of having a noisy environment present. Unfortunately, if one were to naively attempt to mimic this strategy in the quantum realm, it would fail, based on the following premise, that no quantum state $|\psi\rangle$ may be copied to $|\psi\rangle \otimes |\psi\rangle \otimes |\psi\rangle$. Its is a consequence of the following theorem.

Theorem 4.22. *(No-cloning) Let $|e\rangle$ some basis vector. Then, there exists no unitary U such that for all states $|\psi\rangle$,*

$$U(|\psi\rangle \otimes |e\rangle) = |\psi\rangle \otimes |\psi\rangle$$

In other words, one may not map $|\psi\rangle$ *to* $|\psi\rangle \otimes |\psi\rangle$.

Proof. By contradiction, suppose there does exist a unitary such that for all $|\psi\rangle$ $U(|\psi\rangle \otimes |0\rangle) = |\psi\rangle \otimes |\psi\rangle$. (As a side note, one may think of the $|0\rangle$ state as a piece of paper that we are copying the information $|\psi\rangle$ carries onto). Then, it must hold true that

$$U(|0 \otimes |0\rangle) = |0\rangle \otimes |0\rangle.$$

and

$$U(|1\rangle \otimes |1\rangle) = |1\rangle \otimes |1\rangle$$

Now, let $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ for some $\alpha, \beta \in \mathbb{C}$ such that $|\alpha|^2 + |\beta|^2 = 1$. Then

$$\begin{aligned} U(\alpha|0\rangle + \beta|1\rangle) \otimes |0\rangle &= U\alpha|0\rangle \otimes |0\rangle + U\beta|1\rangle \otimes |0\rangle & (4.8) \\ &= \alpha|0\rangle \otimes |0\rangle + \beta|1\rangle \otimes |1\rangle & (4.9) \end{aligned}$$

We may now as, whether this state is actually $|\psi\rangle \otimes |\psi\rangle$, but we see instead that $|\psi\rangle \otimes |\psi\rangle$ is nothing more than

$$\begin{aligned} (\alpha|0\rangle + \beta|1\rangle) \otimes (\alpha|0\rangle + \beta|1\rangle) &= \alpha^2|00\rangle + \alpha\beta(|01\rangle + |10\rangle) + \beta^2|11\rangle & (4.10) \\ &\neq U|\psi\rangle \otimes |0\rangle & (4.11) \end{aligned}$$

And it is herein that lies the contradiction.

Therefore, we can conclude that we may only clone orthonormal basis states. □

Now that we have developed the (albeit rudimentary) idea of what QEC is about and what it entails, we give the following definitions to illustrate how the theory of the higher rank numerical range manifests itself here.

Definition 4.27. *A **code space** is some subspace $\mathcal{C}$ of $\mathcal{H}_{2^n}$.*

Definition 4.28. *A **code word** is some element $|\psi\rangle \in \mathcal{C}$. Plainly stated, it is a quantum state that encodes some data.*

For completion, we also give the precise definition of a correctable code ([22]).

Definition 4.29. *A quantum code is said to be correctable for error operators $\mathcal{E} = \{E_a\}$ if there exists a quantum operation R on the Hilbert space $\mathcal{H}$ such that*

$$R \circ \Big(\sum_a E_a \rho E_a^* \Big) = \rho$$

for all $\rho = P_{\mathcal{C}} \rho P_{\mathcal{C}}$ where $P_{\mathcal{C}}$ denotes the projection onto the code space.

With these in mind, we now state what is referred to as the Knill-Laflamme conditions. See [22] for a more detailed exposition.

Theorem 4.23. *(Knill-Laflamme) Let $\mathcal{C} \subseteq \mathcal{H}_{2^n}$ and $P_{\mathcal{C}}$ some orthogonal projector that projects onto the space $\mathcal{C}$. Then a set of operators*

$$\mathcal{E} = \{E_1, E_2, \ldots, E_n\}$$

is correctable on $\mathcal{C}$ if and only if scalars λ_{ij} exist that satisfy

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

If we look closely at the expression above, we notice that it is identical to the definition of the (complex) higher rank numerical range. The problem of determining whether there exists a quantum error correctable code is equivalent to asking whether the joint k rank numerical range (assuming that $\dim(\mathcal{C}) = k < n$) of the error operators is non-empty. It was this particular realization that led to a new motivation to study both QEC and the higher rank numerical ranges from a purely mathematical point of view.

It might be worth asking whether the real higher rank numerical range has any applications in QEC, or related streams. An explicit connection is yet to be found in this context.

4.1.3 A synopsis of entanglement

Quantum entanglement is an inherent trait to the universe that has, and continues to, elude mathematicians and physicists. Originally appearing in the equations brought forth by Einstein, and several other quantum pioneers, entanglement is the phenomena where linking two (or more) particles appears to result in an instantaneous exchange of information. This is something that goes against the logic of there being a finite upper bound on the rate at which communication can occur, which is what Einstein and physicists got tangled up in. Measuring or altering one part of the entangled system seems to instantly affect other parts of the system. This still, and very well could, remain a mystery as to how nature can demonstrate this.

Nonetheless, it remains a fascinating area of study.

Mathematically, entanglement can be modelled in the following way; suppose that a Hilbert space $\mathcal{H}$ is the tensor product of two Hilbert spaces, call these $\mathcal{H}_A$ and $\mathcal{H}_B$ for which we can write $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. A pure state $|\phi\rangle$ is unentangled (or equivalently; separable, factorable, etc.) if there exists $|\phi_A\rangle \in \mathcal{H}_A$ and $|\phi_B\rangle \in \mathcal{H}_B$ such that $|\phi\rangle = |\phi_A\rangle \otimes |\phi_B\rangle$. If it cannot be written as a single tensor product, then the state contains some measure of entanglement, and hence is not separable. Similarly regarding mixed states, a density matrix is separable if it can be written as a summation of tensor products, i.e. $\rho = \sum_i p_i u_i \otimes v_i$ for some matrices u_i and v_i within the subsystems of $\mathcal{H}$.

Before moving forward, we touch on the idea of the different forms of entanglement in multipartite states. One interesting feature of quantum mechanics is that as one increases the number of particles in a given system, there are different types of entanglement that naturally arise. For example, states in three particles could be entangled in one of two inequivalent ways (these being GHZ- type entanglement and W entanglement). If we let $\mathcal{H}$ denote the space of all states in n particles then the number of distinct forms of entanglement effectively divides $\mathcal{H}$ into equivalence classes, based on the type of entanglement a given state carries. In two particles, it is well known that the Bell state given by

$$|\psi\rangle_0 = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

is maximally entangled. In fact, any state of the form

$$|\psi\rangle = (I \otimes U)|\psi\rangle_0$$

for some unitary U is also maximally entangled. Entanglement in the bipartite case is well known, however, in this thesis, we look at entanglement in three particles and higher.

Conceptually, entanglement is not too difficult a concept to understand. We can intuitively think of it in the following way- suppose that we look at two independent particle systems that have no relationship to one another. What an observer does or measures on one of the systems, has no effect on the other. Hence, we say that the two are completely unentangled. This makes sense mathematically as completely unentangled states can be represented as a single tensor product. Hence, when acting some self-adjoint operator on this state, only one of the subsystems will ever be effected at a given instance. Similarly, a completely entangled state will behave in the opposite way, namely, given two particle systems that are linked together, observing one of the subsystems will result in a change in the other.

Now that we have built on some preliminaries regarding entanglement theory, we now introduce some of the underlying mathematics and research that goes into the area.

One of the more important questions in entanglement theory is finding an efficient method or way of defining the degree in which an arbitrary state is entangled. The problem is difficult to answer when speaking of particle systems involving more than

two constituents, however, when exactly two particles are involved (referred to as the bipartite state) much is already known. We will start here, introducing an example of an entanglement measure. What follows in the next section will be a new measure of entanglement for any number of particles, derived only for pure symmetric quantum states.

Suppose we are given a bipartite state for which we will label $|\psi\rangle$. We can assess the degree that $|\psi\rangle$ is entangled by looking at, what are called the Schmidt coefficients for $|\psi\rangle$. Before defining the measure to be used, we first define some key terms and a theorem.

Theorem 4.24. *(**Schmidt decomposition**) Let $|\psi\rangle$ denote a quantum state in the bipartite system $\mathcal{H}_A \otimes \mathcal{H}_B$. Then there exists orthonormal bases $|a\rangle$ and $|b\rangle$ for $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively, and scalars $\lambda_{(a,b)}$ such that the state can be written in the form*

$$|\psi\rangle = \sum_{(a,b)} \lambda_{(a,b)} |a\rangle \otimes |b\rangle.$$

Proof. We start by letting $|e_i\rangle_{H_1}$ and $|f_j\rangle_{H_2}$ be orthonormal bases for the corresponding Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. Then $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$, a quantum state, can be expressed in the form $|\psi\rangle = \sum_{ij} A_{ij} |e_i\rangle_{H_1} \otimes |f_i\rangle_{H_2}$, where (A_{ij}) are the elements of a matrix A. By the singular value decomposition, the matrix of A has the equivalent form UTV^*. Here, T is a positive semi-definite, diagonal matrix with entries ordered in decreasing fashion, with U and V, unitary matrices. Then,

$$(A_{ij}) = (U_{ik})(T_{kk})(V_{kj})$$

and thus our expression for the state may be seen as

$$|\psi\rangle = \sum_{ijk}(U_{ik})(T_{kk})(V_{kj})|e_i\rangle_{H_1} \otimes |f_i\rangle_{H_2} = \Big(\sum_i U_{ik}|e_i\rangle_{H_1}\Big)\Big(\sum_j V_{kj}|f_j\rangle_{H_2}\Big)\Big(\sum_k T_{kk}\Big)$$

Now (simply relabelling) let $|a\rangle = \sum_i U_{ik}|e_i\rangle_{H_1}$ and $|b\rangle = \sum_j V_{kj}|f_j\rangle_{H_2}$. Each of these are orthonormal bases, inherited from the fact that U and V were unitary to begin with (of which the columns and rows are orthonormal). Thus,

$$|\psi\rangle = \sum_k T_{kk}|a\rangle \otimes |b\rangle.$$

If we let $T_{kk} = \lambda_{(a,b)}$ we attain our result. □

Definition 4.30. *The **Schmidt coefficients** of a matrix $A \in \mathcal{M}_n(\mathbb{C})$ are the square roots of the eigenvalues of the matrix product A^*A, or equivalently, they are the singular values of A. (The term Schmidt coefficients and singular values are used interchangeably when clear about their context).*

Definition 4.31. *The **Schmidt number** of a matrix $A \in \mathcal{M}_n(\mathbb{C})$ is the number of non-zero singular values of A.*

With this in mind, we may now define an entanglement measure- Given some bipartite state $|\psi\rangle$, we can find the Schmidt coefficients of the matrix formed by taking the outer product of the state with itself (in other words, its corresponding density matrix). If, when these singular values are found, they are the same (i.e. $\lambda_1 = \lambda_2 = \frac{1}{2}$) then $|\psi\rangle$ is completely entangled. If there is a maximum separation between the singular values of A, (when $\lambda_1 = 1$ and $\lambda_2 = 0$) then the state is completely unentangled.

Conclusively, if both Schmidt coefficients are non-zero and non-identical, there is some degree of entanglement present, however, not completely entangled.

For two example calculations, we have the following,

Example 4.7. *Suppose we have the state* $|\psi\rangle = \frac{1}{\sqrt{2}} = (|00\rangle + |11\rangle)$. *Then,*

1- Define $A = \frac{1}{\sqrt{2}}|0\rangle\langle 0| + \frac{1}{\sqrt{2}}|1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

2- Compute A^*A

$$A^*A = (\frac{1}{\sqrt{2}})^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^2 = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

3- Find eigenvalues of A^*A

$$e_1 = \lambda_1^2 = \frac{1}{2}$$
$$e_2 = \lambda_2^2 = \frac{1}{2}$$

Conclusion: If you find that each of the singular values are the same, then the state you started with was maximally entangled. If there is maximum distance between singular values (i.e. they are 0 and 1) then the state is completely unentangled.

Example 4.8. *Suppose we are given the state* $|\psi\rangle = \frac{1}{\sqrt{3}} = (|01\rangle + |10\rangle + |11\rangle)$. *Then*

our matrix A *is* $\frac{1}{\sqrt{3}}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ *and hence* $A^*A = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. *This yields eigenvalues of* ≈ 0.873 *and* ≈ 0.127*, and hence there is some degree of entanglement present in the state.*

There are many other entanglement measures known. Other well known measures include the geometric measure of entanglement, and the use of completely positive maps (entanglement witness- see [20] for more on this). We now shift our focus to the entanglement measure we outline in this thesis, namely the surface area measure.

4.2 Introducing the new measure

Using the Majorana representation, we define a unique polytope representation for a given pure symmetric quantum state (later referred to as quantum state polytopes). Using this, we define a new measure of entanglement that compares the surface areas of these geometric objects with those given from states that are known to be maximally entangled. From a purely mathematical standpoint, we also look into the problem of placing n points on the unit sphere in E^3 so as to maximize the surface area of the inscribed polytope. On one final note, we make a remark on the partial trace operator, and show how it acts geometrically in this setting; namely how it changes the structure of these quantum state polytopes and how it influences the

degree of entanglement present in a given system.

There is some work to do before defining the measure itself. To begin, we introduce a number of definitions which will be helpful.

Definition 4.32. *(Pure symmetric quantum state) A state $|\psi\rangle$ is said to be symmetric if it is permutation invariant. I.e. it is a state of the form*

$$|\psi\rangle = N \sum_{\sigma \in S_n} |\psi_{\sigma(1)} \psi_{\sigma(2)} \dots \psi_{\sigma(N)}\rangle$$

where N is some normalization constant and S_n the symmetric group on n elements.

Remark: States that have permutation symmetry are restricted to exist in an $n+1$ dimensional subspace of the underlying Hilbert space (of dimension 2^n). This space is spanned by the Dicke states, which are also the eigenstates of the squared angular momentum operator, J. We will use Ω_n to denote the space of pure symmetric states of n particles. For more to do with symmetric states [26] is a very nice source.

As we will see later on, there is one type of symmetric state that is unique in its own right. That is the GHZ state. We choose to formally define it here.

Definition 4.33. *(GHZ state) We take this state to be*

$$|\psi\rangle = \frac{1}{2}(|0\rangle^{\otimes n} \pm |1\rangle^{\otimes n})$$

Remark: Every pure symmetric state has a unique geometrical representation inscribed on the surface of the unit 2- sphere. This is the Majorana representation, which we will describe in depth shortly.

Definition 4.34. *Let $|\psi\rangle \in \Omega_n$ and also let $z_1, \ldots, z_n$ denote the solutions to p_ψ. Then $|\psi\rangle$ has a geometrical interpretation on the Majorana sphere which is given by*

$$conv(z_1, \ldots, z_n)$$

I.e. the convex hull of the projected solutions onto the Majorana sphere. We shall name this object the ***quantum state polytope*** *(QSP) representation of $|\psi\rangle$. This is equivalently the Majorana representation of pure symmetric states.*

Here we give a brief remark regarding the two different classes a QSP can fall into.

Remark: A pure quantum state $|\psi\rangle$ will admit a polygonal representation if and only if all roots of p_ψ align in the same plane (this plane need not be parallel with the equator of the Majorana sphere, it could be on some oblique cut through it). Similarly, a pure state will admit a polyhedral representation if the number of distinct solutions to the Majorana polynomial is at least 4 such that they do not all lie in the same plane (i.e. they do not fall into the class described previously).

The general construction of QSPs can now be given by the following sequence of steps (this is precisely the Majorana representation),

1- Given an arbitrary state , $|\psi\rangle$, find its **Majorana polynomial**, denoted $p_\psi(z)$. (An explicit derivation of this can be found in [12]).

2- Project the solution to $p_\psi(z)$ onto S^2 by the transformation

$$x_1 = \frac{2Re(z)}{|z|^2+1}$$
$$x_2 = \frac{2Im(z)}{|z|^2+1}$$

and

$$x_3 = \frac{|z|^2-1}{|z|^2+1}$$

3- Obtain the polytope, (of at most n vertices, denoted Π_n) traced out by the collection of solutions to $p_\psi(z)$.

We provide an example calculation next. [2] is a very nice text that outlines the geometry of entanglement well. More analysis on the Majorana representation is given in this source.

Example 4.9. *Starting with the state-*

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes n} + |1\rangle^{\otimes n})$$

Obtain the Majorana polynomial by $f : \Omega_n \to \mathbb{C}[z]$*, which will give*

$$p_\psi(z) = \frac{1}{\sqrt{2}}(1 - z^n)$$

This produces the n*,* n^{th} *roots of unity as solutions,*

$$z_k = e^{\frac{2\pi k i}{n}} = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}$$

The vertices of Π_n *are then located at*

$$\big(\cos\frac{2\pi k}{n}, \sin\frac{2\pi k}{n}, 0\big)$$

This example will turn out to be important to us in subsequent sections as it gives an easy to understand, and manipulate polygon (which happens to always be regular). This provides a nice segue into the next section.

Lemma 4.25. *The QSP of a state $|\psi\rangle$ is unique and vice versa.*

Proof. This follows from the fact that the mapping which takes a point $z \in \mathbb{C}$ to the Riemann sphere is a bijection. □

4.2.1 Regular polygons and special case polyhedra

An arbitrary state will produce some arbitrary arrangement of points on S^2 that can trace out an infinite number of these shapes. However, some of these shapes are of special interest to us because they are easier to analyse mathematically and also easier to conceptualize logically. In the case of quantum state polygons, the 'easier to analyse cases' arise when the polygon is regular. And for polyhedra, the simple cases include; the five platonic solids, double pyramids and ones that are composed entirely of triangular faces.

We now derive a number of geometrical results which will later apply to the measure.

Theorem 4.26. *The area of a regular polygon with n vertices is given by*

$$A = \frac{1}{4} n L^2 \cot \frac{\pi}{n} \tag{4.12}$$

where L is the side length of $\{n\}$.

Proof. In any regular polygon, the distance from its centre to the midpoint of any edge is the same. Call this distance d. Let R_0 denote the distance from the centre of $\{n\}$ to any vertex. We can then trace out a right angle triangle for which we can find the area of.

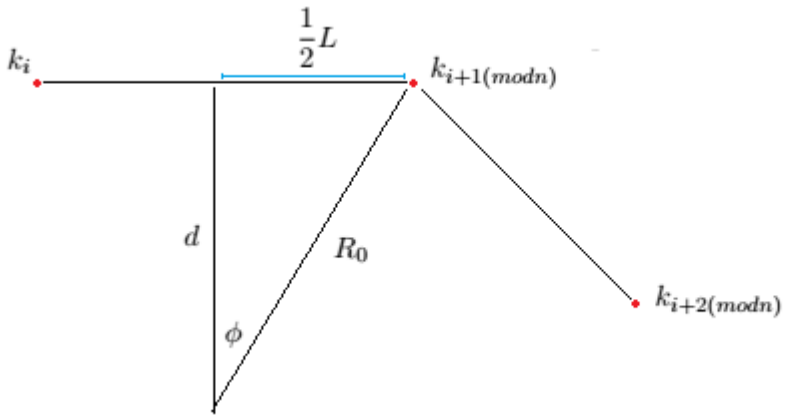

Figure 4.3: A regular polygon decomposed into n triangles.

Its not difficult to see that $\phi = \pi/n$. Because we are partitioning the polygon into $2n$ triangles, there are $2n$ copies of ϕ that sum to 2π, therefore $2n\phi = 2\pi$ simplifies to $\phi = \pi/n$.

The area of each triangle is given by $\frac{1}{2}bh$, where the base b is simply $\frac{1}{2}L$ and height h given by the trigonometric relation $h = \frac{1}{2}L\cot\frac{\pi}{n}$. Therefore, the area of one such triangle is then $A_\triangle = \frac{1}{8}L^2\cot\frac{\pi}{n}$. Since there are $2n$ copies of these triangles in $\{n\}$, we see that the total area of $\{n\}$ is $A = 2nA_\triangle = \frac{1}{4}nL^2\cot\frac{\pi}{n}$. □

Lemma 4.27. *Given a regular polygon of n vertices inscribed in the great circle of*

the Riemann sphere, the distance between any two adjacent roots (i.e. its edge length) is given by

$$L = |dist(z_i, z_{i+1}) = 2|\sin\frac{\pi}{n}| \tag{4.13}$$

Proof. Define two roots $z_{k_1} = e^{\frac{2\pi k_1}{n}i} = \cos(\frac{2\pi k_1}{n}) + i\sin(\frac{2\pi k_1}{n})$ and $z_{k_1} = e^{\frac{2\pi k_2}{n}i} = \cos(\frac{2\pi k_2}{n}) + i\sin(\frac{2\pi k_2}{n})$. Since $|z_{k_1}| = |z_{k_2}| = 1$ we are on the unit circle of S. Without loss of generality we can take $k_2 = 1 + k_1 \,(\mathrm{mod}\, n)$ so that z_{k_1} and z_{k_2} are adjacent.

Now, to find the (straight) Euclidean distance between these roots, we can use the standard distance function without our inputs being the x and y components of our roots. Thus,

$$\begin{aligned} \mathrm{dist}(x, y) &= \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} && (4.14)\\ \mathrm{dist}(z_{k_1}, z_{k_2}) &= \sqrt{(\cos(\frac{2\pi k_1}{n}) - \cos(\frac{2\pi k_2}{n}))^2 - (\sin(\frac{2\pi k_1}{n}) - \sin(\frac{2\pi k_2}{n}))^2} && (4.15) \end{aligned}$$

For the purposes of staying organized, let us first simplify the $x-$component, namely, $(x_1 - x_2)^2$

$$\begin{aligned} (x_1 - x_2)^2 &= (\cos(\frac{2\pi k_1}{n}) - \cos(\frac{2\pi k_2}{n}))^2 && (4.16)\\ &= (\cos(\frac{2\pi k_1}{n}) - \cos(\frac{2\pi(1 + k_1)}{n}))^2 && (4.17)\\ &= (\cos\frac{2\pi k_1}{n} - \cos\frac{2\pi}{n}\cos\frac{2\pi k_1}{n} + \sin\frac{2\pi}{n}\sin\frac{2\pi k_1}{n})^2 && (4.18)\\ &= \cos^2\frac{2\pi k_1}{n}(1 - \cos\frac{2\pi}{n})^2 + \sin^2\frac{2\pi k_1}{n}\sin^2\frac{2\pi}{n} && (4.19)\\ &+ 2\sin\frac{2\pi k_1}{n}\sin\frac{2\pi}{n}\cos\frac{2\pi k_1}{n}(1 - \cos\frac{2\pi}{n}) && (4.20) \end{aligned}$$

Similarly for the $y-$component, we have

$$
\begin{aligned}
(y_1 - y_2)^2 &= (\sin(\frac{2\pi k_1}{n}) - \sin(\frac{2\pi k_2}{n}))^2 && (4.21)\\
&= (\sin(\frac{2\pi k_1}{n}) - \sin(\frac{2\pi(1+k_1)}{n}))^2 && (4.22)\\
&= (\sin\frac{2\pi k_1}{n} - \sin\frac{2\pi}{n}\cos\frac{2\pi k_1}{n} + \cos\frac{2\pi}{n}\sin\frac{2\pi k_1}{n})^2 && (4.23)\\
&= \sin^2\frac{2\pi k_1}{n}(1-\cos\frac{2\pi}{n})^2 + \cos^2\frac{2\pi k_1}{n}\sin^2\frac{2\pi}{n} && (4.24)\\
&- 2\sin\frac{2\pi k_1}{n}\sin\frac{2\pi}{n}\cos\frac{2\pi k_1}{n}(1-\cos\frac{2\pi}{n}) && (4.25)
\end{aligned}
$$

Finally, calculating what is under the root of (4.14), we get

$$
\begin{aligned}
(x_1 - x_2)^2 + (y_1 - y_2)^2 &= \cos^2\frac{2\pi k_1}{n}(1-\cos\frac{2\pi}{n})^2 + \sin^2\frac{2\pi k_1}{n}\sin^2\frac{2\pi}{n} && (4.26)\\
&+ \sin^2\frac{2\pi k_1}{n}(1-\cos\frac{2\pi}{n})^2 + \cos^2\frac{2\pi k_1}{n}\sin^2\frac{2\pi}{n} && (4.27)
\end{aligned}
$$

with the cross terms in the expansions of each component cancelling out. Simplifying the above expression gives way to the following

$$
\begin{aligned}
&= (1-\cos\frac{2\pi}{n})^2(\sin^2\frac{2\pi k_1}{n} + \cos^2\frac{2\pi k_1}{n}) + \sin^2\frac{2\pi}{n}(\sin^2\frac{2\pi k_1}{n} + cos^2\frac{2\pi k_1}{n}) && (4.28)\\
&= 1 - 2\cos\frac{2\pi}{n} + \cos^2\frac{2\pi}{n} + \sin^2\frac{2\pi}{n} = 2 - 2\cos\frac{2\pi}{n} && (4.29)\\
&= 2(1-\cos\frac{2\pi}{n}) = 4\sin^2\frac{\pi}{n} && (4.30)
\end{aligned}
$$

The result now becomes apparent that

$$\begin{aligned} \text{dist}(z_{k_1}, z_{k_2}) &= \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} && (4.31) \\ &= \sqrt{4\sin^2\frac{\pi}{n}} = 2|\sin\frac{\pi}{n}| && (4.32) \end{aligned}$$

Our result should make sense intuitively. Because we choose our roots arbitrarily, there should be no dependence on k_1 in our final expression, only the number of roots in the set from p_ψ (or vertices of the polygon). □

Lemma 4.28. *Suppose that $|z_i| = k, \forall i$ such that $k \neq 1$. Then the distance between any two adjacent roots is*

$$l = dist(z_i, z_{i+1}) = \sqrt{\frac{8}{1+k^2}}\sin\frac{\pi}{n} \tag{4.33}$$

Proof. Now that for every root $|z_i| \neq 1$ (but that still remain on the same horizontal cross section of S^2) we define a distance d_0 above the unit circle. This is effectively the x_3 component described in the stereographic projection mapping, so we can assert $d_0 = (|z|^2 - 1)/(|z|^2 + 1)$. Define r_0 to be the radius of the circle C_0 that all roots z_i lie along. Based on the depiction in figure 4.3, there exists a right triangle that forms among the radii of S^2 and C_0 and distance d_0. Thus, we have the relationship $r_0^2 = 1 - d_0^2$ or $r_0 = \sqrt{1 - \frac{|z|^2-1}{|z|^2+1}} = \sqrt{\frac{2}{1+|z|^2}}$.

Similar to the proof of theorem 4.28, we now place the n roots along C_0, equidistant from one another as to form a regular polygon such that it can be partitioned

into $2n$ right triangles as seen in figure 4.3.

Elementary relations show that the distance between each root, l is given by $l = 2r_0 \sin\frac{\pi}{n}$ which when given the expression for r_0 we can simplify to (4.33).

This is simply an extension of lemma 4.29. Equation (4.33) gives the distance between any two adjacent roots for any polygon cutting the sphere perfectly horizontal. If we take $k = 1$ in (4.33) then we get back (4.13). □

Theorem 4.29. *Let $|\psi\rangle$ be a pure quantum state with Majorana polynomial p_ψ. Suppose that all roots to p_ψ have the same modulus, i.e. $|z_i| = k$ for some real number k. Then, $|\psi\rangle$ admits a polygonal representation with area given by*

$$A = \frac{n}{1+k^2} sin\frac{2\pi}{n} \tag{4.34}$$

Proof. By theorem 4.28, the area of a regular polygon is $1/4\, nL^2 \cot\frac{\pi}{n}$. By lemma 4.30, the edge length of this polygon is given by (4.33), so by substituting this expression in for L in (4.36), we obtain (4.34).

□

Remark: The maximum of (4.34) occurs when $k = 1$ (i.e. $|z_i| = 1, \forall i$) and is given by

$$A = \frac{1}{2} n\, sin\frac{2\pi}{n} \tag{4.35}$$

This area refers to the GHZ state in $n-$particles as seen in example 4.9.

Proof. In the proof of lemma 4.30, we defined a radius r_0 that we used to compute a scaling factor (the factor responsible for adjusting the area as the modulus of the roots increase). Because r_0 and $|z|$ are inversely proportional to one another, the greater $|z|$ becomes, the smaller r_0 gets. However, since the area of the regular polygons described in this setting have a direct proportionality to r_0, we deduce that the area of any given polygon increases, with a decrease in $|z|$. Now it is simply a question of how small we can make $|z|$ such that A is maximal. The answer is 1.

This make sense conceptually as the great circle of S^2 contains the largest area to build a regular polygon on, as any other horizontal circle of S^2 will have a smaller radius and thus smaller region to inscribe any shapes in. □

Example 4.10. *Consider the state* $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes n} + |1\rangle^{\otimes n})$ *for* $n > 1$. *The Majorana polynomial of* $|\psi\rangle$ *is* $p(z) = \frac{1}{\sqrt{2}}(z^n - 1)$ *which gives the* n, n^{th} *roots of unity, that when stereo-graphically projected onto the Riemann sphere trace out a regular polygon with* n *vertices. Thus, its area is exactly defined by equation* (4.35).

Remark: Suppose a number z in the complex plane has a projection $z_0 = (x_1, x_2, x_3)$ on the unit 2-sphere. Then z_0 is given by;

$$x_1 = \frac{2Re(z)}{|z|^2 + 1}, x_2 = \frac{2Im(z)}{|z|^2 + 1}, x_3 = \frac{|z|^2 - 1}{|z|^2 + 1}$$

Proof. Let $z = u + iv$ for $u, v \in \mathbb{R}$. Parametrize the line intersecting $(0, 0, 1)$ and $(u, v, 0)$ with; $x_1 = tu, x_2 = tv$ and $x_3 = 1 - t$ for $-\infty < t < \infty$. Then since

$x_1^2 + x_2^2 + x_3^2 = 1$, we can write

$$(tu)^2 + (tv)^2 + (1-t)^2 = 1$$

$$t(t(u^2 + v^2 + 1) - 2) = 0$$

If we take t to be non-zero then we obtain $t = \frac{2}{u^2+v^2+1} = \frac{2}{|z|^2+1}$ and substituting this back into the original parametrizations yields the desired coordinate expressions.

□

4.2.2 The surface area entanglement measure

Given our previous discussions, we are now able to define a measure of entanglement for which we call the surface area quantum entanglement measure. We now define it here.

Definition 4.35. *(Triangle area measure of entanglement) Let $\{z_i\}_{i\in n}$ be the collection of points on the Majorana sphere. Then the surface area mapping $s : \{z_i\}_{i\in n} \to \mathbb{R}_{\geq 0}$ is given by*

$$s(|\psi\rangle) = \frac{1}{6}\sum_{i,j,k}(z_j - z_i) \times (z_k - z_i) \tag{4.36}$$

The idea behind this measure is to compute the area of each triple of points on the Majorana sphere and sum them together. Although this is an overestimation of the true surface area of the QSP for the original state, it still gives information regarding the degree to which the state is entangled. The greater this area calculation is, the

more spread out the collection of points are on the Majorana sphere, and hence the more entangled the original state is. We may also note that given n distinct points on the Majorana sphere, (4.36) will have $\binom{n}{3}$ terms, as this is the number of distinct triangles in the QSP.

There are two key components to this surface area calculation. The first is the process of mapping any pure symmetric state to its quantum state polytope representation. And the second is finding the surface area of this QSP. We discussed previously the general three step procedure to find any QSP. We elaborate on this more here as it is relevant to this measure.

Suppose we start with any pure symmetric state in two particles. Taking the standard basis we can express this as

$$|\psi\rangle = (a|0\rangle + b|1\rangle) \otimes (c|0\rangle + d|1\rangle) + (c|0\rangle + d|1\rangle) \otimes (a|0\rangle + b|1\rangle)$$

which when expanded and simplified is just

$$2ac|00\rangle + (ad + bc)(|01\rangle + |10\rangle) + 2bd|11\rangle$$

Now, recalling that the first step in the procedure is to find the Majorana polynomial for $|\psi\rangle$; we can simply take this state to be in the form of the quadratic $2acx^2 + (ad + bc)x + 2bd := p_\psi(x)$.

Remark: Note that if we take $a = c = (\frac{1}{2})^{(3/4)}$ and $b = d = -(\frac{1}{2})^{(3/4)}$ we obtain the GHZ state in 2 particles.

This gives us a way (not the most effective however) of obtaining p_ψ from $|\psi\rangle$. We can generalize this expression by noting that any pure symmetric quantum state can be decomposed in the form

$$|\psi\rangle = n!\Big(\prod_{i=1}^{n} a_{2i-1}|0\rangle^{\otimes n} + \prod_{i=1}^{n} a_{2i}|1\rangle^{\otimes n}\Big) + N_1 \sum \prod_i a_i |W\rangle + N_2 \sum \prod_i a_i |D\rangle$$

where $|W\rangle$ is the W state in n particles (for example in three particles we have that $|W\rangle = (1/\sqrt{3})(|100\rangle + |010\rangle + |001\rangle)$) and where $|D\rangle$ are the Dicke states in n particles (see [18] for more information). Furthermore, $N1$ and N_2 are simply complex numbers composed of the original coefficients (the a_i terms) in $|\psi\rangle$. Conceivably, one could obtain the Majorana polynomial from this expression, however, it is going to be computationally expensive for larger values of n. So we point the reader to [12] once again for alternate methods.

However, we will in this instance, use this expression in three particles to find the Majorana polynomial of the state,

$$|\psi\rangle = \frac{1}{\sqrt{5}}\Big(|000\rangle + (|110\rangle + |011\rangle + |101\rangle) + |111\rangle\Big) \tag{4.37}$$

and use apply our surface area calculation to observe the degree of entanglement inherent in this system. We can now state the Majorana polynomial by associating the $|000\rangle$ tensor product as the x^3 term, $|110\rangle + |011\rangle + |101\rangle$ together as the x term

and finally the $|111\rangle$ as the constant term. We now obtain

$$p_\psi(x) = \frac{1}{\sqrt{5}}(x^3 + x + 1)$$

This expression has roots of $x_1 \approx -0.68$, $x_2 \approx 0.34 - 1.16i$, $x_3 \approx 0.34 + 1.16i$ which when stereographically projected on the Majorana sphere we get the three coordinates

$$(0.93, 0, -0.37)$$

$$(0.28, -0.94, 0.19)$$

and

$$(0.28, 0.94, 0.19)$$

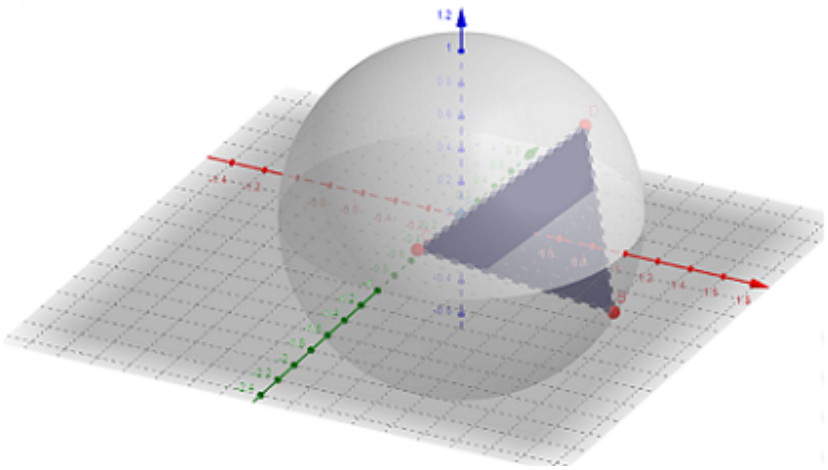

Figure 4.4: The inscribed QSP for $|\psi\rangle$

This now outlines the QSP of $|\psi\rangle$ and we aim to compute its surface area. This can be done via computing

$$\frac{|AB \times AC|}{2}$$

where $AB = (-0.65, -0.94, 0.56)$ and $AC = (-0.65, 0.94, 0.56)$. Then,

$$= \det \begin{bmatrix} e_x & e_y & e_z \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{bmatrix} \tag{4.38}$$

$$= \begin{bmatrix} e_x & e_y & e_z \\ -0.65 & -0.94 & 0.56 \\ -0.65 & 0.94 & 0.56 \end{bmatrix} \tag{4.39}$$

$$\approx 1.0488e_x + 1.22e_z \tag{4.40}$$

Then

$$\frac{1}{2}|AB \times AC| = \frac{1}{2}\sqrt{(1.0488)^2 + (1.22)^2} \approx 0.805$$

And we are done our surface area calculation. However, there is a way to compare the entanglement present in this state against the entanglement in that of another state.

We know that in three particles, the GHZ state is maximally entangled with regards to the GHZ type of entanglement) and gives rise to the maximal surface area polytope, which in this case is simply the equilateral triangle. This will have area,

$$A(3) = \frac{1}{2}(3)\sin\frac{2\pi}{3} = \frac{3\sqrt{3}}{4}$$

And thus, we may define a relative measure of entanglement,

$$s(\psi) = \frac{SA(\Pi_n)}{SA(\Pi_n^{max})} = \frac{0.805}{\frac{3\sqrt{3}}{4}} \approx \frac{0.805}{1.299} \approx 0.62$$

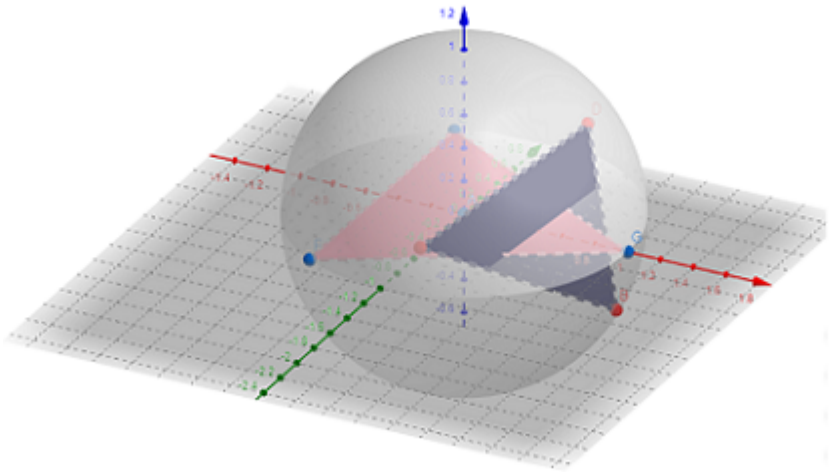

Figure 4.5: The inscribed QSP for the GHZ state in 3 particles as depicted by the red triangle.

Hence we know the original state in 4.37, although not completely unentangled or completely entangled, there is some entanglement present in the system. See figure 4.5 for a visualization of these QSPs.

Conjecture 4.1. *All maximally entangled states* $|\psi\rangle \in \Omega_n$ *yield maximal surface area polytopes.*

Remark: When considering states that follow GHZ entanglement, conjecture 4.2 is certainly true.

Finally, we end with what we call the surface area entanglement measure. It is defined as follows:

Definition 4.36. *(Surface area measure of entanglement) Define the map* $s : \Omega_n \to [0, 1]$ *by*

$$s(|\psi\rangle) = \frac{SA(\Pi_n)}{4\pi}$$

where Π_n *is the QSP of* $|\psi\rangle$*.*

The formulation of this measure is almost identical to that of the relative measure, except instead of dividing out by the maximal surface area QSP, we simply divide by the surface area of the unit sphere. This will of course yield an output between 0 and 1 since 4π is an upperbound for the surface area of any QSP inscribed in $\mathcal{S}^2$.

4.2.3 A possible extension of the measure

If conjecture 4.2 is true, then one could define an analogous measure of entanglement that compares the volumes of quantum states emitting polyhedrons to the same states having maximal surface area, as we know they are the same maximally entangled state previously used. The same interpretation would follow from the definition of the relative surface area measure.

Definition 4.37. *(Volume entanglement measure) Define the map* $v : \Omega_n \to [0, 1]$ *by the following,*

$$v(\psi) = \frac{vol(\Pi_n)}{vol(\Pi_n^{max})}$$

where Π_n *is the QSP of* $|\psi\rangle$ *and* Π_n^{max} *is the QSP of the maximally entangled state in* n *particles.*

Remark: $v(\psi) = 1$ implies that ψ is maximally entangled and similarly $v(\psi) = 0$ implies that ψ is completely separable.

Conjecture 4.2. *All maximal volume polyhedra also have maximal surface area.*

Open Problem: What are the maximal volume polytopes that can be inscribed in the unit 2- sphere?

These two problems will provide a nice transition into the next section where these statements are explored in more detail. It is important to investigate these problems as our measure of entanglement is contingent on what results are known of an optimal distribution of points on the 2-sphere. For the remainder of this chapter, the open problem stated above will be referred to as the n point problem. (Although this does not distinguish between the optimal configuration for volume and surface, where it is ambiguous, we will specify which of these we mean).

4.2.4 Known results for the n point problem

This section of chapter four is devoted to the current progress of the n point problem looking specifically to maximize the volume of the polyhedron Π formed by this collection. We review the results of two main papers, [3] and [29].

Current status: The problem has been solved (using algebraic methods) for up to and including 8 points. Using computer assistance, the maximal volume has been found for $n \leq 30$.

We will start by giving an overview of the techniques Bermann and Hanes used in [3]. The authors first start by defining a necessary condition for a polyhedron inscribed in the 2- sphere to be maximal. This can be stated as a definition.

Definition 4.38. *Let* Π *be a polytope inscribed in the 2- sphere and label its vertices* k_1 *through* k_n*. Then* Π *is said to have maximum volume if for each vertex* k_j *there exists an open set* $U_j \subset \partial\mathcal{S}^2$ *such that for any point* $k'_j \neq k_j$ *in* U_j *we always have*

$$vol(k_1, \ldots, k_n) \geq vol(k_1, \ldots, k_{j-1}, k'_j, k_{j+1}, \ldots, k_n).$$

Throughout the paper, the authors refer to this as 'property Z'. And it is this very technique that they instil to obtain their results for up to eight vertices.

Given n points on the unit sphere, the volume of the convex hull of these points is given as

$$\frac{1}{2}\sum_{i=1}^{n} ||(v_i - v_0) \times (v_{i+1(\mathrm{mod}n)} - v_0)||$$

which is the expression the authors in [3] maximize.

For smaller values of n, this expression is not too difficult to analyse, however for

larger n it can become quick complicated. It is likely a consequence of this that the same method will not work for vertices nine and greater. It would seems that a new method needs to be realized before any further algebraic proofs for a greater number of vertices is to be shown.

Hanes and Bermann (1970) showed that

$$V(n_1, \ldots, n_8) \leq \frac{\sqrt{3}}{4}\big(1 - \frac{1}{3}tan^2\frac{2\pi - \theta}{6}\big) + \frac{9}{4}tan\frac{14\pi + \theta}{54}\big(1 - \frac{1}{3}tan^2\frac{14\pi + \theta}{54}\big)$$

which has a maximum at

$$\sqrt{\frac{475 + 29\sqrt{145}}{250}} \approx 1.8157$$

On a final note regarding this paper. Although it is not completely relevant to the issue of finding maximal value polyhedra, Bermann and Hanes give the following result, stated as a lemma. (In what follows, they use the term *double n pyramid* to denote the polyhedra with n vertices, 2 of which are located at antipodal points, stemming from the centre of a regular triangle of $n - 2$ vertices. The octahedron is an example of a double pyramid).

Lemma 4.30. *If* Π *is a double pyramid inscribed in the* 2*- sphere and has maximal volume then its volume is given by*

$$\frac{n-2}{3}\sin\Big(\frac{2\pi}{n-2}\Big)$$

Remark: One can easily show that

$$\lim_{n\to\infty}\frac{n-2}{3}\sin\Big(\frac{2\pi}{n-2}\Big) = \frac{2\pi}{3}$$

This should be true intuitively because a volume of $\frac{2\pi}{3}$ is simply twice that of a cone with $h = r = 1$, which as the number of vertices slowly grows off to infinity, one obtains exactly the double cone.

Without going over-board on detail, these were the main points of the paper. We now briefly discuss the main results of [29].

This paper essentially provides a computer-aided proof of all maximal volume polyhedra that can be inscribed in the unit 2- sphere for up to and including 30 vertices. Furthermore, Mutoh confirms that the results of Berman and Hanes for up to 8 vertices is correct. In his paper he list all maximal volume polyhedra along with an approximation of this volume. We encourage the reader to go through Mutoh's paper should they be intrigued about what some of these objects look like; as he provides a visual diagram of all maximal polyhedra for $n \leq 30$. We have provided a figure of what they appear to be for 28, 29 and 30 vertices below (taken from his paper).

Remark: Mutoh makes reference to a conjecture made by Goldberg regarding the relationship between the circumscribed and inscribed polyhedra about $\mathcal{S}^2$. He stated that the maximal volume polyhedra inscribed in $\mathcal{S}^2$ is the dual of the minimal volume circumscribed polyhedra about $\mathcal{S}^2$. This, although an interesting problem, also remains unsolved.

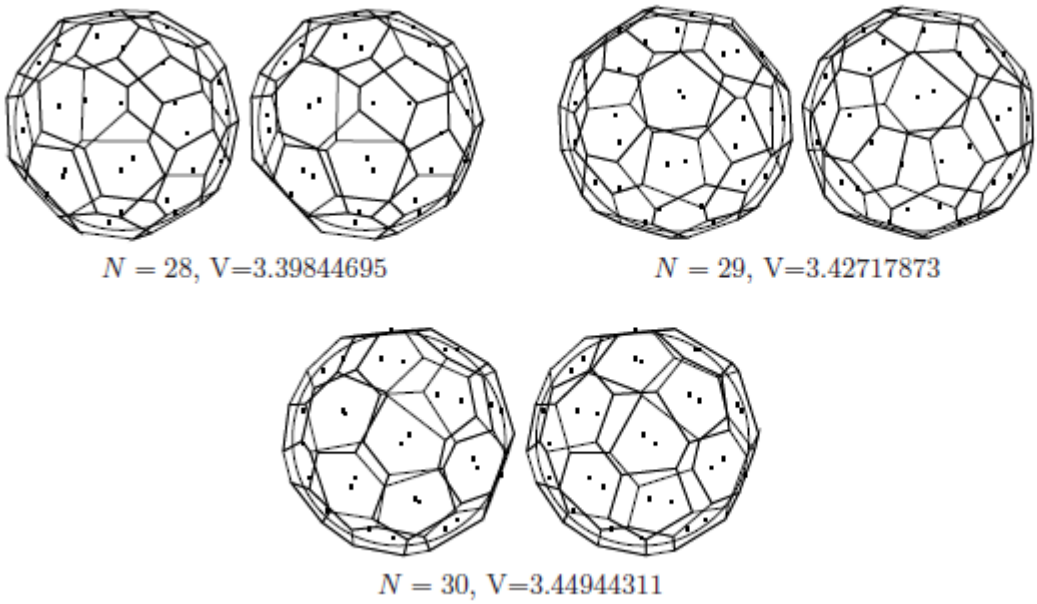

Figure 4.6: Figure taken form [29] depicting the maximal volume polyhedra for the cases of 28, 29 and 30 vertices.

Remark: This remark pertains primarily to the surface area entanglement measure. Although much of the previous two sections talked about the maximal volume polytopes in $\mathcal{S}^2$, if conjecture 4.1 is true, then we will know exactly the objects that have maximal surface area as well, and we will effectively have two measures we can use. The only real obstacle left after this is mapping which pure symmetric states actually have their QSP representations as those maximal volume/ surface area polytopes. This again hinges on conjecture 4.2, that all maximally entangled states have maximal surface area (or equivalently volume, if conjecture 4.2 proves to be true). In any case, it is a measure of entanglement, however there are many questions that remain unsolved regarding it.

4.2.5 A link to the Thomson problem in physics

In the context of our surface area measure, we are interested in how to optimize the area (and with regards to the extension of the measure) and volume of n points placed on the unit sphere. This has proved to be a difficult problem, but it is possible to view this geometry problem as a physics problem. This is the approach taken by the Thomson problem. A statement of the problem is given below.

Thomson Problem: What is the optimal configuration of placing n electrons on the surface of the unit 2-sphere that minimizes the electrostatic potential energy of the system? Mathematically this can be viewed as

$$\min(U(n))$$

where

$$U(n) = \sum_{i<j} U_{ij} = \sum_{i<j} \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}$$

[34] provides a nice summary of the problem and recent results.

In this setting, instead of placing n points on the 2-sphere and maximizing their convex hull (volume) or their surface area, Thomson's problem deals with each of the points being electrons and looking to minimize their global electrostatic potential energy. Although this is not a perfect rephrasing of the problem in the context for which is relevant to us, it could nonetheless prove to be useful as this is a fairly well studied problem in physics.

There are however some obvious similarities between the two problems, at least for

small values of n. In the mathematical context (and for our measure of entanglement), the maximal surface area polytopes for dimension 2, 3 and 4 are the straight line (diameter of the sphere), the equilateral triangle inscribed in the equator and the tetrahedron respectively. These 3 objects also happen to be the verified optimal cases for the Thomson problem. One could conjecture that this is simply a rephrasing of the exact same problem in either field.

Remark: Investigation into the Thomson problem (as recently as [34]) has found the optimal configuration of electrons for the case up to $n = 900$ with computer assistance.

Chapter 5

Conclusions and future work

5.0.1 The real higher rank numerical range

The heart of the new results presented in this paper are located in chapter 3. We investigate the properties of the real numerical range and the real higher rank numerical range, the latter of which has not been looked at in any depth until now.

Among the results presented here we first proved the elementary property that the real numerical range of any matrix A is equivalent to the real numerical range of $\frac{A+A^T}{2}$. We then go onto prove that this does not quite generalize as nicely to the higher rank (real) case. We show here that the best we can do is simply

$$R_{\lceil \frac{k}{2} \rceil}(A) \supseteq R_k\Big(\frac{A+A^T}{2}\Big)$$

Following this we then go onto state the main result of the paper, which is our real elliptical range theorem. We show that the joint real numerical range of an m- tuple

of 2×2 real symmetric matrices forms the boundary of an ellipse. And in particular, we can equivalently state $R(A) = \partial W(A)$. Furthermore, we investigate the similar case only now considering real symmetric matrices of size 3. We concluded that the object in 3 dimensions is likely not an ellipse, however, it does have elliptical cross sections (whenever ϕ, the polar angle is held constant). We use this to motivate the theorem that follows, which we call the elliptical path theorem.

Finally we discuss in the latter parts of chapter three, the starting point for our investigation in the real numerical range and its higher rank version- talking about cross sections of ellipsoids. We give a condition for an n dimensional ellipse to have a hyperspherical cross section based on the real higher rank numerical range.

5.0.2 The surface area entanglement measure

In chapter 4 we introduced the general fundamentals of quantum information theory and briefly, quantum error correction. This was done to eventually introduce a new measure of entanglement which we call the surface area measure. We defined this measure and provided a short example of how the measure might be used in practice for a pure symmetric state in 3 particles. What we state as a final remark here are some of the draw-backs and disadvantages of this measure, in its current state. Starting first, it is usually a difficult task obtaining the Majorana polynomial p_ψ directly from the state $|\psi\rangle$. The method that we provided in this paper describes formulating a degree n polynomial for a state in n particles, however, this will effectively involve

expanding out and collecting like terms from the original $n!$ permuted terms in $|\psi\rangle$. This could be worked by a symbolic calculator, such as Maple, however, it is still not time or computationally effective.

The other problem underlying this measure has to do with finding the quantum state polytope- Π of $|\psi\rangle$. Once one has found p_ψ and successfully factored this polynomial, we now have exactly the vertices of some polyhedra or polygon inscribed in $\mathcal{S}^2$. The question remains- how can we tell which collection of vertices actually form faces of Π? Now there are two answers that could be stated when asked this. One is the simple case of checking to see if each z component of each vertex is zero (or more generally, simply the same value). If this is true, then the state will give a polygon and we will only need to concern ourselves with finding the area of one single face. This is not too difficult to work out with a computer program. However, if once all solutions are plotted on $\mathcal{S}^2$ and do not all lie in the same plane, we are now faced with the problem described previously. The natural question then remains- does there exist an efficient algorithm to test what collection of solutions to p_ψ form faces of Π? This of course is necessary as we want the surface area of this figure. One possible idea is the following- we know that a face of Π must be made of at least 3 solutions to p_ψ, so, begin by testing the $\binom{n}{3}$ possible combination of triples by finding if their convex hull contains any of the other solutions. If it does, then this cannot be a face of Π. At the end of testing all $\binom{n}{3}$ combinations of faces, we keep the ones whose

convex hulls are empty. Now, color in each triple of points kept from before and if they completely enclose a subset of $\mathcal{S}^2$ then this is our polyhedra. if not, repeat the same loop for $\binom{n}{4}$ and so on. (Although this algorithm is not perfect either as this method assumes each face of Π is composed of the same number of vertices, which may not be true in general).

These two points essentially cover the major obstacles for this measure. One additional, more subtle detail one needs to be aware of is that as the number of particles increase, you can get different types of entanglement. For example, in 3 particles one can have GHZ or W entanglement. The number of distinct types of entanglement grow with n. So when we go to calculate the ratio

$$s(|\psi\rangle) = \frac{SA(\Pi_n)}{SA(\Pi_n^{max})}$$

we need to make sure that the state giving Π matches the type of entanglement Π_n^{max} corresponds to.

On one final note, there are also the conjectures to take into consideration. Solving any of these regarding the volume and surface area of a configuration of points on $\mathcal{S}^2$ could assist with this measure.

Bibliography

[1] Bloch sphere - wikipedia, the online encyclopedia. *Wikimedia Foundation*, Sep 2018.

[2] Ingemar Bengtsson and Karol Życzkowski. *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge university press, 2017.

[3] Joel D Berman and Kit Hanes. Volumes of polyhedra inscribed in the unit sphere. *Mathematische Annalen*, 188(1):78–84, 1970.

[4] Reinhold A Bertlmann and Philipp Krammer. Bloch vectors for qudits. *Journal of Physics A: Mathematical and Theoretical*, 41(23):235303, 2008.

[5] Sergey Bravyi, David Gosset, and Robert Koenig. Quantum advantage with shallow circuits. *Science*, 362(6412):308–311, 2018.

[6] Louis Brickman. On the field of values of a matrix. *Proceedings of the American Mathematical Society*, 12(1):61–66, 1961.

[7] Felix E Browder. On the spectral theory of strongly elliptic differential operators. *Proceedings of the National Academy of Sciences*, 45(9):1423–1431, 1959.

[8] Man-Duen Choi, Michael Giesinger, John A Holbrook, and David W Kribs. Geometry of higher-rank numerical ranges. *Linear and Multilinear Algebra*, 56(1-2):53–64, 2008.

[9] Man-Duen Choi, John A Holbrook, David W Kribs, and Karol Zyczkowski. Higher-rank numerical ranges of unitary and normal matrices. *arXiv preprint quant-ph/0608244*, 2006.

[10] Man-Duen Choi, David W Kribs, and Karol Życzkowski. Higher-rank numerical ranges and compression problems. *Linear algebra and its applications*, 418(2-3):828–839, 2006.

[11] Chandler Davis. The Toeplitz-Hausdorff theorem explained. *Canad. Math. Bull*, 14(2), 1971.

[12] AR Usha Devi, AK Rajagopal, et al. Majorana representation of symmetric multiqubit states. *Quantum Information Processing*, 11(3):685–710, 2012.

[13] Wolfgang Dür, Guifre Vidal, and J Ignacio Cirac. Three qubits can be entangled in two inequivalent ways. *Physical Review A*, 62(6):062314, 2000.

[14] Shmuel Friedland and Todd Kemp. Most boson quantum states are almost maximally entangled. *Proceedings of the American Mathematical Society*, 2018.

[15] Howard Georgi. *Lie algebras in particle physics: From isospin to unified theories*. CRC Press, 2018.

[16] A Gómez-Rodríguez. Entangled states and the singular value decomposition. *Revista Mexicana de Física*, 46(5):419–424, 2000.

[17] Karl Gustafson. The Toeplitz-Hausdorff theorem for linear operators. *Proceedings of the American Mathematical Society*, 25(1):203–204, 1970.

[18] Stephan Hartmann. Generalized Dicke states. *arXiv preprint arXiv:1201.1732*, 2012.

[19] Thomas Hawkins. *Emergence of the theory of Lie groups: An essay in the history of mathematics 1869–1926*. Springer Science & Business Media, 2012.

[20] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of n-particle mixed states: necessary and sufficient conditions in terms of linear maps. *Physics Letters A*, 283(1-2):1–7, 2001.

[21] Robert Hübener, Matthias Kleinmann, Tzu-Chieh Wei, Carlos González-Guillén, and Otfried Gühne. Geometric measure of entanglement for symmetric states. *Physical Review A*, 80(3):032324, 2009.

[22] Emanuel Knill and Raymond Laflamme. Theory of quantum error-correcting codes. *Physical Review A*, 55(2):900, 1997.

[23] Chi-Kwong Li. A simple proof of the elliptical range theorem. *Proceedings of the American Mathematical Society*, 124(7):1985–1986, 1996.

[24] Chi-Kwong Li and Yiu-Tung Poon. Convexity of the joint numerical range. *SIAM Journal on Matrix Analysis and Applications*, 21(2):668–678, 2000.

[25] Chi-Kwong Li and Yiu-Tung Poon. Generalized numerical ranges and quantum error correction. *Journal of Operator Theory*, pages 335–351, 2011.

[26] John Martin, Olivier Giraud, PA Braun, Daniel Braun, and Thierry Bastin. Multiqubit symmetric states with high geometric entanglement. *Physical Review A*, 81(6):062347, 2010.

[27] Alan McIntosh. The Toeplitz-Hausdorff theorem and ellipticity conditions. *The American Mathematical Monthly*, 85(6):475–477, 1978.

[28] Jarosław Adam Miszczak. Singular value decomposition and matrix reorderings in quantum information theory. *International Journal of Modern Physics C*, 22(09):897–918, 2011.

[29] Nobuaki Mutoh. The polyhedra of maximal volume inscribed in the unit sphere and of minimal volume circumscribed about the unit sphere. In *Japanese Conference on Discrete and Computational Geometry*, pages 204–214. Springer, 2002.

[30] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.

[31] Rajesh Pereira and Connor Paul-Paddock. Anticoherent subspaces. *Journal of Mathematical Physics*, 58(6):062107, 2017.

[32] Asher Peres. Separability criterion for density matrices. *Physical Review Letters*, 77(8):1413, 1996.

[33] Panayiotis J Psarrakos and Michael J Tsatsomeros. *Numerical range:(in) a matrix nutshell*. Department of Mathematics, Washington State University, 2002.

[34] Wesley JM Ridgway and Alexei F Cheviakov. An iterative procedure for finding locally and globally optimal arrangements of particles on the unit sphere. *Computer Physics Communications*, 2018.

[35] W. Weisstein, Eric. One-sheeted hyperboloid. *from Wolfram MathWorld*.

[36] W. Weisstein, Eric. Two-sheeted hyperboloid. *from Wolfram MathWorld*.