Quantum Foundations Lecture 5

February 12, 2018

Dr. Matthew Leifer
leifer@chapman.edu

HSC112

Example: Effects of the Probability Simplex

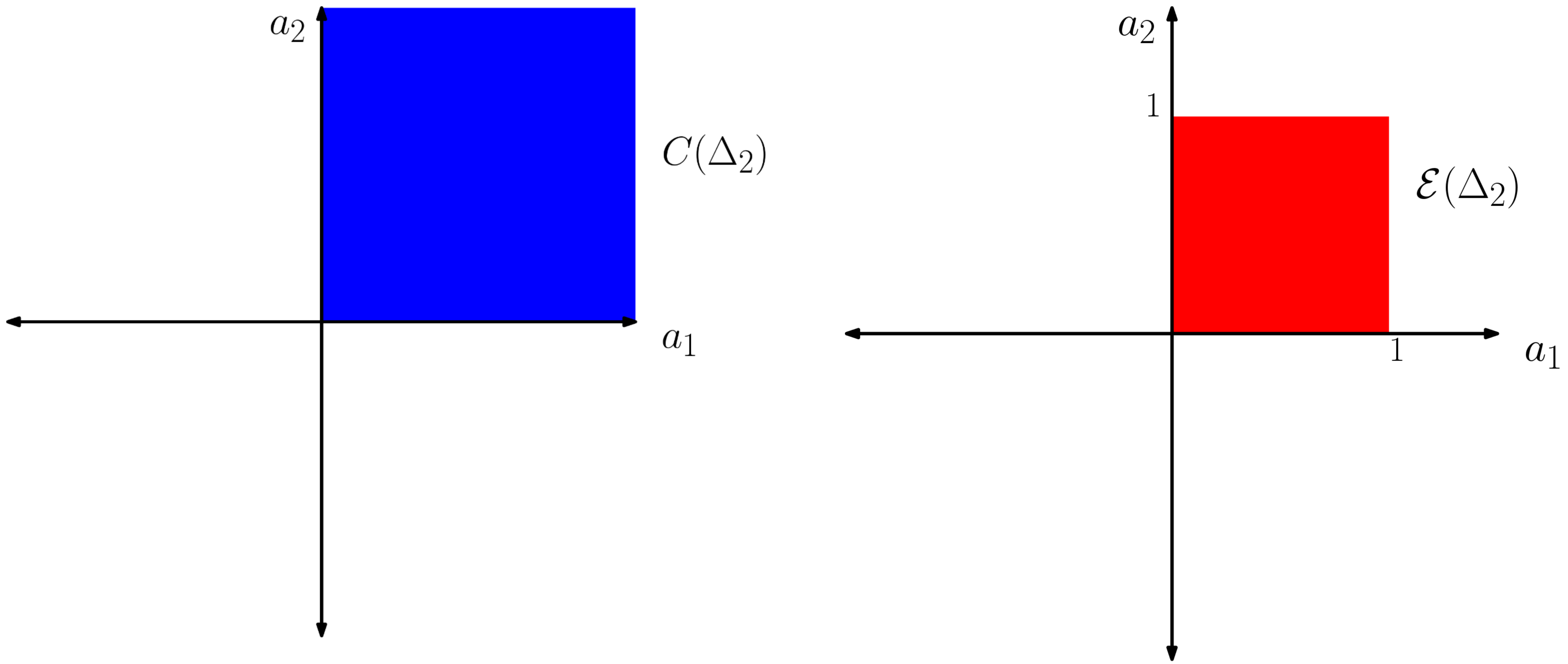
- The extreme points of the 2D probability simplex are $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$. The dual cone consists of vectors $\begin{pmatrix}a_1\\a_2\end{pmatrix}$ such that

$$\begin{pmatrix}a_1\\a_2\end{pmatrix}\cdot\begin{pmatrix}1\\0\end{pmatrix}\geq 0 \text{ or } a_1\geq 0 \text{ and } \begin{pmatrix}a_1\\a_2\end{pmatrix}\cdot\begin{pmatrix}0\\1\end{pmatrix}\geq 0 \text{ or } a_2\geq 0$$

- This is the same as the cone generated by $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$, so the 2D probability cone is self-dual.
- To get the space of effects, we have to add the conditions:

$$\begin{pmatrix}a_1\\a_2\end{pmatrix}\cdot\begin{pmatrix}1\\0\end{pmatrix}\leq 1 \text{ or } a_1\leq 1 \text{ and } \begin{pmatrix}a_1\\a_2\end{pmatrix}\cdot\begin{pmatrix}0\\1\end{pmatrix}\leq 1 \text{ or } a_2\leq 1$$

Example: Effects of the Probability Simplex



Example: Effects of the gbit/squit

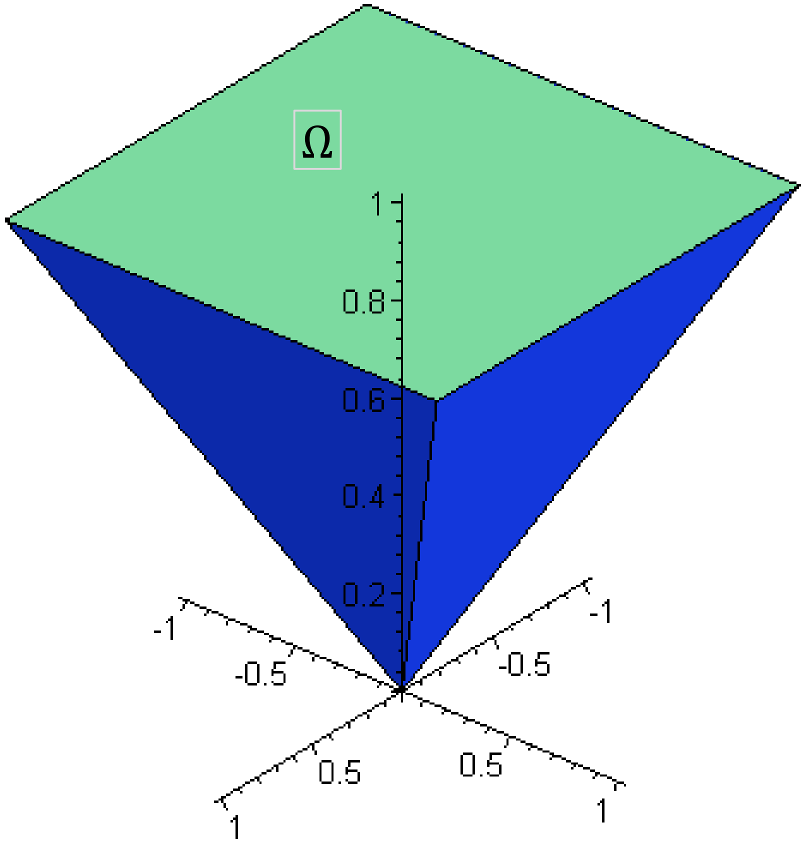
- The extreme points of the gbit are $\begin{pmatrix}-1\\-1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix},\begin{pmatrix}-1\\1\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}$, but we have to go one dimension higher for the cone, so we use:

$$\begin{pmatrix}-1\\-1\\1\end{pmatrix},\begin{pmatrix}1\\-1\\1\end{pmatrix},\begin{pmatrix}-1\\1\\1\end{pmatrix},\begin{pmatrix}1\\1\\1\end{pmatrix}$$

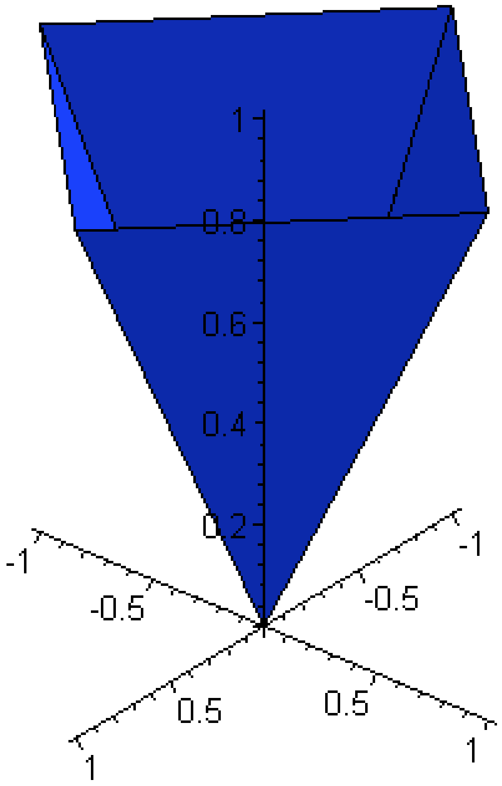
- The dual cone consists of vectors $\boldsymbol{a}=\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}$, which satisfy:

$$-a_1-a_2+a_3\geq 0,\ a_1-a_2+a_3\geq 0,\ -a_1+a_2+a_3\geq 0,\ a_1+a_2+a_3\geq 0.$$

Example: Effects of the gbit/squit



$C(\Omega)$

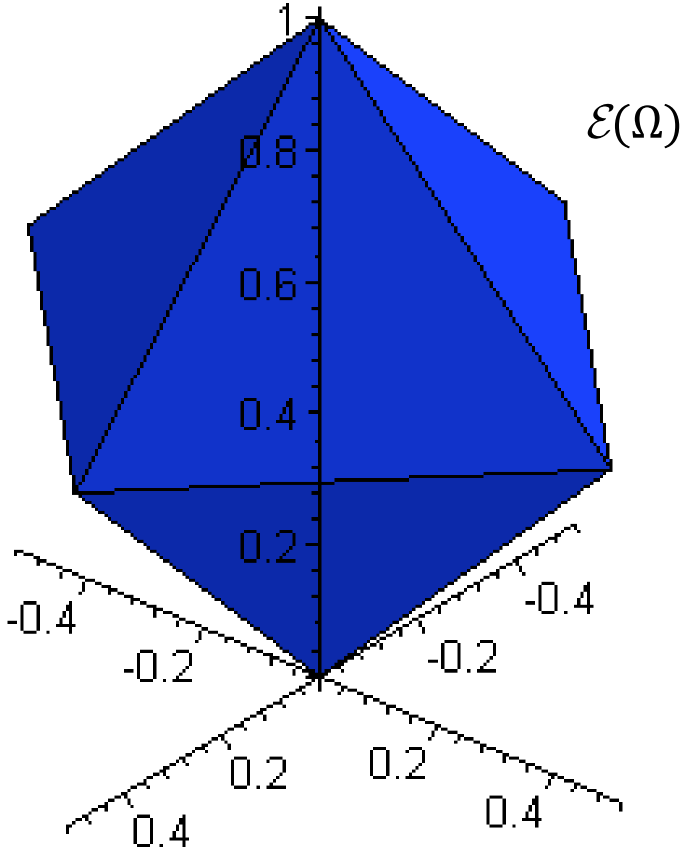


$C^*(\Omega)$

Example: Effects of the gbit/squit

- To get the space of effects, we also have to impose:

$$-a_1 - a_2 + a_3 \leq 1$$
$$a_1 - a_2 + a_3 \leq 1$$
$$-a_1 + a_2 + a_3 \leq 1,$$
$$a_1 + a_2 + a_3 \leq 1.$$



Observables

- An *observable* is a set $\{f_1, f_2, \cdots, f_n\}$ of effects, such that
$$\sum_{j=1}^{n} f_j(\omega) = 1 \text{ for all } \omega \in \Omega$$
- Alternatively, it is a set of effects such that
$$\sum_{j=1}^{n} f_j = u.$$
- We can always define a two-outcome observable from an effect f as $\{f, u - f\}$.
- Observables represent the full set of mathematical objects that can conceivably represent a well-defined measurement.

Schultz Theorem

- **Theorem**: Any closed convex subset of $\mathbb{R}^n$ is the state space of a test space.
  - F. W. Schultz, Journal of Combinatorial Theory A vol. 17, p. 317 (1974)
  - Schultz actually proved something more general, but this is the case we need here.
- Given this, we can dispense with the test space framework and just work with states and effects.

Generalized Probabilistic Theories (GPTs)

- A (finite dimensional) Generalized Probabilistic Theory (GPT) consists of
  - A closed convex set Ω called the *state space*.
  - The cone $C(\Omega)$ formed by lifting Ω.
    - Alternatively, we could start with the cone, define a unit effect u and let
      $$\Omega = \{\omega \in C | u(\omega) = 1\}.$$
  - A subset of the effects $\mathcal{E}(\Omega)$ representing the allowed measurement outcomes in the theory.
  - The observables formed from this subset.
- A theory satisfies the *no-restriction hypothesis* if we allow all effects as measurement outcomes. We will assume this here, but there are interesting GPTs that do not satisfy it.

Distinguishability in GPTs

- A set of states $\{\omega_1, \omega_2, \cdots, \omega_n\}$ is *distinguishable* in a GPT if there exists an observable $\{f_1, f_2, \cdots, f_n\}$ such that
$$f_j(\omega_k) = \delta_{jk}$$
- **Theorem**: In a simplex the set of pure states is distinguishable. In any other GPT there exists a set of pure states that is not distinguishable.
- **Proof**: In the cone generated by the simplex, the n pure states $\{\omega_1, \omega_2, \cdots, \omega_n\}$ are linearly independent, so form a basis for $\mathbb{R}^n$.
- Any basis has a dual basis $\{f_1, f_2, \cdots, f_n\}$ such that $f_j(\omega_k) = \delta_{jk}$
  - Proof: Since the ω_k's are linearly independent, we can independently define the value of f_j on each of them. The resulting set is linearly independent because it is impossible to reconstruct $f_j(\omega_j) = 1$ by a linear combination of f_k's that are all zero on ω_j.

Distinguishability in GPTs

- It remains to prove that $\sum_{j=1}^{n} f_j = u$. Since any ω can be written as a convex combination $\omega = \sum_{k=1}^{n} \alpha_k \omega_k$, we have

$$\left(\sum_{j=1}^{n} f_j\right)(\omega) = \sum_{j,k=1}^{n} \alpha_k f_j(\omega_k) = \sum_{j,k=1}^{n} \alpha_k \delta_{jk}$$
$$= \sum_{j=1}^{n} \alpha_j = 1 = u(\omega).$$

- For the converse, if Ω is not a simplex then the set of pure states is linearly dependent in the lifted cone. Let $\{\omega_1, \cdots, \omega_d\}$ be the largest linearly independent subset. Then $\omega_{d+1} = \sum_{j=1}^{d} a_j \omega_j$ and suppose a_k is nonzero. Then,

$$f_k(\omega_{d+1}) = \sum_{j=1}^{d} a_j f_k(\omega_j) = \sum_{j=1}^{d} a_j \delta_{jk} = a_k \neq 0$$

so f_k cannot be both 1 on ω_k and 0 on ω_{d+1}.

Linear Freedom In GPTs

- The probabilities in a GPT depend on the combination $f_{\boldsymbol{a}}(\omega) = \boldsymbol{a} \cdot \omega$ of a state and an effect. We can represent an effect by the row vector $\boldsymbol{a}^T = (a_1, a_2, \cdots, a_n)$ and then $\boldsymbol{a} \cdot \omega = \boldsymbol{a}^T \omega$ is just matrix multiplication.
- Let A be an invertible matrix, and let $\omega' = A\omega$, $\boldsymbol{a}^{T'} = \boldsymbol{a}^T A^{-1}$. Then
$$\boldsymbol{a}^{T'} \omega' = \boldsymbol{a}^T A^{-1} A \omega = \boldsymbol{a}^T \omega$$
- The probabilities are preserved under this transformation, so it follows that any two cones related by an invertible linear transformation describe the same theory.
- In particular, any simplex is equivalent to the probability simplex, so they all describe classical probability theory.
- We say that a GPT is *self-dual* if its cone can be mapped to a self-dual cone under a linear transformation.

3.v) Rebits

- Consider $\Omega =$ the unit disc,

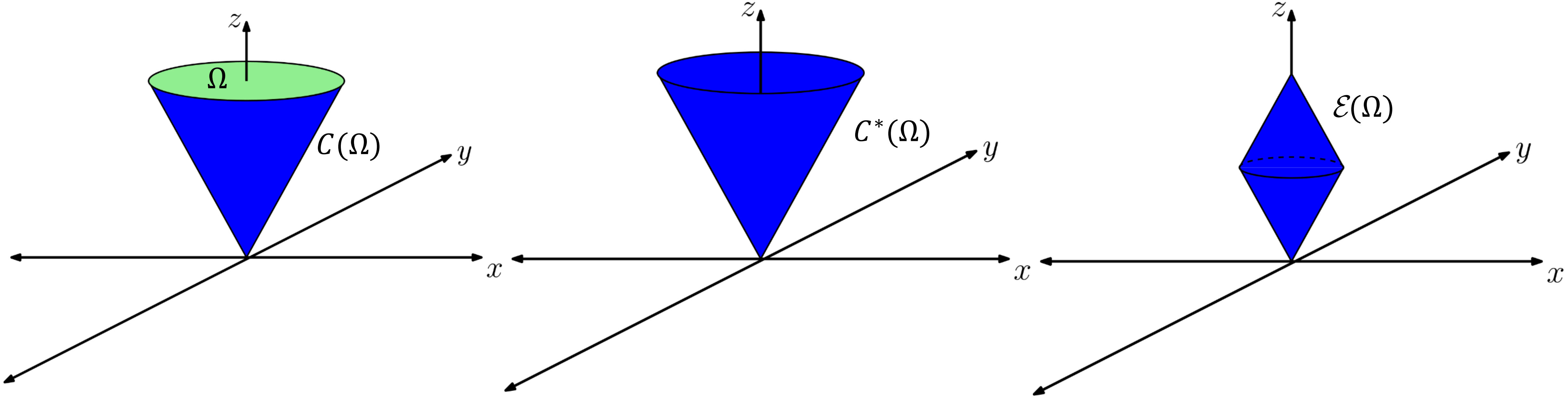
$$\begin{pmatrix} x \\ y \end{pmatrix} \text{ s.t. } x^2 + y^2 \leq 1.$$

- Lifting this to a cone gives

$$\alpha \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \text{ s.t. } x^2 + y^2 \leq 1.$$

- As you have/will show in Hwk 1, this cone is self dual.
- If we also impose $\boldsymbol{a} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \leq 1$ we get $\mathcal{E}(\Omega)$.

3.v) Rebits



Another 3-dimensional Inner Product Space

- Consider the set of 2×2 real symmetric matrices $M^T = M$ where $M^T_{jk} = M_{kj}$.
- This is a real vector space. In fact, it is an inner product space. In Hwk 1, you prove that

$$(N, M) = \mathrm{Tr}\left(N^\dagger M\right) = \sum_{j,k=1}^{n} N^*_{jk} M_{jk}$$

is an inner product on the $n \times n$ complex matrices where $N^\dagger$, defined by $N^\dagger_{jk} = N^*_{kj}$ is called the Hermitian adjoint of N.

- Since we are dealing with real-valued matrices, $N^\dagger = N^T$ in this case, so the inner product is

$$(N, M) = \mathrm{Tr}\left(N^T M\right) = \sum_{j,k=1}^{n} N_{jk} M_{jk}$$

Another 3-dimensional Inner Product Space

- The 2×2 real symmetric matrices are a 3-dimensional vector space, since there are 3 parameters in the matrix.

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix}$$

- Since it is an inner product space, we can find an orthonormal basis. You can check that $\sigma_0/\sqrt{2}, \sigma_1/\sqrt{2}, \sigma_2/\sqrt{2}$ is such a basis, where

$$\sigma_0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

- This means that instead of writing our rebit vectors in $\mathbb{R}^3$ as $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ we can write then as 2×2 matrices $\frac{1}{\sqrt{2}}(a\sigma_0 + b\sigma_1 + c\sigma_2)$.

Rebits in Matrix Space

- Actually, since we have linear freedom in choosing our representation, for the state cone, we normally use the mapping:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \rightarrow \frac{1}{2}(a\sigma_0 + b\sigma_1 + c\sigma_2)$$

i.e. we apply the linear transformation $\boldsymbol{x} \rightarrow \boldsymbol{x}/\sqrt{2}$ and then embed in matrix space. To preserve probabilities, effect vectors will then get multiplied by $\sqrt{2}$ before mapping to matrix space.

- We choose to embed our state space Ω in the σ_1, σ_2 subspace, so a normalized state is of the form

$$\frac{1}{2}(I + x\sigma_1 + y\sigma_2) \quad \text{with} \quad x^2 + y^2 \leq 1.$$

A Panolpy of Vector Spaces

- We should be careful to note that we have now introduced three inner product spaces, and we do not want to mix them up.
  - We have the three dimensional vector space $\mathbb{R}^3$ that we originally represented our rebit in. I will continue to use bold for vectors $\boldsymbol{x}$ in this space, $\boldsymbol{x}^T$ for the corresponding dual vector and $\boldsymbol{x} \cdot \boldsymbol{y}$ for the inner product.
  - We have the three dimension space of 2×2 matrices. I will use capitals M for matrices in this space, M^T for the corresponding dual, and $\mathrm{Tr}\left(N^T M\right)$ for their inner product.
  - We also have the two-dimensional vector space $\mathbb{R}^2$ that the matrices act on. Following quantum conventions, I use Dirac notation for this space:
    - $|x\rangle = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}$ for vectors, $\langle x| = (x_0, x_1)$ for dual vectors and $\langle x|y\rangle = x_0 y_0 + x_1 y_1$ for inner products.
    - We also introduce the basis vectors $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, so we can write any vector as $|x\rangle = x_0|0\rangle + x_1|1\rangle$.

Hermitian Adjoints

- For an $n \times n$ complex matrix M, its *Hermitian adjoint* $M^\dagger$ is given by
$$M^\dagger_{jk} = M^*_{kj}$$
  - For real matrices we can use $M^T = M$, where $M^T_{jk} = M_{kj}$
- Taking the Hermitian adjoint reverses the order of products, i.e. $(MN)^\dagger = N^\dagger M^\dagger$.

$$(MN)^\dagger_{jk} = (MN)^*_{kj} = \left(\sum_m M_{km} N_{mj}\right)^* = \sum_m M^*_{km} N^*_{mj}$$
$$= \sum_m M^\dagger_{mk} N^\dagger_{jm} = \sum_m N^\dagger_{jm} M^\dagger_{mk} = \left(N^\dagger M^\dagger\right)_{jk}$$

Hermitian Matrices

- An $n \times n$ complex matrix is Hermitian if $M^\dagger = M$.
  - For real matrices we can use $M^T = M$, where $M^T_{jk} = M_{kj}$.
- A *projection operator* or *projector* P is a Hermitian matrix that also satisfies $P^2 = P$.
  - The identity matrix I is a projector, since $I_{jk} = \delta_{jk}$ and $I^2 = I$.
- An *outer product* is a matrix of the form $|y\rangle\langle x|$ define to act via
$$(|x\rangle\langle y|)|z\rangle = |x\rangle\langle y|z\rangle$$
where $\langle y|z\rangle$ is the inner product.
- If $|x\rangle$ is a unit vector ($\langle x|x\rangle = 1$) then $|x\rangle\langle x|$ is a projector because
$$|x\rangle\langle x|x\rangle\langle x| = |x\rangle\langle x|$$

Eigenvalues and Eigenvectors

- A vector $|x\rangle$ is an *eigenvector* of a matrix M if

$$M|x\rangle = \lambda|x\rangle,$$

where λ is a scalar called an *eigenvalue* of M.

  - e.g. All vectors are eigenvectors of I, since $I|x\rangle = |x\rangle$.

- In general, more than one eigenvector may correspond to the same eigenvalue. The *eigenspace* of an eigenvalue is the linear subspace spanned by the corresponding eigenvectors.
  - If the eigenspace is one-dimensional, the eigenvalue is called *nondegenerate*.
  - If the eigenspace has dimension ≥ 2, the eigenvalue is called *degenerate*.

Eigenvalues/vectors of Hermitian matrices

- **Theorem**: All the eigenvalues of a Hermitian operator are real and eigenvectors belonging to different eigenspaces are orthogonal.
- A matrix for which all the eigenvalues are nondegenerate is called a *non-degenerate* matrix.
- The eigenvectors of a non-degenerate Hermitian matrix form a complete orthonormal basis.
- As a result, we can write such a matrix as

$$M = \sum_n \lambda_n |x_n\rangle\langle x_n|$$

where λ_j are the eigenvalues and $|x_j\rangle$ are the eigenvectors.

Eigenvalues/vectors of Hermitian matrices

- To see why, note that if $\{|x_j\rangle\}$ is a complete orthonormal basis then any vector $|x\rangle$ can be written as $|x\rangle = \sum_j a_j |x_j\rangle$, and then

$$M|x\rangle = \sum_j a_j M|x_j\rangle = \sum_j a_j \lambda_j |x_j\rangle = \sum_{j,k} a_k \lambda_j \delta_{jk} |x_k\rangle$$

$$= \sum_{j,k} a_k \lambda_j |x_j\rangle\langle x_j | x_k\rangle = \left(\sum_j \lambda_j |x_j\rangle\langle x_j|\right)\left(\sum_k a_k |x_k\rangle\right)$$

- This has the following generalization:
- **Theorem**: *The Spectral Theorem:* Any Hermitian matrix can be written as $M = \sum_j \lambda_j P_j$, where the λ_j are distinct eigenvalues and the P_j are the projectors onto the corresponding eigenspaces, i.e.

$P_j|x_j\rangle = |x_j\rangle$ for any vector in the eigenspace and $P_j|x\rangle = 0$ for vectors orthogonal to the eigenspace.

Positive Matrices

- An $n \times n$ Hermitian matrix is called *positive* if

$$\langle x|M|x\rangle \geq 0 \text{ for all vectors } |x\rangle.$$

- **Theorem**: A Hermitian matrix M is positive iff its eigenvalues are ≥ 0.
- **Proof**: Only if direction: Let $|x\rangle$ be an eigenvector of M with eigenvalue λ. Then,

$$\langle x|M|x\rangle = \lambda\langle x|x\rangle = \lambda\langle x|x\rangle$$

- By positivity $\langle x|M|x\rangle \geq 0$, so $\lambda\langle x|x\rangle$ and hence $\lambda \geq 0$ by positivity of the inner product.
- If direction: M can be written in its spectral decomposition as

$$M = \sum_j \lambda_j |x_j\rangle\langle x_j|$$

where λ_j are the eigenvalues and $|x_j\rangle$ the eigenvectors of M.

Positive Matrices

- If we assume that $\lambda_j \geq 0$ and let $|x\rangle$ be any vector then

$$\langle x|M|x\rangle = \langle x|\left(\sum_j \lambda_j |x_j\rangle\langle x_j|\right)|x\rangle = \sum_j \lambda_j \langle x|x_j\rangle\langle x_j|x\rangle = \sum_j \lambda_j \left|\langle x_j|x\rangle\right|^2 \geq 0$$

Properties of the Rebit Matrix Representation

- A normalized state in our matrix representation is of the form

$$\rho = \frac{1}{2}(I + x\sigma_1 + y\sigma_2)$$

- This is called a *density* matrix in quantum mechanics.
- Let's define the vector $\boldsymbol{n} = \begin{pmatrix} x \\ y \end{pmatrix}$ and find the eigenvalues and eigenvectors of ρ.

$$\rho = \frac{1}{2}\begin{pmatrix} 1+y & x \\ x & 1-y \end{pmatrix}$$

- The characteristic equation is

$$\begin{vmatrix} 1+y-2\lambda & x \\ x & 1-y-2\lambda \end{vmatrix} = 0 \text{ or } 4\lambda^2 - 4\lambda + 1 - x^2 - y^2 = 0.$$

Properties of the Rebit Matrix Representation

- The solution of this quadratic equation is
$$\lambda_\pm = \frac{1}{2}\left(1 \pm \sqrt{x^2 + y^2}\right) = \frac{1}{2}(1 \pm \|\boldsymbol{n}\|)$$
- Since $0 \leq \|\boldsymbol{n}\| = \sqrt{x^2 + y^2} \leq 1$, the eigenvalues satisfy $0 \leq \lambda_\pm \leq 1$, so ρ is a positive matrix.
- Also $\lambda_+ + \lambda_- = 1$, so $\mathrm{Tr}(\rho) = 1$ because the trace is basis independent.
- In this way, we can get *any* positive symmetric 2×2 matrix that satisfies $\mathrm{Tr}(\rho) = 1$, so it is equivalent to define a (normalized) density operator as a positive symmetric 2×2 matrix that satisfies $\mathrm{Tr}(\rho) = 1$.

Properties of the Rebit Matrix Representation

- To find the eigenvectors, let's switch to polar coordinates

$$x = r\sin\theta \quad y = r\cos\theta$$

since then $\|\boldsymbol{n}\| = r$.

- In these coordinates, we have

$$\rho = \begin{pmatrix} 1 + r\cos\theta & r\sin\theta \\ r\sin\theta & 1 - r\cos\theta \end{pmatrix}$$

- It is now straightforward to check that the two orthogonal unit vectors

$$|n+\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix} = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle \quad \text{and} \quad |n-\rangle = \begin{pmatrix} \sin\frac{\theta}{2} \\ -\cos\frac{\theta}{2} \end{pmatrix} = \sin\frac{\theta}{2}|0\rangle - \cos\frac{\theta}{2}|1\rangle$$

are the eigenvectors with $\rho|n\pm\rangle = \lambda_{\pm}|n\pm\rangle$.

Pure Rebit States

- If the state is pure then $\|\boldsymbol{n}\| = \sqrt{x^2 + y^2} = 1$, so $\lambda_+ = 1$ and $\lambda_- = 0.$ As a result, the density operator is
$$\rho = |n+\rangle\langle n+|$$
- This is just the projector onto the one-dimensional subspace spanned by $|n+\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle$.
- In quantum mechanics, we often use the vector $|n+\rangle$ to represent a pure state rather than the projector $|n+\rangle\langle n+|$. This is just a matter of convenience.
- The space of