COPENHAGEN PRE-COURSE

VERN I. PAULSEN

Abstract. These notes contain much of the mathematics that I will be discussing and/or assuming you are familiar with for my lectures in Copenhagen.

Contents

1. Hilbert Spaces

All vector spaces will be over $\mathbb{C}$ unless specified otherwise. Given a vector space V a map $B : V \times V \to \mathbb{C}$ is **sesquilinear** provided:

- $B(v_1 + v_2, w) = B(v_1, w) + B(v_2, w)$
- $B(v, w_1 + w_2) = B(v, w_1) + B(v, w_2)$,
- $\forall \lambda \in \mathbb{C},\ B(\lambda v, w) = \overline{B}(v, w),\ B(v, \lambda w) = \lambda B(v, w)$.

We call B **positive semidefinite** provided that $B(v, v) \geq 0, \forall v \in V$ and **positive**(or **positive definite**) provided that $B(v, v) > 0$ for all $v \neq 0$. A positive sesquilinear map is called an **inner product** and in this case we generally write

$$B(v, w) = \langle v|w\rangle.$$

Proposition 1.1 (Cauchy-Schwartz Inequality)**.** *Let $B : V \times V \to \mathbb{C}$ be sesquilinear and positive semidefinite, then $B(v, w) = \overline{B(w, v)}$ and*

$$|B(v, w)|^2 \leq B(v, v)B(w, w).$$

Corollary 1.2. *Let $B : V \times V \to \mathbb{C}$ be positive semidefinite and sesquilinear, then*

- *$\{x : B(x, x) = 0\} = \{x : B(x, w) = 0 \forall w\}$ is a subspace of V that we denote by $\mathcal{N}$,*
- *there is a well-defined inner product on the quotient space $V/\mathcal{N}$ given by*

$$\dot{B}\,(x + \mathcal{N}, y + \mathcal{N}) = B(x, y).$$

Givne an inner product on V if we set

$$\|v\| = \langle v|v\rangle^{1/2},$$

then this is a norm on V. When $(V, \|\cdot\|)$ is a complete normed space with respect to the norm coming from an inner product then we call V a **Hilbert space**.

If V is a HIlbert space then a set of vectors S is called **orthonormal(o.n.)** provided that $v \in S \implies \|v\| = 1$ and $v, w \in S, v \neq w \implies \langle v|w\rangle = 0$. A set S is called an **orthonormal basis(o.n.b.)** provided that it is an orthonormal set and it is maximal among all orthonormal sets. i.e., $S \subseteq T$ and T also o.n. implies that $S = T$.

Theorem 1.3 (Parseval)**.** *Let $\mathcal{H}$ be a HIlbert space, $\{e_a : a \in A\}$ an o.n.b., then for any $h \in \mathcal{H}$,*

(1) $\|h\|^2 = \sum_{a\in A} |\langle e_a|h\rangle|^2$,
(2) $h = \sum_{a\in A} \langle e_a|h\rangle e_a$.

We need to explain what these unordered sums mean. For example 2) means that given $\epsilon > 0$ there exists a finite set $F_0 \subseteq A$ such that if F is any finite set with $F_0 \subseteq F \subseteq A$, then

$$\|h - \sum_{a\in F} \langle e_a|h\rangle e_a\| < \epsilon.$$

While 1) gives that for any $\epsilon > 0$ there is a finite set F_0 such that for any finite set F, $F_0 \subseteq F \subseteq A$ we have that

$$0 \leq \|h\|^2 - \sum_{a\in F} |\langle e_a|h\rangle|^2 < \epsilon.$$

A good example to keep in mind is that

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n},$$

converges while

$$\sum_{n\in\mathbb{N}} \frac{(-1)^n}{n},$$

does not converge.

Proposition 1.4 (Hilbert Space Dimension)**.** *Let $\mathcal{H}$ be a HIlbert space and let $\{e_a : a \in A\}$ and $\{f_b : b \in B\}$ be two o.n.b.'s for $\mathcal{H}$. Then there is a one-to-one, onto function,*

$$g : A \to B.$$

The existence of such a function g is the definition of what it means for the sets A and B to have the same **cardinality**. So this statement is also written as

$$card(A) = card(B),$$

and we denote this number by $dim(\mathcal{H})$ or sometimes $dim_{HS}(\mathcal{H})$. We will sometimes use the following.

Proposition 1.5. *Let $\mathcal{H}$ be a Hilbert space. Then $\mathcal{H}$ ha an o.n.b. that is at most countable if and only if $\mathcal{H}$ is separable as a metric space, i.e., has a countable dense subset.*

1.1. **Direct Sums.** Given two Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, we set

$$\mathcal{H} \oplus \mathcal{K} = \{(h, k) : h \in \mathcal{H}, k \in \mathcal{K}\}.$$

This is a vector space with $(h_1, k_1) + (h_2, k_2) = (h_1 + h_2, k_1 + k_2)$, and $\lambda(h, k) = \lambda h, \lambda k)$. If we set

$$\langle (h_1, k_1) | (h_2, k_2) \rangle = \langle h_1 | h_2 \rangle_{\mathcal{H}} + \langle k_1 | k_2 \rangle_{\mathcal{K}},$$

then this is an inner product that makes the vector space $\mathcal{H} \oplus \mathcal{K}$ into a Hilbert space, called the **direct sum**. Note that

$$dim(\mathcal{H} \oplus \mathcal{K}) = dim(\mathcal{H}) + dim(\mathcal{K}),$$

which justifies the notation a bit.

We set

$$\mathcal{H}^{(n)} := \mathcal{H} \oplus \mathcal{H} \oplus \cdots \mathcal{H}(n\, copies),$$

which denotes the direct sum of n copies of $\mathcal{H}$ with itself. When we want to form a direct sum of infinitely many copies of $\mathcal{H}$ with itself we cannot use all posible tuples, because the inner products would not converge. Instead we set

$$\mathcal{H}^{(\infty)} := \{(h_1, h_2, \ldots) : h_n \in \mathcal{H} \text{ and } \sum_{n \in \mathbb{N}} \|h_n\|^2 < +\infty\},$$

with inner product,

$$\langle (h_1, h_2, \ldots) | (k_1, k_2, \ldots) \rangle := \sum_{n \in \mathbb{N}} \langle h_n | k_n \rangle.$$

1.2. **Tensor Products.** Given two Hilbert spaces, $\mathcal{H}$ and $\mathcal{K}$, let $\mathcal{H} \otimes \mathcal{K}$ denote the tensor product of these two vector spaces. Given $u = \sum_{i=1}^{n} h_i \otimes k_i$ and $v = \sum_{j=1}^{k} x_j \otimes y_j$ in $\mathcal{H} \otimes \mathcal{K}$, we set

$$\langle u | v \rangle = \sum_{i,j=1}^{n,k} \langle h_i | x_j \rangle_{\mathcal{H}} \cdot \langle k_i | y_j \langle_{\mathcal{K}}.$$

This turns out to define an inner product. If one of the two Hilbert spaces is finite dimensional, then this space is already complete in this inner product, but when they are both infinite dimensional, this space is not complete. However, we still use $clH \otimes \mathcal{K}$ to denote the Hilbert space that is the completion. (Some authors prefer to use $\mathcal{H} \otimes \mathcal{K}$ for the vector space tensor product and $\mathcal{H} \overline{\otimes} \mathcal{K}$ for the completion. Other authors use $\mathcal{H} \odot \mathcal{K}$ for the vector space tensor product and $\mathcal{H} \otimes \mathcal{K}$ for its completon.)

Te following summarizes some of the key properties of the tensor product.

Theorem 1.6. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces.*

(1) *If* $\{e_a; a \in A\}$ *is an o.n.b. for* $\mathcal{H}$ *and* $\{f_b : b \in B\}$ *is an o.n.b. for* $\mathcal{K}$*, then* $\{e_a \otimes f_b : a \in A, b \in B\}$ *is an o.n.b. for* $\mathcal{H} \otimes \mathcal{K}$.
(2) $dim(\mathcal{H} \otimes \mathcal{K}) = dim(\mathcal{H}) \cdot dim(\mathcal{K})$.
(3) *Given* $u \in \mathcal{H} \otimes \mathcal{K}$ *there exist unique vectors* $h_h \in \mathcal{H}$ *such that*

$$u = \sum_{b \in B} h_b \otimes f_b.$$

Similarly, there exist unique vectors $k_a \in \mathcal{K}$ *such that*

$$u = \sum_{a \in A} e_a \otimes k_a.$$

Also,

$$\|u\|^2 = \sum_{b \in B} \|h_b\|^2 = \sum_{a \in A} \|k_a\|^2.$$

1.3. **Identifying direct Sums and Tensor Products.** We shall often use the following identification. Let $\mathcal{H}$ be a Hilbert space and let $\mathbb{C}^n$ be the usual n dimensional Hilbert space. Fix some o.n.b. for $\mathbb{C}^n$, $f_1, ..., f_n$. We define

$$U : \mathcal{H}^{(n)} \to \mathcal{H} \otimes \mathbb{C}^n,$$

by setting

$$U((h_1, ..., h_n)) = \sum_{j=1}^{n} h_j \otimes f_j.$$

This map is one-to-one, onto and **inner product preserving**, namely

$$\langle U(h_1, ..., h_n))|U((k_1, ..., k_n))\rangle_{\mathcal{H} \otimes \mathbb{C}_n} = \sum_{j=1}^{n} \langle h_j|k_j\rangle_{\mathcal{H}} = \langle (h_1, ..., h_n)|(k_1, ..., k_n)\rangle_{\mathcal{H}^{(n)}}.$$

Thus, as Hilbert spaces these spaces are identical. The map U is an example of a unitary map, which we shall discuss more later.

1.4. **Subspaces.** Let $\mathcal{H}$ be a HIlbert space and let $\mathcal{M} \subseteq \mathcal{H}$ be a vector subspace that is also closed in the norm topology. In this case $\mathcal{M}$ is also a Hilbert space. If we set

$$\mathcal{M}^{\perp} := \{h \in \mathcal{H} : \langle h|m\rangle = 0, \forall m \in \mathcal{M}\},$$

then $\mathcal{M}^{\perp}$ is also a closed vector subspace of $\mathcal{H}$. Moreover, every $h \in \mathcal{H}$ has a unique decomposition as $h = m + n$ with $m \in \mathcal{M}$ and $n \in \mathcal{M}^{\perp}$.

1.5. **Bra–ket Notation.** Generally, when we have a vector in $\mathbb{C}^n$, for ease of typing, we write it as a row vector $v = (x_1, ..., x_n)$, yet when we think of vectors and matrices we actually need v to be a column vector. For this reason, matrix theorists really like to think of vectors as columns. Also given another vector $w = (y_1, ..., y_n)$, the inner product is

$$\langle w|v\rangle = \sum_{i=1}^{m} \overline{y_i} x_i.$$

Note that if we do think of v and w as columns, $v = (x_1, ..., x_n)^t$ and $w = (y_1, ..., y_n)^t$ where t denotes the transpose, then the inner product is:

$$\langle w|v\rangle = w^* v,$$

where of course $w^* = (\overline{y_1}, ..., \overline{y_n})$ is the **conjugate transpose** of the column vector w. The fact that matrix theory really wants vectors to be columns is also why we like to have our inner product conjugate linear on the left. If we had made it conjugate linear on the right, then we would have had $\langle w|v\rangle = v^* w$!

Physicists get around this ambiguity with their bra-ket notation. Formally, they always denote vectors by $|v\rangle$, called the "ket of v", and the linear functional

$$f_w : \mathcal{H} \to \mathbb{C},\ f_w(v) = \langle w|v\rangle,$$

induced by the vector w as $\langle w|$, called the "bra of w". This makes the inner product,

$$\langle w||v\rangle.$$

In my notation, $|v\rangle = v$ and $\langle w| = f_w = w^*$.

2. Operator Theory

Let $\mathcal{H}$ and $\mathcal{K}$ denote Hilbert spaces. We let $B(\mathcal{H},\mathcal{K})$ denote the set of **bounded**, linear maps from $\mathcal{H}$ to $\mathcal{K}$. Recall that $T:\mathcal{H}\to\mathcal{K}$ bounded means that,

$$\|T\| := \sup\{\|Th\|_{\mathcal{K}} : h\in\mathcal{H}, \|h\|_{\mathcal{H}}=1\} = \sup\{\frac{\|Th\|_{\mathcal{K}}}{\|h\|_{\mathcal{H}}} : h\neq 0\} < +\infty.$$

When $\mathcal{H}=\mathcal{K}$, we abbreviate, $B(\mathcal{H},\mathcal{H}) = B(\mathcal{H})$.

Given $T:\mathbb{C}^d\to\mathbb{C}^r$ we can always represent T as multiplication by an $r\times d$ matrix $(t_{i,j})$ where

$$t_{i,j} = \langle e_i|Te_j\rangle.$$

A useful bound is that

$$\|T\| \leq \big(\sum_{j=1}^{d}\sum_{i=1}^{r}|t_{i,j}|^2\big)^{1/2} := \|T\|_2,$$

where this latter quantity is the norm of the matrix viewed as a vector in the Hilbert space $\mathbb{C}^{rd}$.

2.1. Adjoint.

Given $T\in B(\mathcal{H},\mathcal{K})$ there is a unique operator $R\in B(\mathcal{K},\mathcal{H})$ satisfying

$$\langle k|Th\rangle_{\mathcal{K}} = \langle Rk|h\rangle_{\mathcal{H}}.$$

This operator is called the **adjoint of T** and is denoted by $T^* := R$.

When T is represented by the matrix $(t_{i,j})$, then T^* is represented by the matrix that is the conjugate, transpose, $T^* = (\overline{t_{j,i}})$.

There are several different types of operators that play an important role. We review their names and some characterizations.

$V\in B(\mathcal{H},\mathcal{K})$ is an **isometry** provided $\|Vh\|_{\mathcal{K}} = \|h\|_{\mathcal{H}}, \forall h\in\mathcal{H}$.

Proposition 2.1. *T.F.A.E.*

1. *V is an isometry,*
2. *V is* **inner product preserving**, *i.e.,*

$$\langle Vh_1|Vh_2\rangle_{\mathcal{K}} = \langle h_1|h_2\rangle_{\mathcal{H}}, \forall h_1,h_2\in\mathcal{H},$$

3. *$V^*V = I_{\mathcal{H}}$.*

A map $U\in B(\mathcal{H},\mathcal{K})$ is called a **unitary** provided U is an isometry and is onto.

Proposition 2.2. *T.F.A¿E.*

1. *U is a unitary,*
2. *U and U^* are isometries,*
3. *$U^*U = I_{\mathcal{H}}$ and $UU^* = I_{\mathcal{K}}$.*
4. *U is invertible and $U^{-1} = U^*$.*

A map $H \in B(\mathcal{H})$ is called **Hermitian** or **self-adjoint** provided that $H = H^*$.

A map $N \in B(\mathcal{H})$ is called **normal** provided that $NN^* = N^*N$.

A map $P \in B(\mathcal{H})$ is called a **projection** provided that there is a closed subspace $\mathcal{M} \subseteq \mathcal{H}$ such that $Ph = m$ where $h = m + n$, $m \in \mathcal{M}, n \in \mathcal{M}^\perp$ is the unique decomposition of h.

Given $T \in B(\mathcal{H}, \mathcal{K})$ we set

$$\mathcal{R}(T) = \{Th : h \in \mathcal{H}\},$$

which is a subspace of $\mathcal{K}$ that we call the **range** of T.

Proposition 2.3. *P is a projection if and only if $P = P^* = P^2$ and in this case $\mathcal{M} = \mathcal{R}(P)$*

A map $F \in B(\mathcal{H}, \mathcal{K})$ is called **finite rank** provided that $\mathcal{R}(F)$ is finite dimensional.

Proposition 2.4. *F is finite rank if and only if there exist finitely many vectors, $h_1, ..., h_n \in \mathcal{H}$ and $k_1, ..., k_n \in \mathcal{K}$ such that*

$$Fh = \sum_{i=1}^{n} \langle h_i | h \rangle k_i.$$

In bra-ket notation, $F = \sum_{i=1}^{n} |k_i\rangle \langle h_i|$.

Back to matrices. If $h = (\alpha_1, ..., \alpha_n)^t \in \mathbb{C}^n$ and $k = (\beta_1, ..., \beta_m)^t \in \mathbb{C}^m$ then

$$kh^* = |k\rangle \langle h| = (\beta_i \overline{\alpha_j}),$$

which is an $m \times n$ rank one matrix.

When $\|h\| = 1$, then

$$hh^* = |h\rangle \langle h| = (\alpha_i \overline{\alpha_j}),$$

is the rank one projection onto the span of h. If $\{v_1, ..., v_n\}$ are orthonormal, then

$$\sum_{i=1}^{n} v_i v_i^* = \sum_{i=1}^{n} |v_i\rangle \langle v_i| ,$$

is the projection onto the n-dimensional subspace that they span.

A map $K \in B(\mathcal{H}, \mathcal{K})$ is called **compact** provided that there is a sequence of finite rank operators $F_n \in B(\mathcal{H}, \mathcal{K})$ such that

$$\lim_n \|K - F_n\| = 0.$$

We let $\mathbb{K}(\mathcal{H}, \mathcal{K})$ denote the set of compact operators from $\mathcal{H}$ to $\mathcal{K}$.

Proposition 2.5. *The set $\mathbb{K}(\mathcal{H}, \mathcal{K}) \subseteq B(\mathcal{H}, \mathcal{K})$ is closed subspace in the operator norm. If $T \in B(\mathcal{H}), K \in \mathbb{K}(\mathcal{H}, \mathcal{K})$ and $R \in B(\mathcal{K})$, then $RKT \in \mathbb{K}(\mathcal{H}, \mathcal{K})$.*

2.2. **Spectrum and Functional Calculus.** If $T \in B(\mathcal{H})$ with $\mathcal{H}$ infinite dimensional, then it is possible that T has no eigenvalues even when $T = T^*$.

For example, if

$$\mathcal{H} = \ell^2_{\mathbb{N}} := \{(a_1, a_2, ...) : \sum_{n\in\mathbb{N}} |a_n|^2 < +\infty\},$$

then this space has an o.n.b. given by $\{e_n : n \in \mathbb{N}\}$ where e_n is the vector that is 1 in the n-th coordinate and 0 elsewhere. The operator defined by

$$Se_n = e_{n+1}$$

is called the **forward unilateral shift** and it is easy to show that it has no non-zero eigenvector. However it's adjoint, S^* is the **backwards unilateral shift** and satisfies

$$S^*e_n = \begin{cases} 0 & n = 1 \\ e_{n-1} & n > 1 \end{cases}.$$

Given $\lambda \in \mathbb{C}$, $|\lambda| < 1$, if we set

$$v_\lambda = (1, \lambda, \lambda^2, ...) = \sum_{n\in\mathbb{N}} \lambda^{n-1} e_n,$$

then $S^*v_\lambda = \lambda v_\lambda$. Thus, although S has no eigenvectors, there is an eigenvector for S^* for every point in the open unit disk.

In infinite dimensions the spectrum plays the role of the eigenvectors. Given $T \in B(\mathcal{H})$ the **spectrum of T** is the set

$$\sigma(T) = \{\lambda \in \mathbb{C} |\ (T - \lambda I_{\mathcal{H}}) \text{ is not invertible } \}.$$

Theorem 2.6. *Let $T \in B(\mathcal{H})$, then $\sigma(T)$ is a non-empty compact set and*

$$\sigma(T) \subseteq \{\lambda \in \mathbb{C} :\ |\lambda| \leq \|T\|\}.$$

In fact,

$$\sup\{|\lambda| : \lambda \in \sigma(T)\} = \lim_n \|T^n\|^{1/n}.$$

This last equation is called the **spectral radius formula**.

Here are a few other facts about the spectrum that we shall often use. Given a polynomial, $p(z) = a_0 + a_1 z + \cdots + a_n z^n$ and $T \in B(\mathcal{H})$ we set $p(T) = a_0 I_{\mathcal{H}} + a_1 T + \cdots + a_n T^n$.

Theorem 2.7. *Let $T \in B(\mathcal{H})$.*

1. *$\sigma(p(T)) = \{p(\lambda) : \lambda \in \sigma(T)\}$.*
2. *If $T = T^*$, then $\sigma(T) \subseteq \mathbb{R}$.*
3. *If U is a unitary, then $\sigma(U) \subseteq \{\lambda :\ |\lambda| = 1\}$*

2.3. **The Continuous Functional Calculus for a Hermitian Operator.** Given a function $f : S \to \mathbb{C}$ we set

$$\|f\|_\infty = \sup\{|f(x)|;\, x \in S\}.$$

Of course, this norm depends on the domain of the function but this will always be clear from the context.

Proposition 2.8. *Let $H \in B(\mathcal{H}), H = H^*$. Then for every polynomial,*

$$\|p(H)\| = \sup\{|p(\lambda)| :\, \lambda \in \sigma(H)\}.$$

Thus, $\|p(H)\| = \|p\|_\infty$ where p is viewed as a function on $\sigma(H)$.

Let $C(\sigma(H))$ denote the set of continuous functions on $\sigma(H) \subseteq \mathbb{R}$. Recall by the Stone-Weierstrass theorem that the polynomials are dense in this set in $\|\cdot\|_\infty$. So given any continuous function f there is a sequence of polynomial $\{p_n\}$ with $\lim_n \|f - p_n\|_\infty = 0$. From this it follows that this sequence is Cauchy in norm, i.e., given $\epsilon > 0$, for m, n sufficiently large, $\|p_n - p_m\|_\infty < \epsilon$. But this means that the operators $\{p_n(H)\}$ are also Cauchy in norm, since

$$\|p_n(H) - p_m(H)\| = \|p_n - p_m\|_\infty.$$

Hence, there will be an operator there will be an operator to which they converge and this operator is denoted by $f(H)$.

Thus, for each $f \in C(\sigma(H))$ we have an operator $f(H)$. We summarize a few of the properties of this construction below.

Theorem 2.9 (The Continuous Functional Calculus for a Self-Adjoint Operator)**.** *Let $H \in B(\mathcal{H}), H = H^*$. Then for every continuous function f on $\sigma(H)$, i.e., $f \in C(\sigma(H))$ there is an operator $f(H)$ these satisfy:*

- $\|f(H)\| = \|f\|_\infty$,
- $\sigma(f(H)) = \{f(\lambda) :\, \lambda \in \sigma(H)\}$,
- $f, g \in C(\sigma(H)) \implies (fg)(H) = f(H)g(H),\ (f+g)(H) = f(H) + g(H)$.

2.4. **Positive Operators.** An operator $P \in B(\mathcal{H})$ is **positive**, denoted $P \geq 0$ provided that

$$\langle h|Ph\rangle \geq 0,\ \forall h \in \mathcal{H}.$$

Proposition 2.10. *T.F.A.E.*

- $P \geq 0$,
- $P = P^*$ *and* $\sigma(P) \subseteq [0, +\infty)$,
- $\exists X \in B(\mathcal{H})$ *such that* $P = X^*X$.

Given $T \in B(\mathcal{H}, \mathcal{K})$ we use the continuous functional calculus tp define

$$|T| = (T^*T)^{1/2}.$$

Note that, unlike numbers, generally, $|T| \neq |T^*|$.

Define continuous functions $f_+, f_- : \mathbb{R} \to \mathbb{R}$ by

$$f_+(t) = \begin{cases} t & t \geq 0 \\ 0 & t < 0 \end{cases} \text{ and } f_-(t) = \begin{cases} 0 & t \geq 0 \\ -t & t < 0 \end{cases}.$$

If $H = H^*$ then we apply the continuous functional calculus to define $H^+ = f_+(H)$ and $H^- = f_-(H)$ and we see that

- $H^+ \geq 0,\ H^- \geq 0,$
- $H = H^+ - H^-$
- $|H| = H^+ + H^-,$
- $H^+H^- = 0.$

Theorem 2.11 (Polar Decomposition)**.** *Let $T \in B(\mathcal{H}, \mathcal{K})$, then there exists a unique unitary $W : \mathcal{R}(|T|)^- \to R(T)^-$ such that $T = W|T|$.*

The proof essentially follows from the fact that

$$\|Th\|^2 = \langle Th|Th\rangle = \langle h|T^*Th\rangle = \langle h||T|^2h\rangle = \langle |T|h||t|h\rangle = \||T|h\|^2.$$

We can always extend W to an operator $\hat{W} : \mathcal{H} \to \mathcal{K}$ by setting $\hat{W}$ equal to 0 on $\mathcal{R}(|T|)^\perp$, i.e.,

$$\hat{W}(|T|h + k) = Th,\ \forall k \in \mathcal{R}(|T|)^\perp$$

and we will still have $T = \hat{W}|T|$. This latter factorization is sometimes what is meant by the polar decomposition.

Moreover, if $dim(\mathcal{R}(|T|)^\perp) = dim(\mathcal{R}(T)^\perp)$ then one can also extend W to be a unitary $U : \mathcal{H} \to \mathcal{K}$ with $T = U|T|$. When $\mathcal{H} = \mathcal{K} = \mathbb{C}^n$, this is always the case, so we may always factor a $n \times n$ matrix T as $T = U|T|$ with U a unitary.

3. More about $\mathbb{K}(\mathcal{H})$

Theorem 3.1 (Positive Compact Operators)**.** *Let $P \in \mathbb{K}(\mathcal{H})$ with $P \geq 0$. Then there exists an o.n.b. $\{\psi_a : a \in A\}$ for $\mathcal{H}$ consisting of eigenvectors for P. Moreover, at most countably many of the corresponding eigenvalues are non-zero and we may arrange the non-zero eigenvalues in a decreasing sequence, $\lambda_1 \geq \lambda_2 \geq ...$ with either at most finitely many eigenvalues non-zero or $\lim_n \lambda_n = 0$.*

Given P as above, set $F_N = \sum_{n=1}^N \lambda_n |\psi_n\rangle \langle\psi_n|$. Then $F_N \geq 0$ and is finite rank, with

$$\|P - F_N\| = \lambda_{N+1} \to 0 \text{ as } N \to +\infty.$$

Thus, we may write

$$P = \sum_{n=1}^{\infty} \lambda_n |\psi_n\rangle \langle\psi_n|,$$

and the converge of this series is in the norm.

Given any $K \in \mathbb{K}(\mathcal{H}, \mathcal{K})$ by the polar decomposition we have that $K = W|K|$ and $|K| \geq 0$ and compact. The non-zero eigenvalues of $|K|$ written

in decreasing order $\lambda_1 \geq \lambda_2 \geq ....$ are called the **singular values of K** and we set

$$s_n(K) = \lambda_n.$$

If we let ψ_n denote the corresponding o.n. sequence of eigenvectors for $|K|$ and set $\phi_n = W\psi_n$ then these vectors are also o.n. and we may write

$$|K| = \sum_{n=1}^{\infty} s_n(K) \, |\psi_n\rangle \, \langle\psi_n| \, ,$$

which yields

$$K = W|K| = \sum_{n=1}^{\infty} s_n(K) \, |\phi_n\rangle \, \langle\psi_n| \, .$$

This latter form is called the **singular valued decomposition(SVD)** of K. It is essentially unique, except that in the case that a single non-zero eigenvalue has multiplicity, then one could choose different o.n. vectors for the corresponding eigenspace.

3.1. **The Schatten p-Classes.** For proofs of the results stated here see [**?**, XI.9] or [**?**, III, Section 7]. Given $1 < p < +\infty$, we set

$$\mathcal{C}_p(\mathcal{H}, \mathcal{K}) = \{K \in \mathbb{K}(\mathcal{H}, \mathcal{K}) : \sum_{n=1}^{\infty} s_n(K)^p < +\infty\},$$

and for $K \in \mathcal{C}_p(\mathcal{H}, \mathcal{K})$ we set

$$\|K\|_p = \big(\sum_{n=1}^{\infty} s_n(K)^p\big)^{1/p}.$$

Here are the key facts about these sets.

(1) For $1 < p < +\infty, \mathcal{C}_p(\mathcal{H}, \mathcal{K})$ is a vector space.
(2) $\|\cdot\|_p$ is a norm on $\mathcal{C}_p(\mathcal{H}, \mathcal{K})$ and it is complete in this norm, i.e., a Banach space.
(3) If $K \in \mathcal{C}_1(\mathcal{H})$ and we pick any o.n.b. $\{e_a : a \in A\}$ for $\mathcal{H}$, then

$$Tr(K) := \sum_{a \in A} \langle e_a | K e_a \rangle$$

converges and its value is independent of the o.n.b. chosen. We call this the **trace of K** and for this reason we call $\mathcal{C}_1(\mathcal{H})$ the **trace class operators**.
(4) If $1 < p, q < +\infty$ with $\frac{1}{p} + \frac{1}{q} = 1$(called Holder conjugates) with $T \in \mathcal{C}_p(\mathcal{H}, \mathcal{K}), R \in \mathcal{C}_q(\mathcal{K}, \mathcal{H})$, then $RT \in \mathcal{C}_1(\mathcal{H}), TR \in \mathcal{C}_1(\mathcal{K})$ and $Tr(RT) = Tr(TR)$. Moreover,

$$|Tr(RT)| \leq \|T\|_p \|R\|_q.$$

(5) Let p, q be Holder conjugates. If we fix R and define a linear functional
$$f_R : \mathcal{C}_p(\mathcal{H}, \mathcal{K}) \to \mathbb{C},\ f_R(T) = Tr(RT),$$
then f_R is a bounded, linear functional with $\|f_R\| = \|R\|_q$. Moreover, every bounded linear functional on $\mathcal{C}_p(\mathcal{H}, \mathcal{K})$ is of this form.

This identifies the dual space of $\mathcal{C}_p(\mathcal{H}, \mathcal{K})$ with $\mathcal{C}_q(\mathcal{K}, \mathcal{H})$ in an isometric manner.

(6) If $T \in B(\mathcal{H}, \mathcal{K})$ and $R \in \mathcal{C}_1(\mathcal{K}, \mathcal{H})$, then $RT \in \mathcal{C}_1(\mathcal{H})$ and $TR \in \mathcal{C}_1(\mathcal{K})$ with $Tr(RT) = Tr(TR)$. The linear functional $f_T : \mathcal{C}_1(\mathcal{K}, \mathcal{H}) \to \mathbb{C}$ is bounded with $\|f_T\| = \|T\|$ and every bounded linear functional on $\mathcal{C}_1(\mathcal{K}, \mathcal{H})$ arises in this manner. That is the dual space of $\mathcal{C}_1(\mathcal{K}, \mathcal{H})$ can be identified with $B(\mathcal{H}, \mathcal{K})$ in this manner.

However, not every bounded linear functional on $B(\mathcal{H}, \mathcal{K})$ is of the form f_R for some $R \in \mathcal{C}_1(\mathcal{K}, \mathcal{H})$.

(7) For each $R \in \mathcal{C}_1(\mathcal{K}, \mathcal{H})$ the linear functional $f_R : \mathbb{K}(\mathcal{H}, \mathcal{K}) \to \mathbb{C}$ defined by $f_R(K) = Tr(RK)$ is bounded with $\|f_R\| = \|R\|_1$ and every bounded linear functional on $\mathbb{K}(\mathcal{H}, \mathcal{K})$ is of this form. That is the dual space of $\mathbb{K}(\mathcal{H}, \mathcal{K})$ can be identified with $\mathcal{C}_1(\mathcal{K}, \mathcal{H})$ in this manner.

An operator $\rho \in B(\mathcal{H})$ is called a **density operator** provided that $\rho \in \mathcal{C}_1(\mathcal{H}),\ \rho \geq 0$ and $Tr(\rho) = 1$.

Proposition 3.2. *Every element of $\mathcal{C}_1(\mathcal{H})$ can be written as a linear combination of 4 density operators.*

Proof. We sketch the key ideas of this proof. First one shows that $T \in \mathcal{C}_1(\mathcal{H}) \implies T^* \in \mathcal{C}_1(\mathcal{H})$. From this it follows that $T = H + iK$ with $H = (T + T^*)/2 \in \mathcal{C}_1(\mathcal{H})$ and $K = (T - T^*)/2i \in \mathcal{C}_1(\mathcal{H})$. Next one shows that $H^+, H^-, K^+, K^- \in \mathcal{C}_1(\mathcal{H})$.

Finally, setting $\rho_1 = H^+/Tr(H^+)$, $\rho_2 = H^-/Tr(H^-)$, $\rho_3 = K^+/Tr(K^+)$, and $\rho_4 = K^-/Tr(K^-)$ defines the four density operators. □

3.2. **Tensor Products of Operators.** Let $R_i \in B(\mathcal{H}_i, \mathcal{K}_i), i = 1, 2$, then there exists a unique operator $R_1 \otimes R_2 \in B(\mathcal{H}_1 \otimes \mathcal{H}_2, \mathcal{K}_1 \otimes \mathcal{K}_2)$ satisfying
$$(R_1 \otimes R_2)(h_1 \otimes h_2) = (R_1 h_1) \otimes (R_2 h_2).$$
Moreover, $\|R_1 \otimes R_2\| = \|R_1\|\|R_2\|$.

In the case that $\mathcal{H} + \mathcal{H}_1 = \mathcal{K}_1$ and $\mathcal{K} = \mathcal{H}_2 = \mathcal{K}_2$, if either $dim(\mathcal{H})$ or $dim(\mathcal{K})$ is finite, then every element of $B(\mathcal{H} \otimes \mathcal{K})$ is a sum of such elementary tensors, but when they are both infinite dimensional this is not the case.

4. Basics of Quantum Viewpoint

4.1. **Postulates of Quantum Mechanics.** To each isolated physical system, there corresponds a Hilbert space $\mathcal{H}$, called the *state space*, and each unit vector in $\mathcal{H}$ represents a possible state, called the *state vector* or *pure state*.

Quantum Measurements. When we want to observe a system, i.e., connect to the "outside world", the system is no longer closed because we interact with it. By *closed*, we mean "not interacting with anything outside the system". By *open*, we mean it is a piece of a larger system.

Quantum measurements are always described by a class of operators $\{M_i\}_{i=\text{one of the outcomes}}$.
The probability that we observe the outcome i, given that the system is in state $|\psi\rangle$ before we measure, is given by $p_i = \|M_i\psi\|^2$ and if we observe the outcome i, then the system changes to the state $\frac{M_i\psi}{\|M_i\psi\|}$. Moreover, as the sum of the probabilities of all possible outcomes must equal 1, we have $\sum_i p_i = 1$.

Keeping in mind that quantum mechanics is inherently probabilistic, we consider a quantum experiment with at most k possible outcomes. Let $\mathcal{H}_s$ and $\mathcal{H}_o$ be Hilbert spaces representing the state space and the outcome space, respectively, and let $\{M_i \in \mathcal{B}(\mathcal{H}_s, \mathcal{H}_o) : 1 \leq i \leq k\}$ be a collection of bounded operators. If the system is in state $\psi \in \mathcal{H}_s, \|\psi\| = 1$ before we measure, then the probability that we observe the outcome i is given by $p_i = \|M_i\psi\|^2$ and if we observe the outcome i, then the system changes to the state $\frac{M_i\psi}{\|M_i\psi\|}$. Moreover, as the sum of the probabilities of all possible outcomes must equal 1, we have $\sum_i p_i = 1$. Hence,

$$1 = \sum_{i=1}^{k} p_i = \sum_{i=1}^{k} \|M_i\psi\|^2 = \sum_{i=1}^{k} \langle M_i\psi | M_i\psi\rangle = \sum_{i=1}^{k} \langle \psi | M_i^* M_i \psi\rangle .$$

Since the above equality holds for every $\psi \in \mathcal{H}$ with $\|\psi\| = 1$, the following lemma forces $\sum_{i=1}^{k} M_i^* M_i = I$. If $T \in \mathcal{B}(\mathcal{H})$, then $T = I \iff \langle \psi | T\psi\rangle = 1$ for every $\|\psi\| = 1$.

Theoretically, given any class of operators $\{M_i \in \mathcal{B}(\mathcal{H}_s, \mathcal{H}_o) : 1 \leq i \leq k\}$ such that $\sum_{i=1}^{k} M_i^* M_i = I$, there is a k-outcome quantum experiment with these measurement operators.

4.2. **Measurement Systems and Distinguishable States.** We include a bit more in the notes than we covered in class to help those who are new to this quantum viewpoint.

Definition 4.1. (Measurement System) Suppose that $\mathcal{H}$ and $\mathcal{K}$ are finite-dimensional Hilbert spaces. A finite family $\{M_i : 1 \leq i \leq k\}$ of operators $M_i : \mathcal{H} \to \mathcal{K}$ is called a *measurement system* if $\sum_i M_i^* M_i = I$. If $\mathcal{H} = \mathcal{K}$, we say that $\{M_i\}$ is a measurement system *on* $\mathcal{H}$.

Definition 4.2. (Perfectly Distinguishable States) A collection of states $\{\psi_1, \ldots, \psi_N\} \subseteq \mathcal{H}$ is called *perfectly distinguishable* if there exists a measurement system $\{M_i : 1 \leq i \leq k\}, k \geq N$ on $\mathcal{H}$ such that $\|M_i(\psi_j)\|^2 = \delta_{i,j}$ for $i, j \in \{1, ..., N\}$.

Theorem 4.3. *A collection of states $\{\psi_1, \ldots, \psi_N\} \subseteq \mathcal{H}$ is perfectly distinguishable if and only if $\psi_i \perp \psi_j$ for all $i \neq j$.*

Proof. ($\implies$) For the forward direction, let us assume that there is a measurement system $\{M_i : 1 \leq i \leq N\}$ such that $\|M_i(\psi_j)\| = \delta_{i,j}$ for $i, j \in \{1, ..., N\}$. Consider ψ_1 and ψ_2. ψ_2 can then be expressed as $\psi_2 = \alpha\psi_1 + \beta\eta$ where $\eta \perp \psi_1, \|\eta\| = 1$. Since $1 = \|\psi\|^2 = |\alpha|^2 + |\beta|^2$, we have $1 = \|M_2(\psi_2)\|^2 = \|M_2(\alpha\psi_1 + \beta\eta)\|^2 = |\beta|^2\|M_2(\eta)\|^2 \leq |\beta|^2\|\eta\|^2 = \|\beta\|^2 \leq 1$. This forces the above inequalities to be equalities so that $|\beta|^2 = 1$ which in turn implies that $\alpha = 0$ which means that ψ_2 and η are collinear and hence $\psi_2 \perp \psi_1$.

($\impliedby$) Let M_i be the (orthogonal) projection onto the one-dimensional subspace spanned by ψ_i. Then $M_i = M_i^* = M_i^* M_i$ for $i = 1, ..., N$ and $\sum_{i=1}^N M_i^* M_i$ is the orthogonal projection onto $\text{span}\{\psi_1, ..., \psi_N\}$. Let M_0 be the orthogonal projection onto $\{\psi_1, \ldots, \psi_N\}^\perp$. Then $\sum_{j=0}^N M_j^* M_j = \sum_{j=0}^N M_j = I$. Furthermore, $M_i(\psi_j) = \delta_{i,j}\psi_j$ for all $i, j \in \{1, ..., N\}$, so that $\|M_i(\phi_j)\|^2 = \delta_{i,j}$ for all $i, j \in \{1, ..., N\}$. This proves that $\{M_i\}_{i=0}^N$ is a measurement system. □

Corollary 4.4. *If* $\dim(\mathcal{H}_s) = N$, *then the system can have at most* N *perfectly distinguishable states.*

Theorem 4.5. *Suppose that* $\{\psi_1, \ldots, \psi_N\}$ *is a collection of linearly independent states. Then there exists a measurement system* $\{M_i : 0 \leq i \leq N\}$ *such that for* $i \neq 0$, $\|M_i(\psi_j)\| \neq 0$ *if and only if* $i = j$.

Proof. For $i = 1, ..., N$, let $V_i = \text{span}\{\psi_j : j \neq i\}$, and let E_i be the projection onto $V_i^\perp$. Then for $j \neq i$, $\psi_j \in V_i \implies E_i(\psi_j) = 0 \implies \|E_i(\psi_j)\|^2 = 0$. Now $0 \leq E_i \leq I \implies 0 \leq E_1 + \cdots + E_N \leq N \cdot I$. Let $M_i = \frac{1}{\sqrt{N}} E_i$ for $i = 1, ..., N$. Then $M_i^* M_i = \frac{1}{N} E_i$, so $\sum_{i=1}^N M_i^* M_i = \frac{1}{N}\sum_{i=1}^N E_i \leq I$, and hence $I - \sum_{i=1}^N M_i^* M_i \geq 0$. Now let $M_0 = (I - \sum_{i=1}^N M_i^* M_i)^{\frac{1}{2}}$. Then $\sum_{i=0}^N M_i^* M_i = \left((I - \sum_{i=1}^N M_i^* M_i)^{\frac{1}{2}}\right)^2 + \sum_{i=1}^N M_i^* M_i = I$, so $\{M_i\}_{i=0}^N$ is a measurement system. For $i \neq 0$, if $j \neq i$, then $\|M_i(\psi_j)\| = \frac{1}{\sqrt{N}}\|E_i(\psi_j)\| = 0$. Therefore by contrapositive, $\|M_i(\psi_j)\| \neq 0$ implies that $i = j$. Conversely, $\|M_i(\psi_i)\| = \frac{1}{\sqrt{N}}\|E_i(\psi_i)\| \neq 0$ since $\psi_i \notin V_i$ and so it has non-zero projection onto $V_i^\perp$. □

So far we have talked about pure states, now we will talk about ensembles (or mixed states).

4.3. **Ensembles or Mixed States.** As motivation for this topic, let $\{M_i : 1 \leq i \leq k\}$ be a measurement system with $M_i : \mathcal{H}_s \longrightarrow \mathcal{H}_o$. Suppose we have the state $\psi \in \mathcal{H}_s$ as input. Recall that $p_i = \|M_i(\psi)\|^2$ should be interpreted as the probability of observing the outcome i, and that if we do observe i, the system is now in the state, $\frac{M_i(\psi)}{\|M_i(\psi)\|}$. That is,

input:$\psi \in \mathcal{H}_s$; output: $\frac{M_i(\psi)}{\|M_i(\psi)\|}$ with probability $p_i = \|M_i(\psi)\|^2$.

So after observation, we will have what now looks like a mixed bag of states $\left\{\frac{M_i(\psi)}{\|M_i(\psi)\|}\right\}_i$, with $\frac{M_i(\psi)}{\|M_i(\psi)\|}$ occurring with probability p_i.

Definition 4.6. An *ensemble of states*, or a *mixed state*, is a finite collection $\{\psi_i, p_i : 1 \leq i \leq N\}$ of states ψ_i with probabilities p_i where $\|\psi_i\| = 1$, $p_i \geq 0$ and $\sum_{i=1}^{N} p_i = 1$.

Suppose we have a measurement system $\{M_i : 1 \leq i \leq N\}$ and an ensemble of states $\{\psi_j, p_j : 1 \leq j \leq k\}$ with $\sum_{j=1}^{k} p_j = 1$, then what is the probability of observing the outcome i?
If ψ_j is our input, then the probability getting outcome i is $\|M_i(\psi_j)\|^2$. So, the probability that we have outcome i is,

$$\sum_{j=1}^{k} p_j \|M_i(\psi_j)\|^2.$$

In the next subsection we discuss a better way to compute the probabilities of outcomes.

4.4. **Von Neumann's Notation: Density Matrices.** For a given state $\psi \in \mathcal{H}_s, \|\psi\| = 1$, a typical unit vector in the one-dimensional subspace spanned by ψ is given by $e^{i\theta}\psi$. In general $e^{i\theta}\psi \neq \psi$ but for any measurement M_j, we can see that $\|M_j(\psi)\|^2 = \|M_j(e^{i\theta}\psi)\|^2$. This shows that measurements don't distinguish between different unit vectors from the one-dimensional subspace spanned by the given state vector ψ and hence states should really refer to one-dimensional subspace and not just a unit vector. This means that *the probabilities of outcomes really depend on the one-dimensional subspace generated by a vector.*

Replacing states by rank one projections and lengths by trace: Recall that given a matrix $A = (a_{ij}) \in M_n$, the *trace* of that matrix is the sum of the diagonal entries: $Tr(Y) = \sum_i a_{ii}$. It is a popular fact that given any two square matrices A and B of the same size, $Tr(AB) = Tr(BA)$. The next proposition establishes this fact for compatible non-square matrices as well. Next, if $\psi \in \mathbb{C}^n, \|\psi\| = 1$, and P_ψ denotes the orthogonal projection onto the subspace spanned by ψ, then $P_\psi = \psi\psi^* = |\psi\rangle\langle\psi|$. ($P_\psi h = \psi\psi^* h = \langle\psi|h\rangle\psi$ where $\langle\psi|h\rangle$ is the component of h in the direction of ψ.) Furthermore,

$$Tr(P_\psi) = Tr(\psi\psi^*) = Tr(\psi^*\psi) = (\psi^*\psi) = \langle\psi|\psi\rangle = 1.$$

Back to Ensemble: Let's get back to the situation where we had a measurement system $\{M_i : 1 \leq i \leq N\}$ and an ensemble of states $\{\psi_j, p_j : 1 \leq j \leq k\}$ with $\sum_{j=1}^{k} p_j = 1$. We know that the probability of observing the outcome i is,

$$\sum_{j=1}^{k} p_j \|M_i(\psi_j)\|^2.$$

Simplifying this expression, we get

$$\begin{aligned}\sum_{j=1}^{k} p_j \|M_i(\psi_j)\|^2 &= \sum_{j=1}^{k} p_j (M_i\psi_j)^*(M_i\psi_j) \\ &= \sum_{j=1}^{k} p_j Tr((M_i\psi_j)^*(M_i\psi_j)) \\ &= \sum_{j=1}^{k} p_j Tr((M_i\psi_j)(M_i\psi_j)^*) \\ &= \sum_{j=1}^{k} p_j Tr(M_i\psi_j\psi_j^* M_i^*) \\ &= \sum_{j=1}^{k} p_j Tr(M_i^* M_i\psi_j\psi_j^*) \\ &= \sum_{j=1}^{k} Tr\left(M_i^* M_i(p_j\psi_j\psi_j^*)\right) \\ &= Tr\left(M_i^* M_i\left(\sum_{j=1}^{k} p_j\psi_j\psi_j^*\right)\right).\end{aligned}$$

Note that $\psi_j\psi_j^* = P_{\psi_j}$. If we set $P = \sum_{j=1}^{k} p_j\psi_j\psi_j^*$, then we have shown that:

Theorem 4.7. *Given an ensemble of states $\{\psi_j, p_j : 1 \leq j \leq k\}$ and a measurement system $\{M_i : 1 \leq i \leq N\}$, the probability of observing the i-th outcome is $Tr\,(M_i^* M_i P)$ where $P = \sum_{j=1}^{k} p_j\psi_j\psi_j^*$.*

The operator $P = \sum_{j=1}^{k} p_j\psi_j\psi_j^*$ associated to an ensemble of states is called the **Von Neumann density operator of the given ensemble**.

We observe that:

(1) If two ensembles have the same density matrix, then we get the same probability for outcomes for any measurement system.
(2) If $\{M_i : 1 \leq i \leq k\}$ and $\{\tilde{M}_i : 1 \leq i \leq k\}$ are two measurement systems such that for every i, $M_i^* M_i = \tilde{M}_i^* \tilde{M}_i$, then also we get the same probability for outcomes for any ensemble.

The following example illustrates the first observation.

Example 4.8. If $\{u_1, ..., u_N\}$ is an orthonormal basis for $\mathbb{C}^N$, then the density matrix P for the ensemble $\{u_j, \frac{1}{N} : 1 \leq j \leq N\}$ is given by $P = \sum_{j=1}^{N} \frac{1}{N} u_j u_j^* = \frac{1}{N} I_N$. If $\{\tilde{u}_1, ..., \tilde{u}_N\}$ is another orthonormal basis for $\mathbb{C}^N$, then the density matrix $\tilde{P}$ for the ensemble $\{\tilde{u}_j, \frac{1}{N} : 1 \leq j \leq N\}$ also turns

out to be $\tilde{P} = \sum_{j=1}^{N} \frac{1}{N} \tilde{u}_j \tilde{u}_j^* = \frac{1}{N} I_N$. This example guarantees the existence of two different ensembles with same density matrix.

Problem 4.9. *Fix $N \geq 3$ and let $u_j = \begin{pmatrix} \cos(\frac{2\pi j}{N}) \\ \sin(\frac{2\pi j}{N}) \end{pmatrix} \in \mathbb{C}^2$. Prove that the density matrix for the ensemble $\{u_j, \frac{1}{N} : 1 \leq j \leq N\}$ is given by $\frac{1}{2} I_2$.*

The above example shows that the density matrix does not distinguish between standard orthonormal basis or any other orthonormal basis as input. So, for computing probabilities, it is the density matrix which is important and not the ensemble.

At this point, let us pause for a while and try to visualise quantum experiments in terms of density matrices. Recall that, if a system is initially in the state ψ, that is, $\psi \in \mathcal{H}_s, \|\psi\| = 1$, and if there is given a measurement system $\{M_i : 1 \leq i \leq k\}$, then after measurement, the system becomes the ensemble $\left\{ \frac{M_i(\psi)}{\|M_i(\psi)\|}, \|M_i\psi\|^2 : 1 \leq i \leq k \right\}$. By associating density matrices with the states of the system before and after the measurement we note that the input is the state ψ and the density matrix corresponding to it is given by $P = \psi\psi^*$. After the measurement, the system becomes the ensemble $\left\{ \frac{M_i(\psi)}{\|M_i(\psi)\|}, \|M_i\psi\|^2 : 1 \leq i \leq k \right\}$, and hence the output is this ensemble which is identified by the density matrix

$$\begin{aligned} &\sum_{i=1}^{k} \|M_i\psi\|^2 \left(\frac{M_i(\psi)}{\|M_i(\psi)\|} \right) \left(\frac{M_i(\psi)}{\|M_i(\psi)\|} \right)^* \\ &= \sum_{i=1}^{k} (M_i\psi)(M_i\psi)^* = \sum_{i=1}^{k} (M_i\psi)(\psi^* M_i^*) \\ &= \sum_{i=1}^{k} M_i(\psi\psi^*)M_i^* = \sum_{i=1}^{k} M_i P M_i^*. \end{aligned}$$

Thus, in terms of density matrices, we observed that if input is identified by the density matrix P, then after measurement, the output is identified by the density matrix $\sum_{i=1}^{k} M_i P M_i^*$. This observation is the key to our next theorem.

Theorem 4.10. *Given an ensemble of states $\{\psi_j, p_j : 1 \leq j \leq J\}$ and a measurement system $\{M_i : 1 \leq i \leq k\}$ on $\mathcal{H}_s$ with density matrix $P = \sum_{j=1}^{J} p_j \psi_j \psi_j^*$, then after measurement, the system becomes the ensemble $\left\{ \frac{M_i(\psi_j)}{\|M_i(\psi_j)\|}, p_j \|M_i\psi_j\|^2 : 1 \leq i \leq k, 1 \leq j \leq J \right\}$ with density matrix $\sum_{i=1}^{k} M_i P M_i^*$.*

Proof. The density matrix for the output ensemble is given by

$$\sum_{j=1}^{J}\sum_{i=1}^{k} p_j||M_i\psi_j||^2 \left(\frac{M_i\psi_j}{||M_i\psi_j||}\right)\left(\frac{M_i\psi_j}{||M_i\psi_j||}\right)^* = \sum_{j=1}^{J}\sum_{i=1}^{k} p_j(M_i\psi_j)(M_i\psi_j)^*$$
$$= \sum_{i=1}^{k}\sum_{j=1}^{J} M_i(\psi_j p_j \psi_j^*)M_i^*$$
$$= \sum_{n=1}^{N} M_i P M_i^*. \qquad \square$$

So, a measurement system takes density matrix as input and yields another density matrix as output.

4.5. **Axiomatization of Quantum Channels.** We are now in a position to axiomatize *quantum channels*.

(1) A quantum channel should be a linear map,

$$\Phi : \mathcal{C}_1(\mathcal{H}_i) \to \mathcal{C}_1(\mathcal{H}_o).$$

(2) If $\rho \in \mathcal{C}_1(\mathcal{H}_i)$ is a density operator, then $\Phi(\rho) \in \mathcal{C}_1(\mathcal{H}_o)$ is a density operator.

The next axiom has to do with how quantum systems combine. Suppose we have two laboratories A and B (for Alice and Bob respectively). We will denote by $\mathcal{H}_{s,A}, \mathcal{H}_{s,B}, \mathcal{H}_{o,A}, \mathcal{H}_{o,B}$, respectively, the state space of lab A, the state space of lab B, the outcome space of lab A, and the outcome space of lab B.

Suppose that each lab has a measurement system. Let $\{M_i : \mathcal{H}_{s,A} \longrightarrow H_{o,A}\}_{i=1}^{K}$ be the measurement system of A and $\{N_j : \mathcal{H}_{s,B} \longrightarrow H_{o,B}\}_{j=1}^{J}$ be the measurement system of B. These define quantum channels,

$$\Phi_A(\rho_A) = \sum_{i=1}^{K} M_i\rho_A M_i^* \ \Phi_B(\rho_B) = \sum_{j=1}^{J} N_j\rho_B N_j^*.$$

If we wish to view these two labs as one single lab, say lab AB, then the state space of this lab is $\mathcal{H}_{s,AB} = \mathcal{H}_{s,A}\otimes\mathcal{H}_{s,B}$ and the output space would be $\mathcal{H}_{o,AB} = \mathcal{H}_{o,A} \otimes \mathcal{H}_{o,B}$ with measurement operators $\{M_i \otimes N_j : \mathcal{H}_{s,AB} \longrightarrow H_{o,AB}\}$, so that there are KJ outcomes. Note that $\sum_{i,j}(M_i \otimes N_j)^*(M_i \otimes N_j) = I$. This measurement system of lab AB, then, defines a quantum channel $\Phi_{AB} : \mathcal{C}_1(\mathcal{H}_{s,AB}) \longrightarrow \mathcal{C}_1(\mathcal{H}_{o,AB})$ given by

$$\Phi_{AB}(W) = \sum_{i,j}(M_i \otimes N_j)W(M_i \otimes N_j)^*.$$

This motivates the next axiom.

(3) Given quantum channels, $\Phi_A : \mathcal{C}_1(\mathcal{H}_{A,i}) \to \mathcal{C}_1(\mathcal{H}_{A,o})$ and $|Phi_B : \mathcal{C}_1(\mathcal{H}_{B,i}) \to \mathcal{C}_1(\mathcal{H}_{B,o})$ there should exist a quantum channel

$$\Phi_{AB} : \mathcal{C}_1(\mathcal{H}_{A,i} \otimes \mathcal{H}_{B,i}) \to \mathcal{C}_1(\mathcal{H}_{A,o} \otimes \mathcal{H}_{B,o})$$

satisfying $\Phi_{AB}(\rho_A \otimes \rho_B) = \Phi_A(\rho_A) \otimes \Phi_B(\rho_B)$.

Finally, doing nothing should be a quantum channel:

(4) Given any Hilbert space, the identity map from $id : \mathcal{C}_1(\mathcal{H}) \to \mathcal{C}_1(\mathcal{H})$ is a quantum channel.

Since every positive operator in $\mathcal{C}_1(\mathcal{H})$ is a positive multiple of a density operator, the first two axioms imply that a quantum channel must send positive operators to positive operators, such a map is called a **positive map**. The fact that density operators span $\mathcal{C}_1(\mathcal{H})$ together with the fact that density operators are mapped to density operators implies that a quantum channel must preserve traces, i.e.,

$$Tr(\Phi(W)) = Tr(W).$$

We will see that axioms 3 and 4 imply that a quantum channel must be "completely" positive. A concept that we need to first discuss.

5. Matrix Norm, Matrix Order, and Operator Matrices

Suppose that $T : V_1 \longrightarrow W_1$ and $R : V_2 \longrightarrow W_2$ are linear maps between vector spaces, then there is a linear map $T \otimes R : V_1 \otimes V_2 \longrightarrow W_1 \otimes W_2$ defined by $(T \otimes R)(v_1 \otimes v_2) = T(v_1) \otimes R(v_2)$ for all $v_1 \in V_1$ and $v_2 \in V_2$. If $\mathcal{H}$ and $\mathcal{K}$ are finite-dimensional Hilbert spaces with $X : \mathcal{H} \to \mathcal{H}$ and $Y : \mathcal{K} \to \mathcal{K}$, linear. Then there is a well-defined linear map denoted $X \otimes Y : \mathcal{H} \otimes \mathcal{K} \to \mathcal{H} \otimes \mathcal{K}$ satisfying $(X \otimes Y)(h \otimes k) = X(h) \otimes Y(k)$.

If $T : \mathcal{H} \longrightarrow \mathcal{H}$, possibly infinite dimensional, and $R : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ are linear, then we can define $T \otimes R : \mathcal{H} \otimes \mathbb{C}^n \longrightarrow \mathcal{H} \otimes \mathbb{C}^n$, in a similar way. Our goal in this subsection is to find a matrix representation for the map $T \otimes R$ in this setting. To do this, let us first address the following question:

I. What is a natural identification of a typical element of $\mathcal{H} \otimes \mathbb{C}^n$?

Recall that if we take the canonical orthonormal basis $\{e_1, ..., e_n\}$ for $\mathbb{C}^n$, then every vector $u \in \mathcal{H} \otimes \mathbb{C}^n$ has a unique representation given by $u = \sum_{i=1}^n h_i \otimes e_i$ where $h_i \in \mathcal{H}$, and

$$\|u\|^2 = \left\langle \sum_{i=1}^n h_i \otimes e_i | \sum_{j=1}^n h_j \otimes e_j \right\rangle = \sum_{i,j=1}^n \langle h_i | h_j \rangle \langle e_i | e_j \rangle = \sum_{i=1}^n \|h_i\|^2 = \|(h_1, ..., h_n)\|^2.$$

In other words, we have the Hilbert space isomorphism

$$\mathcal{H} \otimes \mathbb{C}^n \simeq \underbrace{\mathcal{H} \oplus \cdots \oplus \mathcal{H}}_{n \text{ times}} = \mathcal{H}^{(n)}$$

via the natural identification $\sum_{i=1}^n (h_i \otimes e_i) \simeq \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$.
The next question which we want to address is:

II. What is a natural identification of a linear map in $B(\mathcal{H} \otimes \mathbb{C}^n)$?
Given $A_{ij} \in B(\mathcal{H})$ for $1 \leq i, j \leq n$, we can consider $A = (A_{ij}) \in M_n(B(\mathcal{H}))$ as an operator defined by

$$A \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n A_{1j} h_j \\ \vdots \\ \sum_{j=1}^n A_{nj} h_j \end{pmatrix} \in \underbrace{\mathcal{H} \oplus \cdots \oplus \mathcal{H}}_{n \text{ times}}.$$

It is not hard to see that every such map is bounded. Therefore, we have $M_n(B(\mathcal{H})) \hookrightarrow B(\mathcal{H} \otimes \mathbb{C}^n)$ in a natural way. In fact, every linear map on $\mathcal{H} \otimes \mathbb{C}^n$ has such a matrix representation. The proof is "grubby" but here is the idea: If $A : \mathcal{H} \oplus \cdots \oplus \mathcal{H} \to \mathcal{H} \oplus \cdots \oplus \mathcal{H}$ is linear, then $A \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}$.
The map $\begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \mapsto k_1$ is linear. Similarly, mapping the column vector to k_2 is linear, and so on and so forth. The map $\begin{pmatrix} h_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \mapsto k_1$ is linear, so there is $A_{11} : \mathcal{H} \to \mathcal{H}$ enacting this transformation. If we continue to do this for every h_i and k_j, then we get linear maps $A_{ij} : \mathcal{H} \to \mathcal{H}$ and one can check that $A = (A_{ij})$.
Hence, we have a natural identification $B(\mathcal{H} \otimes \mathbb{C}^n) \simeq M_n(B(\mathcal{H}))$ via $A \simeq (A_{ij})$, thereby allowing us to identify any linear operator $A \in B(\mathcal{H} \otimes \mathbb{C}^n)$ by an $n \times n$ block matrix $(A_{ij}) \in M_n(B(\mathcal{H}))$ whose entries are given by linear maps.

This means, in particular, that when we write down a matrix of operators, then $(A_{i,j})$ has a well-defined norm, namely, its norm as an operator on $\mathcal{H}^{(n)}$ and we can say if it defines a positive operator or not. This is what is meant by the natural **matrix norm** and **matrix order** on $M_n(B(\mathcal{H}))$.

III. Matrix Representation of $T \otimes R$: Suppose that $T : \mathcal{H} \longrightarrow \mathcal{H}$ and $R : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ are linear, $R \in M_n(\mathbb{C}), R = (r_{ij})$, then $T \otimes R : \mathcal{H} \otimes \mathbb{C}^n \longrightarrow \mathcal{H} \otimes \mathbb{C}^n$ has a natural representation as an $n \times n$ block matrix $T \otimes R \in M_n(\mathcal{L}(\mathcal{H}))$ whose entries are given by linear maps.

We know that $(T \otimes R)(h \otimes y) = T(h) \otimes R(y)$, therefore,

$$(T \otimes R)(h \otimes e_j) = T(h) \otimes R(e_j) = T(h) \otimes \left(\sum_{i=1}^{n} r_{ij} e_i \right)$$

$$= \sum_{i=1}^{n} r_{ij} T(h) \otimes e_i \simeq \begin{pmatrix} r_{1j}Th \\ \vdots \\ r_{nj}Th \end{pmatrix} = (r_{ij}T) \begin{pmatrix} 0 \\ \vdots \\ h \\ \vdots \\ 0 \end{pmatrix},$$

where h is in the j-th position and there are 0's everywhere else in the column vector. The ***Kronecker product*** of T and R, then, is the block matrix in $M_n(B(\mathcal{H}))$ given by $(r_{ij}T)$ (so, there are n blocks, each block is of size equal to the dimension of $\mathcal{H}$, and the (i,j)-block is $r_{ij}T$). In other words, the Kronecker product is equal to the tensor product of the linear maps (with respect to the canonical basis for $\mathbb{C}^n$).

A special case is when $R = I_n$ then we have that

$$T \otimes I_n = \begin{pmatrix} T & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & & T \end{pmatrix}.$$

If $T \in M_k$ and $R \in M_n$. Then $T \otimes R$ has matrix representation given by $T \otimes R = \begin{pmatrix} r_{11}T & \cdots & r_{1n}T \\ \vdots & \ddots & \vdots \\ r_{n1}T & \cdots & r_{nn}T \end{pmatrix}$, a block matrix with n blocks, each of size k.

Another way to view operator matrices is as sums of tensors. If we set

$$E_{i,j} = |e_i\rangle \langle e_j|,$$

then

$$(A_{i,j}) = \sum_{i,j=1}^{k} A_{i,j} \otimes E_{i,j} \in B(\mathcal{H}) \otimes M_k.$$

Given subspaces, $V \subseteq B(\mathcal{H})$ and $W \subseteq B(\mathcal{K})$ we can regard $M_k(V) \subseteq M_k(B(\mathcal{H}))$ and $M_k(W) \subseteq M_k(B(\mathcal{K}))$. This means that these subspaces are also endowed with a canonical matrix norm and matrix order, via these inclusions.

Given a linear map $\Phi : V \to W$ we get linear maps, $\Phi^{(k)} : M_k(V) to M_k(W)$ via

$$\Phi^{(k)}((A_{i,j})) = (\Phi(A_{i,j})).$$

We say that Φ is **k-positive** if $\Phi^{(k)}$ maps positive elements of $M_k(V)$ to positive elements of $M_k(W)$. We say that Φ is **completely positive** if it is k-positive for all k.

Similarly, each map $\Phi^{(k)}$ has a norm, but it turns out that these can vary with k. So we call Φ **completely bounded** provided that

$$\|\Phi\|_{cb} := \sup\{\|\Phi^{(k)}\|; k \in \mathbb{N}\} < +\infty.$$

Here is one example. Let $V = W = B(\mathbb{C}^2)$ and define $\Phi(X) = X^t$, the transpose. It is a linear map. It is easy to check that $P \geq 0 \iff P^t \geq 0$, os it is a positive map. Also, $\|X\| + \|X^t\|$ is easily chekced. So Φ is an isometric map.

Now consider the "matrix of matrix units",

$$Q = \begin{pmatrix} E_{1,1} & E_{1,2} \\ E_{2,1} & E_{2,2} \end{pmatrix} \in M_2(B(\mathbb{C}^2)) = B(\mathbb{C}^4).$$

Since $Q = Q^*$ and $Q^2 = 2Q$ we see that the spectrum of Q is $\{0, 2\}$ and so $Q \geq 0$. But

$$\Phi^{(2)}(Q) = \begin{pmatrix} E_{1,1} & E_{2,1} \\ E_{1,2} & E_{2,2} \end{pmatrix} := R.$$

We have that $det(R) = -1$, so it has negative eigenvalues. Hence, R is not positive and Φ is not 2-positive and so not completely positive.

Also, $R^2 = I$ and so $\|R\| = 1$ and $\Phi^{(2)}(R) = Q$ which has norm 2. Thus, $\|Phi^{(2)}\| \geq 2$. In fact, $\|\Phi\|_{cb} = 2$. So this example shows that in general $\|Phi\| \neq \|\Phi\|_{cb}$.

If one considers the transpose map on M_n one can show that it has norm one and cb-norm of n. Thus, the cb-norm of a map can be arbitrarily larger than its norm. In fact, if we consider the transpose map on $B(\ell^2_{\mathbb{N}})$ it is an isometric map with infinite cb-norm.

6. Introduction to C*-algebras

Developments and proofs of many of the results stated in this section can be found in [**?**, **?**, **?**]. Recall that $\mathcal{A}$ is an **algebra** if it is a vector space and also has a product that satisfies:

- $(AB)C = A(BC)$
- $(A + B)C = AC + BC, C(A + B) = CA + CB$,
- $\lambda \in \mathbb{C}, A, B \in \mathcal{A} \implies \lambda(AB) = (\lambda A)B = A(\lambda B)$.

An algebra is called a **Banach algebra** if it has a norm, it is complete in the norm, i.e., a Banach space, and the norm is submultiplicative:

$$\|AB\| \leq \|A\|\|B\|.$$

An algebra is a ***-algebra** if it also has a map $A \to A^*$ satisfying

- $(A^*)^* = A$,
- $(A + B)^* = A^* + B^*$,
- $\lambda \in \mathbb{C}, A \in \mathcal{A} \implies (\lambda A)^* = \overline{\lambda} A^*$,

- $(AB)^* = B^*A^*$.

These properties are reflecting the behaviour of the adjoint of Hilbert space operators.

A *-algebra is a **C*-algebra** if the norm also satisfies

$$\|A^*\| = \|A\| \text{ and } \|A\|^2 = \|A^*A\|.$$

We call $\mathcal{A}$ a **unital C*-algebra** if it also has a unit element, I. In the case one can show that necessarily, $I^* = I$ and $\|I\| = 1$.

The axioms are set up so that any norm closed subalgebra $\mathcal{A} \subseteq B(\mathcal{H})$ such that $A \in \mathcal{A} \implies A^* \in \mathcal{A}$ is a C*-algebra. We will call these **concrete C*-algebras.**

One key theorem is that every abstract C*-algebra is "identical" to a concrete C*-algebra, where means *-isomorphic, a concept that we will define shortly.

Here are some non-concrete C*-algebras. Let X be a compact Hausdorff space and set

$$C(X) = \{f : X \to \mathbb{C} | \text{ f is continuous } \},$$

and set

$$\|f\| = \sup\{|f(x)| : x \in X\},$$

which is finite since X is compact. Define a *-operation by

$$f^*(x) = \overline{f(x)}.$$

Then it is not hard to see that this is a C*-algebra.

Here are a few basic facts about C*algebras.

Cartesian Decomposition: Given $A \in \mathcal{A}$ we have that $H = \frac{A+A^*}{2} = H^* \in \mathcal{A}$ and $K = \frac{A-A^*}{2i} = K^* \in \mathcal{A}$ and $A = H + iK$.

Spectrum: Given a unital C*-algebra $\mathcal{A}$ and $A \in \mathcal{A}$ we set

$$\sigma_{\mathcal{A}}(A) = \{\lambda \in \mathbb{C} | \ (\lambda I - A) \text{ has no inverse in } \mathcal{A}\}.$$

Then $\sigma_{\mathcal{A}}(A)$ is a non-empty compact set and we have

$$\sup\{|\lambda| : \lambda \in \sigma_{\mathcal{A}}(A)\} = \lim_n \|A^n\|^{1/n}.$$

Spectral Permanence: If $\mathcal{A}$ is a C*-subalgebra of $\mathcal{B}$ with $I \in \mathcal{A} \subseteq \mathcal{B}$, then for any $X \in \mathcal{A}$, $\sigma_{\mathcal{A}}(X) = \sigma_{\mathcal{B}}(X)$.

- if $H = H^*$, then $\sigma_{\mathcal{A}}(H) \subseteq \mathbb{R}$.
- If $U^*U = UU^* = I$, then $\sigma_{\mathcal{A}}(U) \subseteq \mathbb{T}$.
- If $P = P^*$, then $P = A^*A$ for some A if and only if $\sigma_{\mathcal{A}}(P) \subseteq [0, +\infty)$.

This last property is used to define the **positive** elements of a C*-algebra.

Given two C*-algebras, $\mathcal{A}, \mathcal{B}$ a map $\pi : \mathcal{A} \to \mathcal{B}$ is called a ***-homomorphism** provided:

- π is linear,
- $\pi(XY) = \pi(X)\pi(Y)$, i.e., is multiplicative,
- $\pi(X^*) = \pi(X)^*$.

We call π a ***-isomorphism** if in addition it is one-to-one and onto.

Proposition 6.1. *If π is a *-homomorphism, then $\|\pi(X)\| \leq \|X\|$ and the range of π, $\mathcal{R}(\pi)$ is closed. Consequently, if π is a *-isomorphism, then π is an isometry.*

Corollary 6.2 (Uniqueness of Norm)**.** *Let $\mathcal{A}$ be a *-algebra and suppose that $\|\cdot\|_1$, $\|\cdot\|_2$ are two norms, both of which make $\mathcal{A}$ into a C*-algebra. Then $\|X\|_1 = \|X\|_2$, $\forall X \in \mathcal{A}$.*

The following theorem characterizes all **abelian,** i.e., $X, Y \in \mathcal{A} \implies XY = YX$, C*-algebras.

Theorem 6.3 (Gelfand-Naimark)**.** *Each unital abelian C*-algebra is *-isomorphic to $C(X)$ for some compact, Hausdorff space X.*

6.1. **States and the GNS Construction.** By a **state** on a unital C*-algebra $\mathcal{A}$ we mean a linear functional, $s : \mathcal{A} \to \mathbb{C}$ such that $s(I) = 1$ and $s(X^*X) \geq 0$, $\forall X \in \mathcal{A}$.

The following alternative characterization of states is often useful.

Proposition 6.4. *Let $s : \mathcal{A} \to \mathbb{C}$ be a linear functional with $s(I) = 1$. Then s is a state if and only if $\|s\| = 1$.*

Theorem 6.5 (The GNS Construction)**.** *Let $\mathcal{A}$ be a unital C*-algebra and let $s : \mathcal{A} \to \mathbb{C}$ be a state. Then there exists a Hilbert space $\mathcal{H}$, a unit vector $\phi \in \mathcal{H}$ and a unital *-homomorphism, $\pi : \mathcal{A} \to B(\mathcal{H})$ such that*

$$s(A) = \langle \phi | \pi(A)\phi \rangle, \ \forall A \in \mathcal{A}.$$

We outline the key ideas of the proof. First define a map

$$B : \mathcal{A} \times \mathcal{A} \to \mathbb{C} \text{ by } B(X, Y) = s(X^*Y).$$

It is easy to check that this map is sesquilinear and positive semidefinite. Thus, if we let

$$\mathcal{N} = \{X \in \mathcal{A} | B(X, X) = 0\},$$

then $\mathcal{N}$ is a subspace and we have a well-defined inner product on $\mathcal{A}/\mathcal{N}$ defined by

$$\langle X + \mathcal{N} | Y + \mathcal{N} \rangle = s(X^*Y).$$

We will get our HIlbert space $\mathcal{H}$ by completing this inner product space.

Next note that for each $A \in \mathcal{A}$ we have a linear map

$$L_A : \mathcal{A} \to \mathcal{A}, \ L_A(X) = AX,$$

given by left multiplication by the element A.

We claim that $L_A(\mathcal{N}) \subseteq \mathcal{N}$. To see this we first show that

$$0 \leq X^*A^*AX \leq \|A\|^2 X^*X.$$

Hence, if $X \in \mathcal{N}$ then

$$0 \leq s(X^*A^*AX) \leq \|A\|^2 s(X^*X) = 0.$$

This implies that $s((AX)^*(AX)) = s(X^*A^*AX) = 0$ and so $AX \in \mathcal{N}$ and the claim is done.

General algebra then tells us that we have a well-defined quotient map,

$$\widehat{L_A} : \mathcal{A}/\mathcal{N} \to \mathcal{A}/\mathcal{N},\ \widehat{L_A}(X+\mathcal{N}) = AX+\mathcal{N}.$$

The above inequality also tells us that this map is bounded on the inner product space $\mathcal{A}/\mathcal{N}$, since,

$$\|\widehat{L_A}(X+\mathcal{N})\|^2 = \langle AX+\mathcal{N}|AX+\mathcal{N}\rangle = s(X^*A^*AX) \leq \|A\|^2 s(X^*X) = \|A\|^2\|X+\mathcal{N}\|^2.$$

By HW1, we can extend this linear map by continuity to a bounded linear map, $\widetilde{L_A} : \mathcal{H} \to \mathcal{H}$ with $\|\widetilde{L_A}\| = \|\widehat{L_A}\|$.

Thus, we have a map, $\pi : \mathcal{A} \to B(\mathcal{H})$, $\pi(A) = \widetilde{L_A}$.

Some checking shows that the map π is a *-homomorphism.

To define the vector, we set $\phi = I + \mathcal{N}$. Then $\|\phi\|^2 = \langle\phi|\phi\rangle = s(I^*I) = s(I) = 1$.

Finally,

$$\langle\phi|\pi(A)\phi\rangle = \langle I+\mathcal{N}|A+\mathcal{N}\rangle = s(A).$$

This completes the outline of the proof.

This construction also leads to the following important theorem.

Theorem 6.6 (GNS Representation Theorem)**.** *Let $\mathcal{A}$ be a unital C*-algebra. Then there exists a Hilbert space and an isometric *-homomorphism, $\pi : \mathcal{A} \to B(\mathcal{H})$. Hence, $\mathcal{A}$ and the concrete C*-subalgebra $\mathcal{B} = \pi(\mathcal{A})$ are *-isomorphic.*

The idea of the proof is to for each state get a *-homomorphism and then take a direct sum of these *-homomorphisms and prove that there are enough states that this direct sum must be isometric.

6.2. **GNS and State Purification.** Suppose that we are given a density operator $\rho \in \mathcal{C}_1(\mathcal{H})$. This defines a linear functional,

$$s_\rho : B(\mathcal{H}) \to \mathbb{C} \text{ via } s_\rho(X) = Tr(X\rho).$$

Note that $s_\rho(I) = Tr(\rho) = 1$ and for any positive X^*X,

$$s_\rho(X^*X\rho) = Tr(X^*X\rho) = Tr(X\rho X^*) \geq 0,$$

since $X\rho X^* \geq 0$. Thus, s_ρ is a state and by GNS has a representation,

$$s_\rho(X) = \langle\phi|\pi(X)\phi\rangle.$$

In the case that $\rho = \sum_{i=1}^N \lambda_i |\phi_i\rangle\langle\phi_i|$ we can make this very explicit. Set $\phi = (\sqrt{\lambda_1}\phi_1, \ldots, \sqrt{\lambda_N}\phi_N) \in \mathcal{H}^{(N)}$, which is a unit vector, and let

$$\pi(X) = Diag(X) \in M_N(B(\mathcal{H})) = B(\mathcal{H}^{(N)}),$$

where by $Diag(X)$ we mean the diagonal operator matrix with X for the diagonal entry.

It is easily seen that $\pi : B(\mathcal{H}) \to M_N(B(\mathcal{H}))$ is a *-homomorphism and that $s_\rho(X) = \langle\phi|\pi(X)\phi\rangle$.

Thus, we have a very concrete GNS-like, in this case. This construction is generally referred to as **state purification**, as in the phrase, "by purifying

the state ensemble, we may regard it as a pure state on a larger HIlbert space". In this sense, the GNS representation shows that every state can be "purified".

Later we will talk about what it means for a state on a C*-algebra to be "pure". GNS does not say that every state is pure, just that it can be represented as a vector state, and we will see that vector states are pure states on $B(\mathcal{H})$.

A natural question is whether or not this concrete construction is the GNS, in an appropriate sense. The following result tells us how to recognize the GNS representation of a state.

Proposition 6.7. *Let $\mathcal{A}$ be a unital C*-algbra, let $s : \mathcal{A} \to \mathbb{C}$ be a state and let $\pi_s : \mathcal{A} \to B(\mathcal{H}_s)$, $\phi_s \in \mathcal{H}_s$ be the GNS representation of the state. Then*

$$\pi_s(\mathcal{A}) := \{\pi_s(A)\phi_s : A \in \mathcal{A}\}$$

*is a dense subset of $\mathcal{H}_s$. Moreover, let $\pi : \mathcal{A} \to B(\mathcal{H})$ be a unital *-homomorphims and let $\phi \in \mathcal{H}$ be a unit vector such that $\langle\phi|\pi(A)\phi\rangle = s(A)$ and such that $\pi(\mathcal{A})\phi$ is dense in $\mathcal{H}$, then there is a unitary $U : \mathcal{H}_s \to \mathcal{H}$ with $U\phi_s = \phi$ and $\pi(A) = U\pi_s(A)U^*$.*

Given a *-homomorphism π a vector ϕ is called **cyclic** if $\pi(\mathcal{A})\phi$ is dense in $\mathcal{H}$. Thus, the proposition says that, up to a unitary equivalence, any (π, ϕ) that gives rise to the state via the formula, $s(A) = \langle\phi|\pi(A)\phi\rangle$ with ϕ cyclic, is the GNS.

In the case that $\rho = \sum_{i=1}^N \lambda_i |\phi_i\rangle \langle\phi_i|$ considered above, with the representation $\pi(A) = Diag(A)$, the vector $\phi = (\sqrt{\lambda_1}\phi_1, ..., \sqrt{\lambda_N}\phi_N)$ might not be cyclic, so this might not be the GNS of the state on $\mathcal{A} = B(\mathcal{H})$. For example, if the vectors $\phi_1, ..., \phi_N$ are not linearly independent, then ϕ won't be cyclic.

However, if we use the spectral decomposition of ρ, then the vectors $\phi_1, ..., \phi_N$ will be orthonormal and in this case one can see that the vector ϕ is cyclic. This is because when the vectors are o.n., then given any set of vectors $h_1, ..., h_N$ we can always find an operator A such that $h_i = A(\sqrt{\lambda_i}\phi_i)$, $\forall i$.

6.3. **The C*-algebra $M_n(\mathcal{A})$.** Given a unital C*-algebra $\mathcal{A}$, we want to discuss how to make $M_n(\mathcal{A})$ into a C*-algebra. First note that it is always a vector space with operations, scalar multiplication $\lambda(A_{I,j}) = (\lambda A_{i,j})$ and addition $(A_{i,j}) + (B_{i,j}) = (A_{i,j} + B_{i,j})$. There is a natural way to make it an algebra too via the formula for matrix multiplication, $(A_{i,j}) \cdot (B_{i,j}) = (\sum_{k=1}^n A_{i,k}B_{k,j})$. If we set $(A_{i,j})^* = (A^*_{j,i})$ then we have a *-algebra. all that we are lacking to make it into a C*-algebra is a norm.

To find a norm, we use the GNS theorem. Take any $\pi : \mathcal{A} \to B(\mathcal{H})$ an isometric *-homomorphism, so that $\|A\| = \|\pi(A)\|$. We now define

$$\|(A_{i,j})\|_\pi := \|(\pi(A_{i,j})\|_{B(\mathcal{H}^{(n)})}.$$

It is easily checked that this norm makes $M_n(\mathcal{A})$ into a C*-algebra.

However, we have that the norm on a C*-algebra is unique, so any other way that we tried to create a norm, as long as it was a C*-norm, would necessarily be this norm.

Now that we know that every $M_n(\mathcal{A})$ is itself a C*-algebra, it makes sense to talk about completely positive maps between any two C*-algebras. Namely, if $\Phi : \mathcal{A} \to \mathcal{B}$ is a linear map, then we say that it is **n-positive** if whenever $(A_{i,j})$ is a positive element of the C*-algebra $M_n(\mathcal{A})$ then $(\Phi(A_{i,j}))$ is a positive element of the C*-algebra $M_n(\mathcal{B})$. As before a map is **completely positive** provided that it is n-positive for every n.

6.4. **Stinespring's dilation Theorem.**

Theorem 6.8 (Stinespring(1955))**.** *Let $\mathcal{A}$ be a unital C*-algebra, $\mathcal{H}$ a Hilbert space, and $\Phi : \mathcal{A} \to B(\mathcal{H})$ a completely positive map. Then there is a Hilbert space $\mathcal{K}$, a unital *-homomorphism $\pi : \mathcal{A} \to B(\mathcal{K})$ and $V \in B(\mathcal{H}, \mathcal{K})$ such that*

$$\Phi(A) = V^*\pi(A)V.$$

Moreover, every map of this form is completely positive.

For a complete proof see either [**?**] or [**?**].

We sketch the key ideas of the proof. First we take the vector space $\mathcal{A}\otimes\mathcal{H}$ and define a sesquilinear form by

$$B(\sum_i X_i \otimes h_i | \sum_j Y_j \otimes k_j) = \sum_{i,j} \langle h_i | \Phi(X_i^* Y_j) k_j.$$

One checks that this is positive semidefinite. To see, note that

$$(*) := B(\sum_{i=1}^N A_i \otimes h_i | \sum_{j=1}^N A_j \otimes h_j) = \langle h | (\Phi(A_i^* A_j)h\rangle,$$

where $h = (h_1, ..., h_n) \in \mathcal{H}^{(N)}$ and $(\Phi(A_i^*A_j)) = \Phi^{(N)}((A_i^*A_j)) \geq 0$, since Φ is N-positive and since $(A_i^*A_j) = X^*X \geq 0$, with X the matrix that has $A_1, ..., A_N$ for its first row and all other rows equal to 0. This shows that $(*) \geq 0$ and so B is positive semidefinite.

Let $\mathcal{N}$ be the null space of B. Our Hilbert space $\mathcal{K}$ will be the completion of the inner product space $(\mathcal{A} \otimes \mathcal{H})/\mathcal{N}$.

Now as in GNS for each $A \in \mathcal{A}$ we define a linear map $L_A : \mathcal{A}\otimes\mathcal{H} \to \mathcal{A}\otimes\mathcal{H}$ via $L_A(\sum_i X_i \otimes h_i) = \sum_i (AX_i)\otimes h_i$ and check that $L_A(\mathcal{N}) \subseteq \mathcal{N}$. This allows us to define a linear map on the quotient, $\widehat{L_A} : (\mathcal{A} \otimes \mathcal{H})/\mathcal{N} \to (\mathcal{A} \otimes \mathcal{H})/\mathcal{N}$ which we show is bounded and so extends to an operator, $\pi(A) : \mathcal{K} \to \mathcal{K}$. This defines our *-homomorphism.

To define $V : \mathcal{H} \to \mathcal{K}$ we set $V(h) = I_{\mathcal{A}} \otimes h + \mathcal{N}$ and check that it is linear and bounded.

Finally, to see that this gives us what we want we compute,

$$\begin{aligned}\langle h | V^*\pi(A)Vk\rangle_{\mathcal{H}} = \langle Vh | \pi(A)Vk\rangle_{\mathcal{K}} &= \langle I_{\mathcal{A}} \otimes h + \mathcal{N} | A \otimes k + \mathcal{N}\rangle_{\mathcal{K}} \\ &= B(I_{\mathcal{A}} \otimes h | A \otimes k) = \langle h | \Phi(A)k\rangle_{\mathcal{H}}.\end{aligned}$$

Since this is true for all pairs of vectors, we have that $V^*\pi(A)V = \Phi(A)$.

6.5. **More on Tensor Products.** Given $A = (a_{i,j}) : \mathbb{C}^n \to \mathbb{C}^n$ and $B = (b_{k,l}) : \mathbb{C}^d \to \mathbb{C}^d$ we have a linear map $A \otimes B : \mathbb{C}^n \otimes \mathbb{C}^d \to \mathbb{C}^n \otimes \mathbb{C}^d$. We would like to look at matrix representations of htis map. Recall that to write down a matrix for a linear map one wants an **ordered** basis for the space. If $\{e_i : 1 \leq i \leq n\}$ and $\{f_k : 1 \leq k \leq d\}$ are the canonical onb's, then we know that $\{e_i \otimes f_k : 1 \leq i \leq n, 1 \leq k \leq d\}$ is an orthonormal basis for $\mathbb{C}^n \otimes \mathbb{C}^d$.

There are two natural ways to order this basis, one is as

$$e_1 \otimes f_1, e_2 \otimes f_1, ..., e_n \otimes f_1, e_1 \otimes f_2, ..., e_n \otimes f_2, ..., e_n \otimes f_d,$$

when we group these into blocks of n, this corresponds to the decomposition

$$\mathbb{C}^n \otimes \mathbb{C}^d \sim (\mathbb{C}^n \otimes f_1) \oplus \cdots \oplus (\mathbb{C}^n \otimes f_d) \sim (\mathbb{C}^n)^{(d)}.$$

With respect to this ordering,

$$A \otimes B \sim (b_{k,l}A) \in M_d(M_n).$$

Alternatively, we may order the basis as,

$$e_1 \otimes f_1, e_1 \otimes f_2, ..., e_1 \otimes f_d, e_2 \otimes f_1, ..., e_2 \otimes f_d, ..., e_n \otimes f_d,$$

when we group these into blocks of size d this corresponds to the decomposition

$$\mathbb{C}^n \otimes \mathbb{C}^d \sim (e_1 \otimes \mathbb{C}^d) \oplus \cdots (e_n \otimes \mathbb{C}^d) \sim (\mathbb{C}^d)^{(n)}.$$

With repsect to this ordering,

$$A \otimes B \sim (a_{i,j}B) \in M_n(M_d).$$

In particular these two $(nd) \times (nd)$ matrices are unitarily equivalent by the permutation unitary that takes one ordering to the other.

When $B = I_d$ this gives us two matrix representations,

$$A \otimes I_d \sim Diag(A) \sim (a_{i,j}I_d).$$

6.6. **The Finite Dimensional Version of Stinespring.** We want to look at what Stinespring's theorem says in the case that $\mathcal{A} = M_d$ and $\mathcal{H} = \mathbb{C}^r$, so that we have a CP map $\Phi : M_d \to B(\mathbb{C}^r) = M_r$.

In this case the Hilbert space $\mathcal{K}$ is obtained by completing $(M_d \otimes \mathbb{C}^r)/\mathcal{N}$. But this space is finite dimensional and every finite dimensional inner product space is already complete, so that $\mathcal{K} = (M_d \otimes \mathbb{C}^r)/\mathcal{N}$ and in particular,

$$dim(\mathcal{K}) \leq d^2 r.$$

Now let $\{E_{i,j} = |e_i\rangle \langle e_j| : 1 \leq i, j \leq d\}$ be the canonical basis for M_d and let the Stinespring representation, be $\Phi(X) = V^*\pi(X)V$ with $V : \mathbb{C}^r \to \mathcal{K}$ and $\pi : M_d \to B(\mathcal{K})$. Because π is a unital *-homomorphism, it follows that $\pi(E_{i,i})$ is the orthogonal projection onto some subspace $\mathcal{M}_i$ of $\mathcal{K}$. Because

$$I_{\mathcal{K}} = \pi(I_d) = \sum_i \pi(E_{i,i}),$$

we see that

$$\mathcal{K} = \mathcal{M}_1 \oplus \mathcal{M}_2 \oplus \cdots \mathcal{M}_d.$$

Moreover, because $E^*_{i,j}E_{j,j} = E_{j,j}$ and $E_{i,j}E^*_{i,j} = E_{i,i}$ we see that $\pi(E_{i,j}$ defines an isometry from $\mathcal{M}_j$ onto $\mathcal{M}_i$. This guarantees that $dim(\mathcal{M}_i) = dim(\mathcal{M}_j)$ and if that if we use the maps $\pi(E_{i,j})$ to identify these as all the same space $\mathcal{M}$, then

$$\mathcal{K} = \mathcal{M}^{(d)},$$

and the maps $\pi(E_{i,j})$ just act as permutations of the j-th copy of $\mathcal{M}$ to the i-th copy.

We also have that, $d \cdot dim(\mathcal{M}) = dim(\mathcal{K}) \leq d^2 r$, so that

$$m := dim(\mathcal{M}) \leq dr.$$

With these identifications, for $X = (x_{i,j}) = \sum_{i,j=1}^{d} x_{i,j}E_{i,j}$, we have that

$$\pi(X) = (x_{i,j}I_{\mathcal{M}}).$$

But up to a permutation, we may also regard

$$\mathcal{K} \sim (\mathbb{C}^d)^{(m)},$$

and now

$$\pi(X) = Diag(X),$$

the block diagonal matrix of m copies of X, and now $V : \mathbb{C}^r \to \mathcal{K} = (\mathbb{C}^d)^{(m)}$ has the form

$$Vh = (V_1 h, \ldots, V_m h)^t,$$

for maps $V_i : \mathbb{C}^r \to \mathbb{C}^d$, i.e., $d \times r$ matrices.

With these identifications,

$$\Phi(X) = V^* Diag(X) V = \sum_{i=1}^{m} V_i^* X V_i.$$

This last form of Φ is often called a **Choi-Kraus** representation of Φ.

Note that our proof shows that the Choi-Kraus representation can always be taken to have fewer than dr terms.

A few things to note. If Φ is UCP, then

$$I_r = \Phi(I_d) = \sum_{i=1}^{m} V_i^* V_i.$$

On the other hand if Φ is CPTP, then

$$Tr(X) = Tr(\Phi(X)) = Tr(\sum_{i=1}^{m} V_i^* X V_i) = Tr((\sum_{i=1}^{m} V_i V_i^*)X), \forall X,$$

from which it follows that

$$\sum_{i=1}^{m} V_i V_i^* = I_d.$$

Thus, we see that every CPTP $\Phi : B(\mathbb{C}^d) = \mathcal{C}_1(\mathbb{C}^d) \to B(\mathbb{C}^r) = \mathcal{C}_1(\mathbb{C}^r)$ corresponds to the quantum channel induced by an m outcome measurement system,

$$\{V_1^*, ..., V_m^*\},$$

between the initial space $\mathbb{C}^d$ and the final space $\mathbb{C}^r$.

Perhaps the key takeaway of this subsection is the following.

Corollary 6.9. *Every CPTP map $\Phi : M_d \to M_r$ is the quantum channel induced by an m-outcome measurement system, $\{V_1^*, ..., V_m^* : \mathbb{C}^d \to \mathbb{C}^r$.*

Thus, in finite dimensions the set of CPTP maps and the set of quantum channels induced by measurement systems coincide.

6.7. **K Outcome POVM's.** Recall that a K outcome measurement system on an input state space $\mathcal{H}_i$ is given by an output Hilbert space $\mathcal{H}_o$ and operators $T_k : \mathcal{H}_i \to \mathcal{H}_o$ such that if we are in state ψ then the probability of observing outcome k is

$$p_k = \langle T_k\psi | T_k\psi\rangle = \langle \psi | T_k^* T_k \psi\rangle.$$

The operators $P_k = T_k^* T_k$ are positive, sum to one and

$$p_k = \langle \psi | P_k \psi\rangle.$$

Thus, if we are only interested in the probabilities of outcomes, all that matters are the positive operators P_k. Also note that $\sum_{k=1}^K P_k = I_{\mathcal{H}_i}$.

A set of operators $P_1, ..., P_K \in B(\mathcal{H})$ is called a **K outcome positive operator-valued measure** or **K-POVM** provided that they are positive and sum to the identity. A K-POVM is called a **K outcome projection-valued measure** or **K-PVM** when each P_k is a projection, i.e., $P_k = P_k^2 = P_k^*, \forall k$.

Given a K-POVM, define $\Phi : \ell_K^\infty \to B(\mathcal{H})$ by

$$\Phi(\sum_{k=1}^K a_k \delta_k) = \sum_{k=1}^K a_k P_k.$$

It is not hard to see that this map is unital and positive. Hence, by Stinespring's theorem it is UCP. Conversely, given any unital positive map $\Phi : \ell_K^\infty \to B(\mathcal{H})$ if we set $P_k = \Phi(\delta_k)$ then this defines a K-POVM.

Thus, studying K-POVM's is the same as studying UCP maps on ℓ_K^∞.

By Stinespring's dilation theorem, given a K-POVM on $\mathcal{H}$, there is another Hilbert space $\mathcal{K}$, an isometry $V : \mathcal{H} \to \mathcal{K}$ and a unital *-homomorphism, $\pi : \ell_K^\infty \to B(\mathcal{K})$ such that

$$P_k = V^* \pi(\delta_k) V, 1 \leq k \leq K.$$

Note that if we set $E_k = \pi(\delta_k)$ then $\{E_1, ..., E_K\}$ is a K-PVM on $\mathcal{K}$.

Thus, each K-POVM dilates to a K-PVM and this process is often called a **purification** of the K-POVM.

In this simple case it is possible to carry out this process directly. If we set $\mathcal{K} = \mathcal{H}^{(K)}$, let E_k denote the projection onto the k-th copy of $\mathcal{H}$, and define

$V : \mathcal{H} \to \mathcal{H}^{(K)}$ by $Vh = (P_1^{1/2}h, ..., P_K^{1/2}h)$, then it is not hard to see that $P_k = V^*E_kV, 1 \leq k \leq K$. Thus, we have obtained a Stinespring-like dilation. However, this dilation will often not satisfy the minimality property needed to be equivalent to *the* Stinespring dilation.

7. C*-algebras of Groups

These are used in the study of unitary representations and arise quite a bit in QI. We will only look at the case of discrete groups, i.e., we will not worry about groups with topologies.

Given a group G we will always let e denote its identity element. By a *unitary representation of G on the Hilbert space $\mathcal{H}$*, we mean a map $\pi : G \to B(\mathcal{H})$ such that π is a unital homomorphism, i.e., $\pi(e) = I_{\mathcal{H}}$, $\pi(gh) = \pi(g)\pi(h)$, into the group of unitary operators on $\mathcal{H}$. Thus, $\pi(g^{-1}) = \pi(g)^{-1} = \pi(G)^*$. Sometimes we set $\pi(g) = U_g$ in which case $U_gU_h = U_{gh}$.

When we have a unitary representation, then it makes sense to take linear combinations, such as $\sqrt{2}\pi(g) + 3i\pi(h)$. But in the group this would have no meaning. The *-algebra of the group is an object that allows such expressions.

The **-algebra of G*, denoted $\mathbb{C}(G)$, is the vector space with a basis denoted $\{u_g : g \in G\}$, so typical elements look like $a = \sum_i \alpha_i u_{g_i}$ and $b = \sum_j \beta_j u_{g_j}$ where both sums are over finitely many group elements and $\alpha_i, \beta_j \in \mathbb{C}$. The product is given by

$$a \cdot b = \sum_{i,j} \alpha_i\beta_j u_{g_ig_j}$$

and

$$a^* = \sum_i \overline{\alpha_i} u_{g_i^{-1}}.$$

Now it should be clear that every unitary representation π of G on $\mathcal{H}$ induces a unital *-homomorphism, $\tilde{\pi} : \mathbb{C}(G) \to B(\mathcal{H})$ via

$$\tilde{\pi}(\sum_i \alpha_i u_{g_i}) = \sum_i \alpha_i \pi(g_i).$$

Conversely, given a unital *-homomorphism, $\tilde{\pi} : \mathbb{C}(G) \to B(\mathcal{H})$ we get a unitary representation by setting $\pi(g) = \tilde{\pi}(u_g)$.

Often to avoid the additional subscripts, it is better to write a typical element as $\sum_g \alpha_g u_g$ with the understanding that only finitely many of the α_g's are non-zero.

There is a second, equivalent, description of $\mathbb{C}(G)$ that is often used. Namely, $\mathbb{C}(G)$ can also be defined as the finitely supported complex valued functions on G. The correspondence is that we think of the finite sum $\sum_g \alpha_g u_g$ as the function $f : G \to \mathbb{C}$ with $f(g) = \alpha_g$. Thus, in this description, if we define $\delta_g : G \to \mathbb{C}$ to be the function given by

$$\delta_g(h) = \begin{cases} 1, & h = g \\ 0, & h \neq g \end{cases},$$

then in the identification of the two descriptions $u_g \sim \delta_g$.

The product looks quite different in this representation. That is because if we think of $f_1 \sim \sum_g \alpha_g u_g = a$ and $f_2 \sim \sum_h \beta_h u_h = b$, then we need a formula for the product at a typical point k. To see this note that

$$\big(\sum_g \alpha_g u_g\big)\big(\sum_h \beta_h u_h\big) = \sum_{g,h} \alpha_g \beta_h u_{gh} =$$
$$\sum_k \big(\sum_{gh=k} \alpha_g \beta_h\big) u_k = \sum_k \big(\sum_{h=g^{-1}k} \alpha_g \beta_h\big) u_k =$$
$$\sum_k \big(\sum_g \alpha_g \beta_{g^{-1}k}\big) u_k =$$
$$\sum_k \big(\sum_g f_1(g) f_2(g^{-1}k)\big) u_k.$$

Thus, the formula for the product of two functions is

$$(f_1 \star f_2)(k) = \sum_g f_1(g) f_2(g^{-1}k),$$

which is called the **convolution product**.

So the summary is that the convolution product is just the product that we have defined above when we regard finite sums as finitely supported functions. The convolution product is preferred by many since it is the one that generalizes most easily to the case of continuous groups and measures. But for discrete groups it is perhaps less intuitive.

Note that the *-operation from the functions viewpoint is

$$f^*(g) = \overline{f(g^{-1})}.$$

Finally, to form the C*-algebra of a group, we first need to define a norm on $\mathbb{C}(G)$. We do this by setting

$$\|\sum_g \alpha_g u_g\| = \sup\{\|\sum_g \alpha_g \pi(g)\| | \pi \text{ a unitary representation }\}.$$

Note that the supremum is finite since the sums are finite. In fact,

$$\|\sum_g \alpha_g u_g\| \leq \sum_g |\alpha_g|,$$

since the norm of every unitary is 1. It is easily checked that with this norm $\mathbb{C}(G)$ satisfies all the axioms needed to be a C*-algebra, except that it may not be complete as a normed space.

So we define $C^*(G)$ to be the completion of $\mathbb{C}(G)$ in this norm. Then it is readily checked that we obtaina C*-algebra in this fashion.

The key property of this C*-algebra is that every time we have a unitary representation $\pi : G \to B(\mathcal{H})$, then it induces a unital *-homomorphism, $\tilde{\pi} : \mathbb{C}(G) \to B(\mathcal{H})$, which then extends by continuity to a unital *-homomorphism, $\tilde{\pi} : C^*(G) \to B(\mathcal{H})$. Conversely, every unital *-homomorphism $\tilde{\pi}$ of $C^*(G)$ defines a unitary representation of G by setting $\pi(g) = \tilde{\pi}(u_g)$.

Thus we have one-to-one correspondences between:

- unitary representations of G,
- unital *-homomorphisms of $\mathbb{C}(G)$,
- unital *-homomorphisms of $C^*(G)$.

We now look at a few examples.

Example 7.1 $(\mathbb{Z}, +)$**.** This group is also known as the infinite cyclic group. It is a little annoying as a first example, because we have written groups multiplicatively and this example is additive. So that $e = 0$ and if we let $g = 1$ then $g^n = 1 + \cdots + 1 = n$. Note that to define a unitary representation, all we need to do is select any unitary $\pi(1) = U \in B(\mathcal{H})$ and set $\pi(n) = U^n$ with the understanding that $U^0 = I_{\mathcal{H}}$. Thus,

$$\|\sum_n \alpha_n u_n\| = \sup \|\sum_n \alpha_n U^n\|,$$

where the supremum is over all unitaries on all Hilbert spaces.

If we first consider the case that $\mathcal{H}$ is one-dimensional, then a unitary is just a complex number on the unit circle $\mathbb{T}$. If we set $\pi(1) = \lambda \in \mathbb{T}$, then $\tilde{\pi}(\sum_n \alpha_n u_n) = \sum_n \alpha_n \lambda^n$. If we consider the Laurent polynomial, $p(z) = \sum_n \alpha_n z^n$ which we regard as a continuous function on $\mathbb{T}$, i.e., as an element of $C(\mathbb{T})$, then $\tilde{\pi}(\sum_n \alpha_n u_n) = p(\lambda)$. Thus, we see that

$$\|\sum_n \alpha_n u_n\| \geq \sup_{\lambda \in \mathbb{T}} |p(\lambda)| = \|p\|_\infty,$$

where this last quantity is the norm of $p(z) \in C(\mathbb{T})$.

On the other hand, given any unitary U we know that its spectrum $\sigma(U)$ is a subset of $\mathbb{T}$. Thus, by the spectral theorem we know that

$$\|\sum_n \alpha_n U^n\| = \|p(U)\| = \sup\{|p(\lambda)|; \lambda \in \sigma(U)\} \leq \sup\{|p(\lambda)| : \lambda \in \mathbb{T}\}.$$

These two inequalities show that

$$\|\sum_n \alpha_n u_n\|_{\mathbb{C}(\mathbb{Z})} = \|\sum_n \alpha_n z^n\|_{C(\mathbb{T})}.$$

From this it follows that $C^*(\mathbb{Z})$ and $C(\mathbb{T})$ are *-isomorphic via the map that sends $u_n \to z^n$.

Example 7.2 $(\mathbb{Z}_n, +)$**.** We now look at the cyclic group of order n. The analysis is similar to the infinite cyclic case. Every unitary representation, $\pi : \mathbb{Z}_n \to B(\mathcal{H})$ is determined by the image of the generator, $U = \pi(1)$, except now since $n = 0$, the unitary must satisfy $U^n = I_{\mathcal{H}}$. We set $\omega = e^{2\pi i/n}$ and let $X_n = \{\omega^j : 0 \leq j \leq n-1\}$. The fact that $U^n = I_{\mathcal{H}}$ means that $\sigma(U) \subseteq X_n$.

Every element of $\mathbb{C}(\mathbb{Z}_n)$ can be written as $\sum_{j=0}^{n-1} \alpha_j u_j$ and since every finite dimensional normed space is complete, we have that $C^*(\mathbb{Z}_n) = \mathbb{C}(\mathbb{Z}_n)$.

A similar analysis to above shows that

$$\|\sum_{j=0}^{n-1}\alpha_j u_j\|_{C^*(\mathbb{Z}_n)} = \sup\{|\sum_{j=0}^{n-1}\alpha_j\lambda^j| : \lambda \in X_n\} = \|\sum_{j=0}^{n-1}\alpha_j z^j\|_{C(X_n)}.$$

Thus, $C^*(\mathbb{Z}_n)$ and $C(X_n)$ are *-isomorphic.

There is a bit more that can be said in this case, since X_n is a finite discrete set. If we let $\delta_j : X_n \to \mathbb{C}$ be defined by

$$\delta_j(\omega^i) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases},$$

then it is easy to see that $\delta_j^2 = \delta_j = \overline{\delta_j}$, so that these elements are projections in $C(X_n)$. Also, $C(X_n) = span\{\delta_j : 0 \leq j \leq n-1\}$.

On the other hand give a unitary U with $U^n = I_{\mathcal{H}}$, we will have projections E_j onto the eigenspace for the eigenvalue ω^j and "diagonalizing" U we may write it as

$$U = \sum_{j=0}^{n-1}\omega^j E_j.$$

Using some algebra we can see that

$$E_j = 1/n\sum_{k=0}^{n-1}\big(\omega_j U\big)^k.$$

This suggests that in the group algebra, if we set

$$e_j = 1/n\sum_{k=0}^{n-1}(\omega^- j)^k u_k,$$

then these elements should satisfy:

$$e_j = e_j^2 = e_j^*,\ e_i e_j = 0, i \neq j,\ \ 1 = \sum_{j=0}^{n-1} e_j,$$

and that in the isomorphism between $\mathbb{C}(\mathbb{Z}_n)$ and $C(X_n)$ the image of e_j is δ_j. We leave it to the reader to verify these facts.

Thus, the group algebra $\mathbb{C}(\mathbb{Z}_n) \simeq C(X_n)$ can be thought of as the algebra generated by an element $u = u_1$, satisfying $u^n = 1$ or as generated by n orthogonal projections that sum to the identity.

7.1. **Free Products of Groups.** One other concept that we shall use is the concept of the free product of groups. Given two groups G and H, their free product, denoted $G \star H$ is the unique group, containing G and H as subgroups, with the property that whenever K is another group and we are given group homomorphisms, $\pi : G \to K$ and $\rho : H \to K$, then there is a unique group homomorphism, $\gamma : G \star H \to K$ such that $\gamma(g) = \pi(g)$ and $\gamma(h) = \rho(h)$ for all $g \in G$, $h \in H$.

Operationally, $G \star H$ can be viewed as the set of **words** in G and H, where a word w is an alternating of elements of G and H. The identity elements e_G and e_H are identified as the same element in $G \star H$ and this element is the identity of $G \star H$. Thus,

$$w_1 = g_1 h_1 g_2,\ w_2 = g_2^{-1} h_2 g_4 h_3,\ w_3 = h_4 g_5 h_5,$$

are all examples of words. The operation of multiplication of words is a process called **concatenation**. Briefly, this operation just strings the elements of two words together, multiplying them whenever possible. Thus,

$$w_1 w_3 = g_1 h_1 g_2 h_4 g_5 h_5,$$

while,

$$w_1 w_2 = g_1 h_1 (g_2 g_2^{-1}) h_2 g_4 h_3 = g_1 h_1 e_G h_2 g_4 h_3 = g_1 (h_1 h_2) g_4 h_3.$$

It is easily checked that the inverse of a word is just the string of inverses written in the reverse order, so that

$$w_1^{-1} = g_2^{-1} h_1^{-1} g_1^{-1}.$$

Note that $G \star H$ and $H \star G$ are the same group. The homomorphism γ is usually denoted $\pi \star \rho$ so we have that, we also have $\pi \star \rho = \rho \star \pi$.

Finally, given unitary representations $\pi : G \to B(\mathcal{H})$ and $\rho : H \to B(\mathcal{H})$ we obtain a unitary representation $\pi \star \rho : G \star H \to B(\mathcal{H})$.

This operation extends so that one can form the free product of any collection of groups. Some of these have special notations. The free product of n copies of $(\mathbb{Z}, +)$ is denoted $\mathbb{F}_n$ and is called the **free group on n generators**. If we write $(\mathbb{Z}, +)$ multiplicatively, so that it is just powers of some generator g and we write a second copy as powers of some generator h, then typical elements of $\mathbb{F}_2$ look like words in powers of g and h.

Since all we need to define a unitary representation of $\mathbb{Z}$ on $\mathcal{H}$ is a choice of a unitary $U \in B(\mathcal{H})$ with $\pi(1) = U$, every time we choose two unitaries $U, V \in B(\mathcal{H})$ we obtain a representation of $\mathbb{F}_2$ on $\mathcal{H}$. The representation sends words in the generators g, h to the corresponding word in U and V. Similarly, every choice of n unitaries in $B(\mathcal{H})$ defines a representation of $\mathbb{F}_n$ on $\mathcal{H}$.

The free product of n copies of $(\mathbb{Z}_n, +)$ is denoted $\mathbb{F}_{n,k}$. Every choice of n unitaries, $U_1, ..., U_n \in B(\mathcal{H})$ with $U_i^k = I_{\mathcal{H}}$ defines a unitary representation of $\mathbb{F}_{n,k}$.

The group algebra $\mathbb{C}(\mathbb{F}_{n,k})$ can be thought of as being generated by n unitaries, $u_1, ..., u_n$ each satisfying $u_x^k = 1$, with no relations between distinct unitaries. Or since we may write each $u_x = \sum_{j=0}^{k-1} \omega^j e_{x,j}$ in terms of its spectral projections, where $\omega = e^{2\pi i/k}$, it is also generated by families of projections $\{e_{x,j} : 1 \leq x \leq n,\ 0 \leq j \leq k-1\}$ with $e_{x,i} e_{x,j} = 0, \forall i \neq j$ and $\sum_{j=0}^{k-1} e_{x,j} = 1,\ \forall x$.

7.2. **Conditional Joint Bivariate Densities.** Suppose that we have two separate labs A and B in some joint quantum state. If A can perform n_1 different measurements, each with k_1 outcomes and B can perform n_2 different measurements each with k_2 outcomes then there is a joint conditional probability density

$$p(i, j|x, y),$$

which represents the probability that if A performs measurment x, and B performs measurement y, then they see outcomes i and j respectively.

Bell showed that in the case $n_1 = n_2 = k_1 = k_2 = 2$ that a certain mathematical model for describing these densities, gave a strictly larger set of densities than was allowed by Einstein's theory of "hidden local variables". Consequently, by producing quantum experiments that exhibited densities in this larger set but not in the smaller set, the first proofs that entanglement is a real phenomenon were given. Later Tsirelson realized that there was more than one mathematical model for these densities and it is only very recently that it has been shown that these various mathematical models to not all give rise to the same sets of densities. The first difference was shown by our own W. Slofstra.

We will now describe these differing mathematical models for representing these various sets of densities.

7.3. **Classical or Local Densities.** In abstract probability theory we are given a set Ω a collection $\mathcal{E}$ of subsets of Ω called **events**, and a map $P : \mathcal{E} \to [0, 1]$ that assigns a probability to each event. For each measurement x, outcome i is an event $A_{x,i}$ with

$$\cup_{i=1}^{k_1} A_{x,i} = \Omega \text{ and } A_{x,i} \cap A_{x,j} = \emptyset, i \neq j.$$

Similarly, for each measurment y, there would be events $B_{y,j}$ such that

$$\cup_{j=1}^{k_2} B_{y,j} = \Omega \text{ and } B_{y,i} \cap B_{y,j} = \emptyset, i \neq j.$$

In this case the joint density is given by

$$p(i, j|x, y) = P(A_{x,i} \cap B_{y,j}).$$

The set of all possible $(p(i, j|x, y))$ that one can obtain this way is a subset of $\mathbb{R}^{n_1 n_2 k_1 k_2}$ denoted by

$$C_{loc}(n_1, n_2, k_1, k_2),$$

or

$$LOC(n_1, n_2, k_1 k_2),$$

is known as the set of **local conditional densities** or **local correlations**. If Einstein's "local hidden variable theory" had been correct, these would be the only densities that nature could produce.

7.4. **The Basic Model for Quantum Conditional Densities.** In the basic model, A is assumed to have a finite dimensional state space, $\mathcal{H}_A$ and POVM's $\{P_{x,i} : 1 \leq x \leq n_1, 1 \leq i \leq k_1\}$, while B has another finite dimensional state space $\mathcal{H}_B$ and POVM's $\{Q_{y,j} : 1 \leq y \leq n_2, 1 \leq j \leq k_2\}$. Then the operator $P_{x,i} \otimes Q_{y,j}$ on $\mathcal{H}_A \otimes \mathcal{H}_B$ is the measurement operator for outcome (i, j) when the pair of measurements (x, y) is performed. Thus, if the labs are in a joint pure quantum state $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ then we obtain

$$p(i, j|x, y) = \langle \psi | (P_{x,i} \otimes Q_{y,j})\psi \rangle.$$

The set of all possible densities that one can obtain this way by varying, the POVM's finite dimensional Hilbert spaces and state vectors, is denoted by

$$C_q(n_1, n_2, k_1, k_2),$$

or

$$Q(n_1, n_2, k_1, k_2), \text{ or } Q_{\otimes}(n_1, n_2, k_1, k_2)$$

and is generally referred to as the set of **quantum correlations** or **quantum conditional densities**.

Since every PVM is a POVM, if we required that $P_{x,a}$ and $Q_{y,b}$ are both PVM's in the above definition we would get a potentially smaller set of joint densities. On the other hand, by the last section's dilation theorem, the family of POVM's $P_{x,a}$ can be written as $P_{x,i} = V^* E_{x,i} V$, where $E_{x,i}$ is a family of PVM's on a space $\mathcal{K}_A$ and $V : \mathcal{H}_A \to \mathcal{K}_A$ is an isometry. Similarly, if we write $Q_{y,j} = W^* F_{y,j} W$ for some family of PVM's $F_{y,j}$ on a space $\mathcal{K}_B$ and isometry $W : \mathcal{H}_B \to \mathcal{K}_B$ and set $\gamma = (V \otimes W)\psi \in \mathcal{K}_A \otimes \mathcal{K}_B$, then we have that

$$p(i, j|x, y) = \langle \psi | (P_{x,i} \otimes Q_{y,j})\psi \rangle = \langle \gamma | (E_{x,i} \otimes F_{y,j})\gamma \rangle.$$

Thus, we see that if in our definition of elements of $C_q(n_1, n_2, k_1, k_2)$ we had replaced POVM's by PVM's we would have obtained the same set.

For this reason, you will often see either of the definitions used in the literature, but it is important to know that the set that is being defined is independent of this ambiguity.

In the case that $\psi = \psi_A \otimes \psi_B$, one has that

$$p(i, j|x, y) = \langle \psi_A | P_{x,i} \psi_A \rangle \langle \psi_B | Q_{y,j} \psi_B \rangle,$$

and it is possible to show that this density is in C_{loc}.

7.5. **Quantum Spatial Correlations.** If we keep all the definitions the same as in the last subsection, but remove the requirement that the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ be finite dimensional we obtain a larger set denoted

$$C_{qs}(n_1, n_2, k_1, k_2).$$

Again, using the same reasoning as above, we see that this set is the same if we use PVM's in place of POVM's in the definition.

7.6. **Quantum Commuting Correlations.** In this model, instead of each lab having its own private state space, it is assume that there is a universal state space $\mathcal{H}$ that all of the measurement operators act on. The assumption that the labs are "separated" is interpreted as saying that the outcome doesn't depend on the order that the measurements are applied, in other words that

$$P_{x,i}Q_{y,j} = Q_{y,j}P_{x,i}, \ \forall x, y, i, j.$$

If the combined system is in state $\psi \in \mathcal{H}$ then the conditional probability densities are given by

$$p(i, j|x, y) = \langle \psi | P_{x,i}Q_{y,j}\psi \rangle.$$

The set of all such densities is denoted

$$C_{qc}(n_1, n_2, k_1, k_2).$$

It is not at all clear in this case, if the set C_{qc} remains the same if we replace POVM's by PVM's. This is because there is only one Hilbert space involved and if we tried to apply our above ideas, we would get potentially different spaces when we dilated the $P_{x,i}$'s and the $Q_{y,j}$'s. However, using POVM's or PVM's both yield the same set C_{qc}.

This fact was first proved in [**?**] and a direct statement and proof of this fact can be found in [**?**] and [**?**]. All three proofs use C*-algebras of free groups and the theory of tensor products of these algebras. We outline the ideas here.

By Corollary 11.12, the set of n_1 PVM's $\{P_{x,i} : 1 \leq x \leq n_1, \ 0 \leq i \leq k_1 - 1\}$ on $B(\mathcal{H})$ each with k_1 outcomes, induces a UCP map,

$$\Phi : C^*(\mathbb{F}_{n_1,k_1}) \to B(\mathcal{H}),$$

with $\Phi(e_{x,i}) = P_{x,i}$. Similarly, $\{Q_{y,j} : 1 \leq y \leq n_2, \ 0 \leq j \leq k_2 - 1\}$ induces another UCP map

$$\Psi : C^*(\mathbb{F}_{n_2,k_2}) \to B(\mathcal{H}).$$

Because the range of Φ commutes with the range of Ψ, this pair of maps induces a UCP, $\Phi \odot \Psi$ map on the **maximal C*-tensor product** of these two algebras, $C^*(\mathbb{F}_{n_1,k_1}) \otimes_{max} C^*(\mathbb{F}_{n_2,k_2})$ defined on generators by

$$\Phi \odot \Psi(e_{x,i} \otimes e_{y,j}) = \Phi(e_{x,i})\Psi(e_{y,j}).$$

Applying the Stinespring dilation to this map one obtains the desired commuting families of commuting PVM's.

The key takeaway at this point is that to fully understand these commuting densities, one needs the theory of group C*-algebras and some results from the theory of tensor products of C*-algebras. The tensor theory is covered in the later chapter where we examine C*-algebras in greater depth.

7.7. **Quantum Approximate Correlations.** These are the set of densities that can be approximated by densities in C_q, in other words the closure of the subset $C_q(n_1, n_2, k_1, k_2)$ of $\mathbb{R}^{n_1 n_2 k_1 k_2}$ in the usual Euclidean topology. We will denote this set by

$$C_{qa}(n_1, n_2, k_1, k_2).$$

One could ask about the closure of $C_{qs}(n_1, n_2, k_1, k_2)$ but it turns out that it has the same closure as C_q. Again, this set is the same if we use POVM's or PVM's in its definition.

Originally Tsirelson believed that

$$C_q(n_1, n_2, k_1, k_2) = C_{qs}(n_1, n_2, k_1, k_2) = C_{qa}(n_1, n_2, k_1, k_2) = C_{qc}(n_1, n_2, k_1, k_2),$$

and so whether or not various pairs are equal became known as the **Tsirelson conjectures**. These sets are all equal in the case that $n_1 = n_2 = k_1 = k_2$.

Interest in these problems ramped up when [**?**] and [**?**], proved that the equality of C_{qa} and C_{qc} for all n's and k's is equivalent to a famous problem in operator algebras, the **Connes Embedding Problem**, sometimes known as the **Connes' Embedding Conjecture**, although Connes himself never committed to whether or not the statement was true or not. This conjecture is quite famous, because if it were true then it would answer many other questions in mathematics.

We now know that all of these sets are different for large enough values of the n's and k's, and so that the Connes' Embedding Conjecture is false.

The first gap between these sets was discovered by W. Slofstra [**?**], who used the theory of non-local games to first show that $C_q \neq C_{qc}$ and then that $C_q \neq C_{qa}$ for the n's and k's large enough. Now it is known that C_q is not closed as long as $n_1, n_2 \geq 5$ and $k_1, k_2 \geq 2$.

In fact in [**?**] it is shown that for $t \in \left[\frac{\sqrt{5}-1}{2\sqrt{5}}, \frac{\sqrt{5}+1}{2\sqrt{5}}\right]$, and for $1 \leq x, y \leq 5$, if we set

$$p(0,0|x,x) = t,\ p(0,1|x,x) = p(1,0|x,x) = 0,\ p(1,1|x,x) = 1-t,$$

and for $x \neq y$, set

$$p(0,0|x,y) = \frac{t(5t-1)}{4},\ p(0,1|x,y) = p(1,0|x,y) = \frac{5t(1-t)}{4},$$
$$p(1,1|x,y) = \frac{(1-t)(4-5t)}{4}.$$

Then $p \in C_{qa}(5,2)$ and $p \in C_q(5,2) \iff t$ is rational.

In [**?**], Coladangelo and Stark show that $C_q(4,3) \neq C_{qs}(4,3)$.

Finally, in the monumental paper $MIP^* = RE$[**?**] prove that $C_{qa} \neq C_{qc}$ for sufficiently large values of the n's and k's, using methods from everywhere, but especially, complexity theory and non-local games. Thus, showing that the Connes' Embedding Problem has a negative answer.

We summarize some of what is known about these sets and their relations below.

- $C_{loc} \subseteq C_q \subseteq C_{qs} \subseteq C_{qa} \subseteq C_{qc}$,
- C_{loc}, C_{qa}, C_{qc} are closed for all n_1, n_2, k_1, k_2,
- $C_loc \subsetneq C_q$ for all $n_1, n_2, k_1, k_2 \geq 2$,
- $C_q = C_{qs} = C_{qa} = C_{qc}$ when $n_1 = n_2 = k_1 = k_2 = 2$,
- C_q and C_{qs} are not closed when $n_1, n_2 \geq 5$ and $k_1, k_2 \geq 2$,
- Consequently, $C_{qs} \subsetneq C_{qa}$ when $n_1, n_2 \geq 5$ and $k_1, k_2 \geq 2$,
- $C_q \subsetneq C_{qs}$ for $n_1, n_2 \geq 4$ and $k_1, k_2 \geq 3$,
- $C_{qa} \subsetneq C_{qc}$ for some sufficiently large values of n_1, n_2, k_1, k_2.

8. Non-Local Games

When one performs a measurement of a quantum system the outcome is random and the probabilities that one can obtain in this fashion turn out to have some very counterintuitive properties. In this section we will discuss some games where using quantum mechanics to "randomly roll the dice" leads to much higher probabilities of winning then our classical intuition tells us is possible. We will also look at *prover systems* which are games that are designed to test if a solution to a problem has been found. These games can be classically won with probability 1 if and only if a solution to the problem has been found. Prover systems can be used in situations where there are so many equations and so many variables that it is impractical to write down all the values of all the variables and check that they satisfy all the equations. Instead, if players keep giving correct answers each time that they play a round of a prover system game, then the Referee can feel fairly certain that the players have indeed solved the problem. Surprisingly, there are many prover systems for which quantum-assisted players can design a strategy that will win with probability 1 even when no actual solution exists! So players that have access to quantum phenomena can fool a prover system.

The games that we will look at all involve three parties. There are two players, who we will call Alice and Bob. Alice and Bob are not competing, but instead they are cooperating to try and return correct *answers* to *questions* posed by the third party, who we will call the Referee. The questions are called the *inputs* and the answers are called the *outputs* of the game.

One property of these games is that whether or not the pair of answers given by Alice and Bob is correct, depends on the pair of questions Alice and Bob received and not on just their individual questions. So although Alice and Bob both know the *rules*, i.e., they both know which pairs of answers are right for a pair of questions, they must both give their replies without knowing what question the other was asked. Formally, this is what is meant by saying that they are *non-communicating*. It is easiest to imagine that they are in separated soundproof rooms so that they cannot hear what the Referee tells the other player or what the other player tells the Referee.

Let's look at a couple of famous example of these games to illustrate these ideas.

8.1. **The CHSH Game.** The initials stand for Clauser, Holt, Shimony, and Horne [**?**]. This game was designed to illustrate one of the original inequalities that could be used to experimentally show that *entanglement* is an actual phenomena [**?**].

In this game Alice and Bob are each given a number from the binary field $\mathbb{Z}_2 = \{0, 1\}$, say x and y, respectively, and must return binary numbers, a and b. They will win if the sum of their numbers $a + b$ is of the same parity as the product xy. That is, they win if $a + b = xy(mod\, 2)$.

Suppose that they also know that the Referee will pick each of the four possible input pairs, $(0, 0), (0, 1), (1, 0), (1, 1)$ randomly and with equal probability of $1/4$. I think that many of you will see that a good strategy is for them to agree ahead of time that no matter what they receive they will always answer with a 0. In this case $a + b = 0$, which is what the value of xy will be equal to $3/4$ of the time, i.e., for all input pairs except $(1, 1)$. This strategy has a winning probability or *value* of $3/4$.

This strategy is what is meant by a *deterministic* strategy. Deterministic means that Alice's and Bob's outputs are both functions of their inputs, and that they use the same function every time that they play the game. Another deterministic strategy with the same value is if they always return 1. For this game it is straightforward to list all the possible deterministic strategies, i.e., pairs of functions, and check that the strategy of always returning 0 or of always returning 1 are the two strategies with the highest probability of winning.

Next we want to introduce the idea of *random strategies*. With a random strategy, if a player receives the same input x at two different rounds of the game, they might give different answers. An example of a *classical shared random strategy* for the CHSH game would be if the Referee rolled a die each time he selected an input and along with their inputs, he told Alice and Bob the number on the die. Alice and Bob would then use this random number to adjust their strategy. It is a little harder to see, but it is still the case that their highest winning probability is still $3/4$. Roughly, this is because at each round they are still selecting a deterministic strategy.

Now suppose that instead of a roll of the die, the external input is a pair of laser beams, one beam shining into each players room. Each player designs two quantum experiments to perform on their laser beam and each measurement has two outcomes, which we can label with a 0 and a 1. A key property of quantum mechanics is that outcomes of experiments can be random. They will perform one experiment if they receive a 0 and the second experiment if they receive a 1. They perform the experiment and report their outcome back to the Referee. So in this way they randomly generate outputs for their give inputs.

If the two laser beams are not entangled, then their optimal winning probability is still $3/4$. But if the laser beams are entangled in just the right way their winning probability can be as large as about .85, in fact, $cos^2(\pi/8)$, to be precise. This is roughly because when things are entangled, the outcome

of one experiment subtly influences the outcome of the other experiment. It is no wonder that Einstein referred to entanglement as "...spooky action at a distance...".

It is important to emphasize that the laser beams contain no information about the questions that were asked, that is they don't vary with the questions, they are just constantly in the same entangled state. It is this particular entangled state and the experiments that they will perform that are chosen to achieve this higher winning probability for this particular game.

So far we have been a bit informal. Now we will show that there is a density in $C_q(2,2,2,2)$ that gives a higher probability of winning than .75, i.e., the best classical strategy.

For this we let $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^2$ and let

$$\psi = \frac{1}{\sqrt{2}}\big(e_0 \otimes e_0 + e_1 \otimes e_1\big).$$

Let us set

$$P(\theta) = \begin{pmatrix} cos^2(\theta) & cos(\theta)sin(\theta) \\ cos(\theta)sin(\theta) & sin^2(\theta) \end{pmatrix},$$

which can be seen to be the orthogonal projection onto the line spanned by $cos(\theta)e_0 + sin(\theta)e_1$. We will let $P(\theta)^\perp = I_2 - P(\theta) = P(\theta + \pi/2)$. for each input $x \in \{0,1\}$, Alice will use measurement operators,

$$P_{x,0} = P(\theta_{A,x}),\ P_{x,1} = P_{x,0}^\perp.$$

Similarly, for each input $y \in \{0,1\}$, Bob will use measurement operators,

$$Q_{y,0} = P(\theta_{B,y}),\ Q_{y,1} = Q_{y,0}^\perp.$$

In this case we have that

$$\begin{aligned} p(0,0|x,y) &= \langle \psi | (P_{x,0} \otimes Q_{y,0})\psi \rangle = \\ &\frac{1}{2}\big(\langle e_0 \otimes e_o | (P_{x,0} \otimes Q_{y,0}) e_0 \otimes e_0 \rangle + \langle e_0 \otimes e_0 | (P_{x,0} \otimes Q_{y,0}) e_1 \otimes e_1 \rangle\big) + \\ &\frac{1}{2}\big(\langle e_1 \otimes e_1 | (P_{x,0} \otimes Q_{y,0}) e_0 \otimes e_0 \rangle + \langle e_1 \otimes e_1 (P_{x,0} \otimes Q_{y,0}) e_1 \otimes e_1 \rangle\big) = \\ &\frac{1}{2}\big(cos^2(\theta_{A,x})cos^2(\theta_{B,y}) + 2cos(\theta_{A,x})sin(\theta_{A,x})cos(\theta_{B,y})sin(\theta_{B,y}) + sin^2(\theta_{A,x})sin^2(\theta_{B,y})\big), \end{aligned}$$

with similar formulas for the $p(1,0|x,y), p(0,1|x,y), p(1,1|x,y)$.

For this particular strategy the probability of winning is

$$\begin{aligned} &\frac{1}{4}\big(p(0,0|0,0) + p(1,1|0,0) + p(0,0|0,1) + p(1,1|0,1) + \\ &\qquad p(0,0|1,0) + p(1,1|1,0) + p(0,1|1,1) + p(1,0|1,1)\big). \end{aligned}$$

We leave it to the reader to show that when we pick

$$\theta_{A,0} = 0,\ \theta_{A,1} = \pi/4,\ \theta_{B,0} = \pi/8,\ \theta_{B,1} = -\pi/8,$$

then this probability will be $cos^2(\pi/8) > .75$.

8.2. **Mermin's Magic Square and Linear Constraint Games.** The type of prover system game used by Slofstra comes from games built around solving systems of linear equations over the binary field. Suppose that we are given a set of m linear equations in n variables all over the binary field $\mathbb{Z}_2$. We can write this in matrix-vector form as $Mx = c$, where M is an $m \times n$ matrix with entries from $\mathbb{Z}_2$, c is an n-tuple of constants from $\mathbb{Z}_2$ and x is a vector of unknowns–the variables. We make a game of it as follows:

The Referee sends Alice a number i between 1 and m, representing one of the equations and sends Bob a number j between 1 and n representing a variable. Alice replies with a binary n-tuple $a = (a_1, ..., a_n)$ and Bob replies with a single bit b, i.e., an element of $\mathbb{Z}_2$. They automatically win if $m_{i,j} = 0$. When $m_{i,j} \neq 0$, they win if Alice's n-tuple is a solution to the i-th equation and if $a_j = b$, i.e., if Bob correctly predicted the j-th entry of her solution.

Given a conditional probability density $p(a, b|x, y)$ we will say that is a *perfect* density for the game if $p(a, b|x, y) = 0$ whenever (a, b) is a losing pair of outputs for the input pair (x, y). Thus, a perfect density never produces a losing output pair.

This game turns out to be a *prover system*. That is, the game has a perfect deterministic strategy if and only if the system of equations has an actual solution.

However, there are many linear systems for which this game has been shown to have a *perfect* quantum density when there is no classical solution. That is, by using entanglement Alice and Bob can produce correct solutions for every round, even though there is no classical solution.

A famous example of this phenomena is *Mermin's Magic Square Game*[**?**]. This is a system of 6 equations in 9 variables, but the system of equations is best pictured by arranging the variables in a square:

$$\begin{matrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7 & x_8 & x_9 \end{matrix}$$

in which case the equations are that each row should have an even sum, and the first two columns should have an even sum, but the third column should have an odd sum. A little thought shows that it is impossible to find a solution to these equations. Nonetheless, using entanglement, there is a perfect quantum strategy for this game so that no matter what row or column Alice is given and no matter what variable Bob is given they will always give replies that satisfy the rules.

Perhaps the reader has realized by now that, besides entanglement, one other key feature that allows these strange outcomes is that our games are *memoryless*. That is if at some round Alice assigns value 1 to a particular variable, then there is nothing that prevents her from giving that same variable value 0 at some later round.

Slofstra [**?**] obtained his separation by implicitly creating a linear system with about 200 equations in about 200 variables that had a perfect density in the set C_{qa} and no perfect density in the set C_q.

8.3. **Finite Input-Output Games.** We now make the ideas above a bit more formal.

By a **two-person finite input-output game** we mean a tuple $\mathcal{G} = (I_A, I_B, O_A, O_B, \lambda)$ where I_A, I_B, O_A, O_B are finite sets and

$$\lambda : I_A \times I_B \times O_A \times O_B \to \{0, 1\}$$

is a function that represents the rules of the game, sometimes called the *predicate*. The sets I_A and I_B represent the inputs that Alice and Bob can receive, and the sets O_A and O_B, represent the outputs that Alice and Bob can produce, respectively. A referee selects a pair $(v, w) \in I_A \times I_B$, gives Alice v and Bob w, and they then produce outputs (answers), $a \in O_A$ and $b \in O_B$, respectively. They win the game if $\lambda(v, w, a, b) = 1$ and lose otherwise. Alice and Bob are allowed to know the sets and the function λ and cooperate before the game to produce a strategy for providing outputs, but while producing outputs, Alice and Bob only know their own inputs and are not allowed to know the other person's input. Each time that they are given an input and produce an output is referred to as a **round** of the game.

A **deterministic strategy** for a game is a pair of functions, $h : I_A \to O_A$ and $k : I_B \to O_B$ such that if Alice and Bob receive inputs (v, w) then they produce outputs $(h(v), k(w))$. A deterministic strategy wins every round of the game if and only if

$$\forall (v, w) \in I_A \times I_B, \lambda(v, w, h(v), k(w)) = 1.$$

Such a strategy is called a **perfect deterministic strategy**.

A **random strategy** for a game $\mathcal{G}$ is a conditional probability density $p(a, b|v, w)$, which represents the probability that, given inputs $(v, w) \in I_A \times I_B$, Alice and Bob produce outputs $(a, b) \in O_A \times O_B$. Thus, $p(a, b|v, w) \geq 0$ and for each (v, w),

$$\sum_{a \in O_A, b \in O_B} p(a, b|v, w) = 1.$$

In this paper we identify random strategies with their conditional probability density, so that a random strategy will simply be a conditional probability density $p(a, b|v, w)$.

A random strategy is called **perfect** if

$$\lambda(v, w, a, b) = 0 \implies p(a, b|v, w) = 0, \forall (v, w, a, b) \in I_A \times I_B \times O_A \times O_B.$$

Thus, for each round, a perfect strategy gives a winning output with probability 1.

We next discuss **local** random strategies, which are also sometimes called **classical**, meaning not quantum. They are obtained as follows: Alice and Bob share a probability space (Ω, P), for each input $v \in I_A$, Alice has a random variable, $f_v : \Omega \to O_A$ and for each input $w \in I_B$, Bob has a

random variable, $g_w : \Omega \to O_B$ such that for each round of the game, Alice and Bob will evaluate their random variables at a point $\omega \in \Omega$ via a formula that has been agreed upon in advance. This yields conditional probabilities,

$$p(a, b|v, w) = P(\{\omega \in \Omega \mid f_v(\omega) = a, g_w(\omega) = b\}).$$

Note that this density will be an element of $C_{loc}(n_1, n_2, k_1, k_2)$, where $n_1 = |I_A|$ and $n_2 = |I_B|$ are the cardinalities of Alice and Bob's input sets, respectively, and $k_1 = |O_A|, k_2 = |O_B|$ are the respective cardinalities of Alice and Bob's output sets.

A local density $p(a, b|v, w)$ will be a perfect strategy for a game $\mathcal{G}$ if and only if

$$\forall (v, w) \in I_A \times I_B,\ P(\{\omega \in \Omega \mid \lambda(v, w, f_v(\omega), g_w(\omega)) = 0\}) = 0,$$

or equivalently,

$$\forall (v, w) \in I_A \times I_B,\ P(\{\omega \in \Omega \mid \lambda(v, w, f_v(\omega), g_w(\omega)) = 1\}) = 1.$$

If we have a perfect local strategy and set

$$\Omega_1 = \cap_{v \in I_A, w \in I_B} \{\omega \in \Omega \mid \lambda(v, w, f_v(\omega), g_w(\omega)) = 1\},$$

then $P(\Omega_1) = 1$ since I_A and I_B are finite sets; in particular, Ω_1 is non-empty. If we choose any $\omega \in \Omega_1$ and set $h(v) = f_v(\omega)$ and $k(w) = g_w(\omega)$, then it is easily checked that this is a perfect deterministic strategy.

Thus, a perfect classical random strategy exists if and only if a perfect deterministic strategy exists. An advantage to using a perfect classical random strategy over a perfect deterministic strategy, is that it is difficult for an observer to construct a deterministic strategy even after observing the outputs of many rounds.

For $t \in \{loc, q, qs, qa, qc\}$, we say that $p(a, b|v, w)$ is a **perfect t-strategy** for a game provided that it is a perfect strategy that belongs to the set C_t.

8.4. **Values of Strategies.** Suppose that we are given a game $\mathcal{G}$ and a probability density on its set of inputs, i.e., a map,

$$\Pi : I_A \times I_B \to [0, 1]$$

satisfying $\sum_{x \in I_A, y \in I_B} \Pi(x, y) = 1$. In this case, given a strategy $p(a, b|x, y)$ the probability that we will win the game is called the **value of the strategy**, denoted $\omega(\mathcal{G}, \Pi, p)$ and is given by summing the probability that we receive (x, y) times the probability that we give a winning reply (a, b). This can seen to be given by the formula

$$\omega(\mathcal{G}, \Pi, p) := \sum_{x,y,a,b} \Pi(x, y) p(a, b|x, y) \lambda(x, y, a, b).$$

Note that a perfect strategy will always have value 1, no matter what density the Π is equal to. Conversely, if $\Pi(x, y) \neq 0$, $\forall x, y$, then any strategy with value 1, will have to be perfect.

For each $t \in \{loc, q, qs, qa, qc\}$ we set

$$\omega_t(\mathcal{G}, \Pi) := sup\{\omega(\mathcal{G}, \Pi, p) : p \in C_t\}.$$

Since the closure of C_q and C_{qs} are both equal to C_{qa}, and taking a supremum over a set is the same as taking it over its closure, we have

$$\omega_q(\mathcal{G}, \Pi) = \omega_{qs}(\mathcal{G}, \Pi) = \omega_{qa}(\mathcal{G}, \Pi).$$

So we are really only interested in the cases $t \in \{loc, q, qc\}$, although sometimes it is nice to use the other instances.

We caution the reader that our notation is not standard, often ω_{loc} is denoted ω, ω_q is denoted ω^* and ω_{qc} is denoted $\tilde{\omega}$. Also many authors include the density Π as part of the definition of the game.

In this language, our first example was that for the uniform density on inputs,

$$\omega_{loc}(CHSH, \Pi) = .75,\ \omega_q(CHSH, \Pi) = cos^2(\pi/8).$$

Since $C_q(2,2,2,2) = C_{qc}(2,2,2,2)$ we also have that

$$\omega_{qc}(CHSH, \Pi) = cos^2(\pi/8).$$

In the paper MIP*=RE, a game $\mathcal{G}$ is produced such that, for the uniform density on inputs one has

$$\omega_{qc}(\mathcal{G}, \Pi) = 1,\ \omega_q(\mathcal{G}, \Pi) \leq 1/2,$$

thus showing that $C_{qa} \subsetneq C_{qc}$. Thereby, disproving the Connes' Embedding Conjecture.

8.5. **The Graph Colouring and Graph Homomorphism Games.** A *graph* G is specified by a vertex set $V(G)$ and an edge set $E(G) \subseteq V(G) \times V(G)$, satisfying $(v, v) \notin E(G)$ and $(v, w) \in E(G) \implies (w, v) \in E(G)$. The **c-coloring game** for G has inputs $I_A = I_B = V(G)$ and outputs $O_A = O_B = \{1, ..., c\}$ where the outputs are thought of as different colors. They win provided that whenever Alice and Bob receive adjacent vertices, i.e., $(v, w) \in E$, their outputs are different colors and when they receive the same vertex they output the same color. Thus, $(v, w) \in E(G) \implies \lambda(v, w, a, a) = 0, \forall a,\ \lambda(v, v, a, b) = 0, \forall v \in V(G), \forall a \neq b$ and the rule function is equal to 1 for all other tuples.

We claim that this game is a prover system for the chromatic number $\chi(G)$, in the sense that a perfect deterministic strategy exists if and only if $\chi(G) \leq c$. This is one of the HW problems.

Given two graphs G and H, a *graph homomorphism from G to H* is a function $f : V(G) \to V(H)$ with the property that $(v, w) \in E(G) \implies (f(v), f(w)) \in E(H)$. The **graph homomorphism game** from G to H has inputs $I_A = I_B = V(G)$ and outputs $O_A = O_B = V(H)$. They win provided that whenever Alice and Bob receive inputs that are an edge in G, their outputs are an edge in H and that whenever Alice and Bob receive the same vertex in G they produce the same vertex in H.

A perfect deterministic strategy exists for the graph homomorphism game if and only if there is an actual graph homomorphism from G to H.

Finally, it is not difficult to see that if K_c denotes the complete graph on c vertices then a graph homomorphism exists from G to K_c if and only if

G has a c-coloring. This is because any time $(v, w) \in E(G)$ then a graph homomorphism must send them to distinct vertices in K_c. Indeed, the rule function for the c-coloring game is exactly the same as the rule function for the graph homomorphism game from G to K_c.

Given a graph G we set $\chi_t(G)$ equal to the least c for which there exists a perfect t-strategy for the c-coloring game for G. The above inclusions imply that

$$\chi(G) = \chi_{loc}(G) \geq \chi_q(G) \geq \chi_{qs}(G) \geq \chi_{qa}(G) \geq \chi_{qc}(G).$$

In [**?**] it is shown that $\chi_q(G) = \chi_{qs}(G)$ for all graphs. Currently, it is unknown if there are any graphs that separate $\chi_q(G), \chi_{qa}(G)$ and $\chi_{qc}(G)$ or whether these three parameters are always equal. Examples of graphs are known for which $\chi(G) > \chi_q(G)$. In fact, $\chi_q(G)$ can be exponentially smaller. For details, see [**?**] and [**?**].

Similarly, we say that there is a *t-homomorphism* from G to H if and only if there exists a perfect t-strategy for the graph homomorphism game from G to H. It is unknown if q-homomorphisms, qa-homomorphisms and qc-homomorphisms are distinct or coincide.

References

[1] Arveson, William B.; Subalgebras of C^*-algebras. *Acta Math.* 123 (1969), 141?224.

[2] Choi, Man Duen; Completely positive linear maps on complex matrices. *Linear Algebra Appl.* 10 (1975), 285?290.

[3]

[4] Coladangelo, A; Stark, J., *Unconditional separation of finite and infinite-dimensional quantum correlations*, arXiv:1804.05116

[5] Conway, John:; *A course in functional analysis*,

[6] Davidson, Kenneth R.; *C*-algebras by example*,

[7] Dunford, Nelson; Schwartz, Jacob T, *Linear operators. Part II: Spectral theory*, Reprint of the 1963 original. Wiley Classics Library. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1988.

[8] Dykema, K.; Paulsen, V.; *Synchronous Correlation Matrices and Connes' Embedding Conjecture*, Journal of Mathematical Physics, 57, 015215 http://dx.doi.org/10.1063/1.4936751

[9] Dykema, K.; Paulsen, V.; Prakash, J., *Non-closure of the set of quantum correlations via graphs*, Comm. Math. Phys. 365 (2019), no. 3, 1125?1142. https://doi.org/10.1007/s00220-019-03301-1

[10] Fritz, T.;

[11] Goberg, I.C. and Krein, M.G.; *Introduction to the Theory of Linear Nonselfadjoint Operators*, Translations of Mathematical Monographs, American Mathematical Society, 1969

[12] S. J. Harris, S. K. Pandey, and V. I. Paulsen, *Entanglement and Non-locality.*

[13] Helton, J.W.; Meyer, K.P.; Palsen, V.I.; Satriano, M., *Algebras, synchronous games and chromatic numbers of graphs*, New York J. Math. 25 (2019), 328?361.

[14] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, Henry Yuen, *MIP*=RE*, arXiv:2001.04383

[15] M. Junge, M. Navascues, C. Palazuelos, D. Perez-Garcia, V. B. Scholz, R. F. Werner, *Connes' embedding problem and Tsirelson's problem*, J. Math. Phys. 52, 012102 (2011).

[16] Kim, S. J.; Paulsen, V.I.; Schaffhauser, C., *A synchronous game for binary constraint systems*, Journal of Mathematical Physics 59, 032201 (2018); doi: 10.1063/1.4996867
[17] Ozawa, N.; *About the Connes Embedding Conjecture—Algebraic approaches*, Jpn. J. Math., 8 (2013), 147–183.
[18] Paulsen, Vern I.; *Completely bounded maps and operator algebras*, Cambridge University Press, 2002.
[**?**] Paulsen, V.; Severini, S.; Stahlke, D.; Tovodor, I.; Winter, A., *Estimating quantum chromatic numbers*, J. Funct. Anal. 270 (2016), no. 6, 2188?2222.
[19] Paulsen, Vern I.; Todorov, Ivan G.; *Quantum chromatic numbers via operator systems*, Q. J. Math. 66(2015), no. 2, 677-692.
[20] Pedersen, G.; *C*-algebras and their automorphism groups*,
[21] Slofstra, W.;

Institute for Quantum Computing and Department of Pure Mathematics, University of Waterloo, Waterloo, ON, Canada N2L 3G1
Email address: vpaulsen@uwaterloo.ca