Lecture notes on the Statistical Structure of Quantum Theory

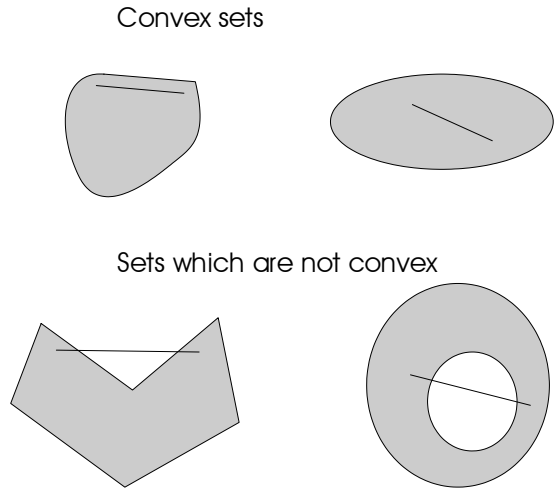
Peter Harremoës

January 17, 2012

1 Convexity

1.1 Convex sets

Let $\mathcal{H}$ be a vector space. We will identify the vectors in $\mathcal{H}$ with points. Let C be a subset of $\mathcal{H}$. Let $\vec{u}, \vec{v} \in C$. If $\vec{u} \neq \vec{v}$ then the map $\alpha \curvearrowright \alpha \cdot \vec{u} + (1-\alpha) \cdot \vec{v}$ is a parametrization of a straight line through $\vec{u}$ and $\vec{v}$ if the domain is $\mathbb{R}$. If the domain of the parametrization is $[0;1]$ then the curve is a line with end points $\vec{u}$ and $\vec{v}$. If C is a circle or a square in two dimensions the whole line $\alpha \cdot \vec{u} + (1-\alpha) \cdot \vec{v}, \alpha \in [0;1]$ is within C. The same will hold if C has the shape of an ball in 3 dimensions or $[a;b]^d$ in $\mathbb{R}^d$. Not all subsets of a vector space satisfy this property. For instance $\{0,1\}$ as subset of the vector space R, or a torus as a subset of a three dimensional vector space do not satisfy the property. The set $C \subseteq \mathcal{H}$ is said to be *convex* if for all $\vec{u}, \vec{v} \in C$ and all $\alpha \in [0;1]$ the vector $\alpha \cdot \vec{u} + (1-\alpha) \cdot \vec{v}$ belongs to C. If $\alpha \in [0;1]$ then $\alpha \cdot \vec{u} + (1-\alpha) \cdot \vec{v}$ is called a *convex combination* or a *mixture* of the two points/vectors $\vec{u}$ and $\vec{v}$.

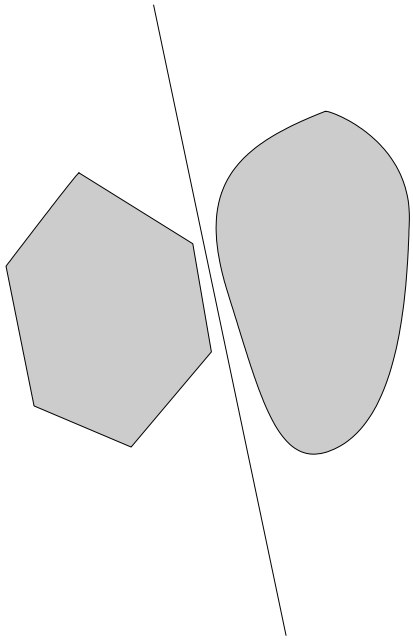


Convex sets and sets which are not convex.

Example 1 *Let $M_+^1(U)$ be the set of probability vectors on a finite space U. Then $M_+^1(U)$ is convex.*

A *hyperplane* PL in $\mathcal{H}$ is a subset of the form $\{\vec{u} \in \mathcal{H} \mid \Phi(\bar{u}) = \lambda\}$ where Φ is linear and λ is a constant. Let C and K be subsets of $\mathcal{H}$. Then C and K are *separated* by the hyperplane PL if

$$\Phi(\bar{u}) \begin{cases} < \lambda & \text{for} \quad \bar{u} \in K \\ \geq \lambda & \text{for} \quad \bar{u} \in C \end{cases} .$$



Two convex sets separated by a hyperplane.

The *distance* $dist\,(C,K)$ *between the sets* is defined by

$$dist\,(C,K) = \inf_{\vec{u}\in C, \vec{v}\in K} \|\vec{v}-\vec{u}\|\,.$$

Theorem 2 *Assume that $\mathcal{H}$ is a Hilbert space. Let $C \subseteq \mathcal{H}$ be convex sets and assume that C is closed. Let $\vec{u} \in \mathcal{H}$ be a vector. Put $d_{\min} = dist\,(\vec{u},C)$. If $\vec{u} \notin C$ then there exists $\vec{v} \in C$ such that $dist\,(\vec{u},\vec{v}) = d_{\min}$, and the hyperplane $\left\{\vec{w} \in \mathcal{H} \mid \langle \vec{w}-\vec{u} \mid \vec{v}-\vec{u}\rangle = d_{\min}^2\right\}$ separates the sets C and $\{\vec{u}\}$.*

Proof. Let $\vec{v}_n \in C$ be a sequence such that $\|\vec{v}_n - \vec{u}\| \to d_{\min}$. For all $m,n \in \mathbb{N}$

$$\frac{1}{2}\vec{v}_m + \frac{1}{2}\vec{v}_n \in C.$$

Therefore

$$\begin{aligned} d_{\min}^2 &\leq \left\| \frac{1}{2}\vec{v}_m + \frac{1}{2}\vec{v}_n - \vec{u} \right\|^2 \\ &= \frac{1}{4}\left\|\vec{v}_m - \vec{u} + \vec{v}_n - \vec{u}\right\|^2 \\ &= \frac{1}{4}\left(2\left\|\vec{v}_m - \vec{u}\right\|^2 + 2\left\|\vec{v}_n - \vec{u}\right\|^2 - \left\|\vec{v}_m - \vec{v}_n\right\|^2\right) \\ &= \frac{1}{2}\left\|\vec{v}_m - \vec{u}\right\|^2 + \frac{1}{2}\left\|\vec{v}_n - \vec{u}\right\|^2 - \frac{1}{4}\left\|\vec{v}_m - \vec{v}_n\right\|^2. \end{aligned}$$

Then

$$\begin{aligned}\left\|\vec{v}_m-\vec{v}_n\right\|^2 &\leq 4d_{\min}^2-2\left\|\vec{v}_m-\vec{u}\right\|^2-2\left\|\vec{v}_n-\vec{u}\right\|^2\\ &\to 0 \text{ for } m,n\to\infty,\end{aligned}$$

and $\vec{v}_n$ is a Cauchy sequence and converges to some vector $\vec{v}$. The set C is closed which proves that $\vec{v}\in C$. The equation $\left\|\vec{v}-\vec{u}\right\|=d_{\min}$ holds by continuity. Let $\vec{w}$ be an element in C. Then $\alpha\cdot\vec{w}+(1-\alpha)\cdot\vec{v}\in C$ by convexity. Therefore

$$\begin{aligned}\left\|\vec{v}-\vec{u}\right\|^2 &\leq \left\|\alpha\vec{w}+(1-\alpha)\,\vec{v}-\vec{u}\right\|^2\\ &= \left\|\alpha\left(\vec{w}-\vec{u}\right)+(1-\alpha)\left(\vec{v}-\vec{u}\right)\right\|^2\\ &= \alpha^2\left\|\vec{w}-\vec{u}\right\|^2+(1-\alpha)^2\left\|\left(\vec{v}-\vec{u}\right)\right\|^2+2\alpha\left(1-\alpha\right)\left\langle\vec{w}-\vec{u}\mid\vec{v}-\vec{u}\right\rangle.\end{aligned}$$

For $\alpha=0$ equality holds. Taking the derivative with respect to α in $\alpha=0$ gives

$$0\leq -2\left\|\left(\vec{v}-\vec{u}\right)\right\|^2+2\left\langle\vec{w}-\vec{u}\mid\vec{v}-\vec{u}\right\rangle,$$

which is the desired inequality. ■

We see that a point belongs to a closed convex set if and only if it cannot be separated from the set. The vector $\vec{v}\in C$ which minimizes $\left\|\vec{v}-\vec{u}\right\|$ is called the *projection of* $\vec{u}$ *on* C. The following theorem is known as the *Separation Theorem for Convex Sets.*

Theorem 3 *Let $C,K\subseteq\mathcal{H}$ be convex sets and assume that C is closed and K is compact. Then $C\cap K=\varnothing$ if and only C and K are separated by a hyperplane.*

Proof. Put $d_{\min}=dist\left(C,K\right)$. Choose $\bar{v}_0\in K$ such that $dist\left(C,\bar{v}_0\right)\leq d_{\min}+1$. Then $\tilde{C}=\left\{\vec{u}\in C\mid dist\left(\vec{u},K\right)\leq d+1\right\}$ is compact and there exists $\vec{u}_1\in\tilde{C}$ and $\vec{v}_1\in K$ such that $\left\|\vec{u}_1-\vec{v}_1\right\|=d$. Then $\vec{u}_1$ is the projection of $\vec{v}_1$ on C and $\vec{v}_1$ is the projection of $\vec{u}_1$ on K. Then the set

$$\left\{\vec{w}\in\mathcal{H}\mid\left\langle\vec{w}-\vec{v}_1\mid\vec{u}_1-\vec{v}_1\right\rangle=\frac{1}{2}d_{\min}^2\right\}$$

is a separation hyperplane. ■

The Separation Theorem can be generalized to infinite dimensional spaces with suitable topologies, but then the proof requires the axiom of choice (or some of its equivalents).

Theorem 4 *Let C be a closed convex subset of a finite dimensional Hilbert space $\mathcal{H}$. Let P be a probability measure on C. Then the vector defined by*

$$\vec{v}=E\left(\vec{u}\right)=\int_C\vec{u}\;dP\vec{u}$$

is element in C.

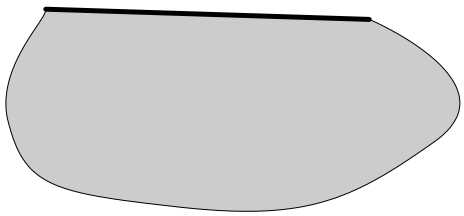
Proof. Assume that $\vec{v} \notin C$. Then there exists a linear map Φ such that $\Phi(\vec{u}) \geq 0$ for $\vec{u} \in C$ and $\Phi(\vec{v}) < 0$. Then $\Phi(\vec{u})$ can be considered as a positive random variable, and

$$\begin{aligned} 0 &\leq E(\Phi(\vec{u})) \\ &= \Phi(E(\vec{u})) \\ &\Phi(\vec{v}), \end{aligned}$$

and we have a contradiction. ■

The point/vector $\vec{v} = E(\vec{u}) = \int_C \vec{u}\, dP\vec{u}$ is called a *convex combination* or *mixture* of the points/vectors in $\text{supp}(P)$. For a set $A \in \mathbb{R}^d$ (which is not assumed to be convex) the *convex hull of* A is the set of vectors of the form $\int_C \vec{u}\, dP\vec{u}$ where P denotes probability measures on A.One sees that the convex hull is the smallest convex set containing A as a subset. In infinite dimensional spaces it is more complicated, and one has to distinguish between finite probability measures (probability vectors) and more general probability measures.

A subset F of a convex set C is said to be a *face* if for all $\vec{u}, \vec{v} \in C$ and $\alpha \in]0;1[$ we have that $\alpha \cdot \vec{u} + (1-\alpha) \cdot \vec{v} \in F$ implies that $\vec{u}, \vec{v} \in F$. A face of the convex set C consisting of one point is called an *extreme point* of C. The set of extreme points in C is denoted $\partial_e C$.



A convex with one of its faces marked.

By the dimension of a convex/face set we will mean the dimension of the affine space spanned by the set.

Lemma 5 *A non-empty compact convex set $C \subseteq \mathbb{R}^d$ contains at least one extreme point.*

Proof. The proof is by induction in the dimension of C. For $\dim(C) = 0$ C contains only one point which must be an extreme point.

Assume that the lemma is true for all convex sets of dimension less than n and that $\dim(C) = n+1$. Let $\Phi : \mathbb{R}^d \to \mathbb{R}$ be a linear map such that $\Phi(C)$ is not a point. Then $\Phi(C)$ is an interval $[a;b]$. Then set $\Phi^{-1}(b) \cap C$ is a face of C of dimension less than n and therefore $\Phi^{-1}(b) \cap C$ contains at least one extreme point, and this point must also be an extreme point of C. ■

The following important theorem is due to Caratheodory.

Theorem 6 *Let C be of finite dimension d. Assume that C is convex and compact, and that $\vec{u} \in C$ is a point. Then there exists extreme points $\vec{v}_1, \vec{v}_2, ..., \vec{v}_{d+1}$ and a probability vector $(\alpha_1, \alpha_2, ..., \alpha_{d+1})$ such that*

$$\vec{u} = \sum_{j=1}^{d+1} \alpha_j \cdot \vec{v}_j \ .$$

Proof. The proof is by induction in d. For $d = 0$ it is obvious.

Assume that the theorem is true for convex sets of dimension less that n and that $d = n+1$. The convex set C is compact and therefore C contains an extreme point $\vec{v}_1$. Let β_0 be the smallest $\beta \in \mathbb{R}$ such that $\vec{w} = (1-\beta)\,\vec{v} + \beta\vec{u} \in C$. Put

$$F = \{\vec{v} \in C \mid \exists \gamma \in]0;1[\,, \exists \tilde{v} \in C : \gamma \cdot \vec{v} + (1-\gamma) \cdot \tilde{v} = \vec{w}\}\,.$$

One easily checks that F is a face containing $\vec{w}$, and the dimension must be at most d. Therefore $\vec{w}$ can be written as a convex combination of at most d extreme points. By replacing $\vec{w}$ by a convex combination of d extreme points in the formula

$$\vec{u} = (1-\beta)^{-1}\,(\vec{w} - \beta \cdot \vec{v})$$

and using that $\beta \leq 0$ we get $\vec{v}$ written as a convex combination of at most $d+1$ extreme points. ■

Another way to state the theorem is that the map $M_+^1\,(\partial_e C) \to C$ is surjective. In general it is not injective. For instance the centre of a circle is a convex combination of any pair of antipodic points. If the map $M_+^1\,(\partial_e C) \to C$ is injective then C is said to be a simplex.

1.2 Convex functions

Let f be a real function with a convex set C as domain. Then f is said to be *convex* if for all $\vec{u}, \vec{v} \in C$ and all $\alpha \in [0;1]$ the following inequality is satisfied

$$f\,(\alpha \cdot \vec{u} + (1-\alpha) \cdot \vec{v}) \leq \alpha \cdot f\,(\vec{u}) + (1-\alpha) \cdot f\,(\vec{v})\,.$$

If f is convex then $-f$ is said to be *concave*.

The *hypergraph* of f is the set $\{(\vec{v}, y) \in C \times \mathbb{R} \mid y \geq f\,(\vec{v})\}$. We see that f is convex if and only if the hypergraph is convex. Using the results from the previous section on the hyper graph of a convex function we get

$$f\,(E\,(\vec{u})) \leq E\,(f\,(\vec{u}))\ .$$

The inequality is called Jensen's inequality after the Danish mathematician J. L. W. V. Jensen (1859-1925).

2 States and measurements

A physical experiment consists of a physical arrangement $\mathcal{O}$ and a result $\mathcal{R}$. These data may be of an arbitrary nature: They may be discrete if the measurement instrument make registrations of an even for instance the appearance

of a particle, they can represent a scalar or vector depending on whether the measurement instrument has one or more scales, or the result may be the entire trace of particle in a bobble chamber. To give a unified treatment we will assume that the results are the elements in a finite set $\mathfrak{U}$ and denote by $\mathcal{B}(\mathfrak{U})$ the set of all subsets of $\mathfrak{U}$. Similarly we will assume that the set of physical arrangements O (and the possible values of all other variables introduced in this chapter) is a measurable set. An individual result is rarely completely determined by the preparation. If one considers the repeated experiments the frequency of the different results will be uniquely determined.

Often $\mathcal{O}$ and $\mathcal{R}$ will be composed of a number of other variables. In the situations we shall consider O can be divided into a *preparation* $\mathcal{P}$ and a *measurement procedure* μ. We will assume that $\mathcal{P}$ and $\mathcal{M}$ are independent and thereby we exclude correlation between $\mathcal{P}$ and $\mathcal{M}$. We put $\mathcal{O} = (\mathcal{P}, \mathcal{M})$. By the preparation an experimental setup is established by giving initial conditions and input data. By the measurement procedure the "prepared object" is coupled to the measurement apparatus which results in the observation $\mathcal{R}$. The "object" can be considered as a "black box", a coupling or an information channel between preparation and measurement apparatus. Let $\mathfrak{P}$ be the set of preparations and $\mathfrak{M}$ the set of measurement procedures. Instead of making just one specific preparation we can equip the set of preparations with a probability measure and perform the different preparations with probabilities according to the probability measure. This is called randomization. A measurement $\mu \in \mathfrak{M}$ which maps gives $\mathfrak{P}$ into probability measures in $M_+^1(\mathfrak{U})$ can be extended to a map $M_+^1(\mathfrak{P}) \to M_+^1(\mathfrak{U})$. The set $M_+^1(\mathfrak{P})$ can be equipped with the pseudo-metric *dist* defined by

$$dist\,(s_1, s_2) = \sup_{\mu \in \mathfrak{M}} \|\mu_{s_1} - \mu_{s_2}\|_{tot} \ .$$

We define the *state space* $\mathfrak{S}$ as the completion of $M_+^1(\mathfrak{P})$ with respect to *dist*, and the elements in $\mathfrak{S}$ are called states. Hereby we get a mapping $M_+^1(\mathfrak{P}) \to \mathfrak{S}$. This means that 2 preparations gives the same state if it is not possible to distinguish them by any measurement. The state space is automatically bounded because total variation is bounded. By a measurement we will understand any affine mapping $\mathfrak{S} \to M_+^1(\mathfrak{U})$. With these definitions the state space depend on the set of measurements considered. Therefore it is misleading to say for instance "the electron is in state ϕ". Instead one should say "our knowledge about the electron is completely described by ϕ.

Example 7 *At time $t = 0$ a classical particle is send from the position $\bar{x} = \bar{0}$ with velocity $\bar{v} \in \mathbb{R}^3$. This shall be our preparations $\mathfrak{P} = \mathbb{R}^3$. Assume that the particle is not subjected to any forces. For any $B \subseteq \mathbb{R}^3$ define a measurement by a detection at time $t > 0$ whether the particles in $B \subseteq \mathbb{R}^3$. In general classical mechanics is characterized by $\mathfrak{S} = M_+^1(\mathfrak{P})$ and $\mathfrak{M} = \mathcal{B}(\mathfrak{P})$.*

Example 8 *A Stern-Gerlach-apparatus contains an anisotropic magnetic field, which splits a beam of electrons in 2 beams of identical intensity. As the preparation we consider a source emitting the electrons one by one such that only*

the electrons localized in one of the beams, will continue. The Stern-Gerlach-apparatus can be rotated around the axis formed by the beam. The apparatus is therefore characterized by an angle $\theta_1 \in \mathbb{T}/2\pi\mathbb{Z}$ *which is our set of preparations. another Stern-Gerlach-apparatus is placed after the first, and at last there is two detectors which detects whether an electron is drawn in one or the other direction. This apparatus can be rotated too, and therefore it is characterized by an angle* θ_2*. Let* θ^i *denote the vector* $(\cos(\theta_i), \sin(\theta_i))$*, and let* d_i *be detection of the i'th detector. As we shall see we will have (under ideal circumstances) that*

$$\begin{aligned} P(d_1 \mid d_2\ ) &= \cos^2\left(\frac{\theta_1 - \theta_2}{2}\right) \\ &= \frac{1 + \theta^1 \cdot \theta^2}{2}\ . \end{aligned}$$

If P *is a probability distribution on the set of preparations then the measurement* θ_2 *is mapped into a probability distribution on the event space given by*

$$\begin{aligned} P(d_1 \mid d_i\ ) &= \int \frac{1 + \theta^1 \cdot \theta^2}{2}\, dP\theta^1 \\ &= \frac{1 + \left(\int \theta^1\, dP\theta^1\right) \cdot \theta^2}{2}\,. \end{aligned}$$

The mapping $P \to \int \theta^1\, dP\theta^1$ *is therefore exactly the mapping* $\mathfrak{S}(\mathbb{T}) \to \mathfrak{S}$*. This shows that* $\mathfrak{S}$ *is isomorphic with* $\left\{(a,b) \in \mathbb{R}^2 \mid a^2 + b^2 \leq 1\right\}$*. This set can via the mapping*

$$(a,b) \to \frac{1}{2}\begin{pmatrix} 1+a & b \\ b & 1-a \end{pmatrix}$$

be identified with the set of density matrices on the Hilbert space $\mathbb{R}^2$*. In elementary quantum mechanics the state space can usually be identified with the set of density matrices/operators on a Hilbert space (often complex).*

By the above construction $\mathfrak{S}$ will be a metrically complete convex set. Any metrically complete convex set K can be obtained by a similar construction. Let $\mathfrak{U}$ be equal to K, and let the set of measurement procedures $\mathfrak{M}$ be affine mappings $K \to [0;1]$. Then there exists an affine mapping $A : M_+^1(K) \to K$ where a probability distribution s on K is mapped into the corresponding convex combination in K. For $s \in M_+^1(K)$ and $m \in \mathfrak{M}$ define $M(A(s))$ a distribution on $\mathfrak{U} = [\mathfrak{o};\mathfrak{1}]$ which should be the corresponding measurement. Then $\mathfrak{S} = K$.

Definition 9 *A measurement is said to be* simple *if the map* $\phi \to \mu_\phi(D)$ *is extreme among all functionals* $\mathfrak{S} \to [0;1]$ *for all subsets* $D \subseteq \mathfrak{U}$*.*

With this definition the simple measurements are extreme in the set of all measurements. As we shall see later sometimes there may exist extreme measurements which are not simple.

Definition 10 *Let $\mathfrak{S} = M_+^1(U)$ be a simplex and let $\mu : M_+^1(U) \to M_+^1(V)$ be a measurement. Then μ is given by its values on the extreme points $\delta_x, x \in U$. The measurement is extreme if and only if $\mu(\delta_x)$ is extreme in $M_+^1(V)$ for all $x \in V$, but the extreme points in $M_+^1(V)$ are of the form $\delta_y, y \in V$. Therefore the extreme measurements are given by a map $f : U \to V$ and*

$$\mu_\phi(D) = \phi\left(f^{-1}(D)\right)$$

where ϕ is a probability measure in $M_+^1(U)$. The functional $\phi \to \phi\left(f^{-1}(D)\right)$ from $M_+^1(U)$ to $[0;1]$ maps the extreme points δ_x into extreme points of $[0;1]$ and therefore the functional is extreme. We see that the extreme measurements on a simplex are simple measurements of the form $\phi\left(f^{-1}(D)\right)$. All other measurements are mixtures of the simple measurements. The simple measurements correspond observations to of which subsets $f^{-1}(D)$ of U the "result" $x \in V$ belongs to.

Example 11 *In the Stern-Gerlach experiment the state space can be identified with the unit circle. The simple measurements can then be identified with projections of the unit circle into a diameter of the circle. If the state is given by the density matrix*

$$S = \frac{1}{2}\begin{pmatrix} 1+a & b \\ b & 1-a \end{pmatrix}$$

then the simple measurements are given by

$$(a,b) \curvearrowright \frac{1 + \begin{pmatrix} a \\ b \end{pmatrix} \cdot \theta^2}{2}$$

where $\theta^2 = \begin{pmatrix} x \\ y \end{pmatrix}$ is a unit vector. Put

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

This can also be written as

$$\frac{1 + \begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}}{2} = Tr(ST) \ .$$

3 Group representations on convex sets

In physics it is often possible to perform certain "actions" on the set of preparations. The instruments may be rotated in space. Then it is often possible to make the same measurements on the rotated preparation as on the original preparation. Similarly a measurement instrument may be rotated so that the rotations acts on the set of measurements. The set of rotations form a group

because there exists an inverse to any rotation. We see that two preparations can be distinguished by measurements if and only if the preparations can be distinguished after the action of a certain rotation. Therefore the rotations induces an action on the state space. In this sections we shall see that the assumption that a group acts on the state space only leaves few possibilities of the shape of the state space.

Definition 12 *Let G be a (topological) group and let a finite dimensional state space $\mathfrak{S}$. Then an action of G on $\mathfrak{S}$ is a map which sends a group element $g \in G$ into an affine map $a_g : \mathfrak{S} \to \mathfrak{S}$ is such a way that the following conditions are fulfilled*

- *The neutral element $e \in G$ is mapped into the identity.*
- *For all $g, h \in G$ the following equation is satisfied $a_{gh} = a_g \circ a_h$.*
- *The action is continuous, i.e. for any sequence g_n converging to g in G the sequence a_{g_n} converges to a_g.*

Representation may be complicated, but here the focus will be on the simplest representation which are of special interest in quantum theory.

Definition 13 *Let G be a group acting on a finite dimensional state space $\mathfrak{S}$. Then the representation is said to be irreducible if any invariant convex subset is either a point or spans $\mathfrak{S}$.*

In elementary quantum mechanics we are interested in irreducible representations, and the word is used because the phenomenons studied have no internal structure. When a free neutron decays and suddenly do not move according to the Dirac equation for a free spin 1/2 particle (it disappear) this shows exactly that the neutron has a more complicated structure than just being a spin 1/2-particle.

Theorem 14 *If the state space $\mathfrak{S}$ has dimension at most 3 and there exists a connected group G acting irreducibly on $\mathfrak{S}$, then $\mathfrak{S}$ is isomorphic to the set of density matrices on a (real or complex) Hilbert space.*

Proof. The group of maps $\mathfrak{S} \to \mathfrak{S}$ is compact. Therefore the range of G in this group is compact, so we may assume that G it self is compact and is equipped with the unique normalized *Haar-measure* m. If $\vartheta \in \mathfrak{S}$ then

$$\vartheta_0 = \int_G a_g(\vartheta) \ dmg$$

is invariant under the action of G. Then the state space can be embedded into a real 3-dimensional Hilbert space with ϑ_0 in the origin. Let $(\cdot \mid \cdot)$ be the inner product in this vector space. Then

$$\langle v \mid w \rangle = \int_G (a_g(v) \mid a_g(w)) \ dmg$$

is an inner product which is invariant under the action of G. Therefore G can be identified with a connected subgroup of the group $O(n)$, $n \leq 3$. Now, all connected subgroups of $O(n), n \leq 3$ are isomorphic to $SO(m), m \leq 3$. Therefore $\mathfrak{S}$ is isomorphic to the unit ball in 1,2 or 3 dimensions. The group $SO(1)$ is trivial and can be excluded. Then we just have to remark that the unit circle and the unit ball are isomorphic to the set of density matrices in 2 real or complex dimensions via the map

$$(a,b,c) \curvearrowright \frac{1}{2}\begin{pmatrix} 1+a & b+ic \\ b-ic & 1-a \end{pmatrix}.$$

■

3.1 Representations of compact groups

Let G be a compact group with Haar-measure μ. Let G act irreducible on the state space $\mathfrak{S}$. We will assume that the representation is non-trivial. Choose $\phi \in \mathfrak{S}$ such that $G(\phi) \neq \phi$. Then $G(\phi)$ spans the convex set $\mathfrak{S}$. For a probability measure $\nu \in \mathfrak{S}(\mathfrak{G})$ define

$$m(\nu) = \int g(\phi)\ d\nu g\ .$$

Then m is a map $\mathfrak{S}(\mathfrak{G}) \to \mathfrak{S}$. This map can be extended to a map $\mathfrak{S}'(\mathfrak{G}) \to \mathbb{C}(\mathfrak{G})$ where $\mathfrak{S}'(\mathfrak{G})$ is the set of Radon distributions on G. Further $L^2(G,\mu)$ is embedded in $\mathfrak{S}'(\mathfrak{G})$. The situation is summarized in the commutative diagram

$$\begin{array}{ccc} \mathfrak{C}(G) & \xrightarrow{m} & \mathfrak{C} \\ \downarrow & & \downarrow \\ \mathfrak{C}'(G) & \xrightarrow{m} & \mathbb{C}(G)\ . \\ \uparrow{\scriptstyle i} & & \\ L^2(G,\mu) & & \end{array} \tag{1}$$

By compactness of G $L^2(G,\mu) = \oplus_I \mathcal{H}_i$ where $\mathcal{H}_i$ are minimal invariant subspaces. This can be rewritten as $L^2(G,\mu) = \oplus_J \mathcal{L}_j$ where $\mathcal{L}_j$ is of the form $\mathcal{H}_i$ if $\mathcal{H}_i = \mathcal{H}_i$ or of the form $\mathcal{H}_i \oplus \mathcal{H}_i$ if $\mathcal{H}_i \neq \mathcal{H}_i$. The elements in $\mathcal{L}_j$ are bounded functions by compactness of G. If (f_k) is a generating set for $\mathcal{L}_j$ then $(\mathrm{Re}(f_k), \mathrm{Im}(f_k))$ is a generation set of reel functions. For $f \in \mathcal{L}_j$ we have $\int f\, d\mu = \langle f \mid 1\rangle = 0$ because $\mathbb{C} \perp \mathcal{L}_j$. Therefore there exists a generating set (b) for $\mathcal{L}_j$ such that $1 + b \geq 0$ and $\int (1+b)\, d\mu = 1$ which shows that $1 + b \in \mathfrak{S}(\mathfrak{G})$. Therefore $1 + \mathcal{L}_j \cap \mathfrak{S}(\mathfrak{G}) = (\mathfrak{1} + \mathcal{L}_j)_+$ spans the convex set $1 + \mathcal{L}_j$. Now, $m\left((1+\mathcal{L}_j)_+\right) \subseteq \mathfrak{S}$ is an invariant convex subset and it will either be a point or span $\mathfrak{S}$. If $m\left((1+\mathcal{L}_j)_+\right)$ is a point then $m(1+\mathcal{L}_j)$ is a point which implies that $m(\mathcal{L}_j) = 0$. If $m\left((1+\mathcal{L}_j)_+\right)$ is a point for all $i \in I$

then $m(\mathfrak{S}'(\mathfrak{G})) = m(\mathbb{C})$. and $m(\mathfrak{S}(\mathfrak{G}))$ is a point which contradicts that the span of $m(\mathfrak{S}(\mathfrak{G}))$ is non-trivial. Therefore there exists $i \in I$ such that $m\left((1+\mathcal{L}_j)_+\right)$ spans $\mathfrak{S}$. Put $K = m^{-1}(m(1))$. Then K is an invariant convex subset of subset of $1+\mathcal{L}_j$ such that $K-1$ is an invariant subspace of $\mathcal{L}_j$. There are 4 possibilities for what $K-1$ can be: 0, $\mathcal{H}_i$, $\mathcal{H}_i$ or $\mathcal{L}_i$. If $\mathcal{H}_i \subseteq K-1$ then we also have $\mathcal{H}_i \subseteq K-1$ and therefore $\mathcal{L}_i \subseteq K-1$. This shows that $m(1+\mathcal{L}_j) \subseteq m(K) = m(1)$ which contradicts that $\mathfrak{S}$ is non-trivial. In a similar way one can exclude that K-1 is equal to $\mathcal{H}_i$ or $\mathcal{L}_i$. Therefore $K-1=0$ and $K=1$. This shows that m is injective on $(1+\mathcal{L}_j)_+$. The above argument shows how all irreducible convex representations can be constructed using the irreducible unitary representations.

If G is commutative the unitary irreducible representations of G are 1-dimensional so that $\dim(\mathcal{H}_i) = 1$. Therefore $\dim(\mathcal{L}_i)$ equals 1 or 2.

$\dim(\mathcal{L}_i) = 1$: Then $\mathfrak{S}$ has the form $[0;1]$ and is isomorphic to the diagonal density matrices on a 2 dimensional Hilbert space. The group of automorphisms is $\mathbb{Z}_2$.

$\dim(\mathcal{L}_i) = 2$: Then $\mathfrak{S}$ can be embedded in the disc Δ which is isomorphic to the density matrices on a 2-dimensional reel Hilbert space. The group of automorphisms is $\mathbb{T} \times \mathbb{Z}_2$.

If G is connected then there are only 2 possibilities: Either $\mathfrak{S}$ is trivial or $\mathfrak{S}$ is isomorphic to Δ.

Example 15 *If G is the group of rotations in 2 dimensions $\mathfrak{S}$ is trivial or $\mathfrak{S}$ is isomorphic to Δ. The trivial representation is called the spin-0 representation. Assume that $\mathfrak{S}$ is isomorphic to Δ. Then the representation is given by $\theta \to n\theta$ for some $n \in \mathbb{Z}$. For $n = 0$ we have the trivial representation, and further the representations n and $-n$ are isomorphic representations, which shows that the representation is given by a number $n \in \mathbb{N}_0$. A quantum mechanic system characterized by the number n is said to have spin $n/2$.*

3.2 Spin

First we shall study rotations in to dimensions. The orthogonal group $O(2)$ is the group of orthogonal transformations. Orthogonal transformations have determinant 1 or -1. The ones with determinant -1 are the reflections. We shall focus on the orthogonal transformations with determinant 1. They form a subgroup called $SO(2)$ that can be identified with $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. The Haar measure on T is simply the uniform distribution U on $\mathbb{T}$that is equal to the Lebesgue measure divided with 2π. Assume that $\mathbb{T}$ acts irreducible on a state space $\mathfrak{S}$. We shall now find the exact shape of $\mathfrak{S}$ and classify the actions of $\mathbb{T}$ on $\mathfrak{S}$. One possibility is that $\mathbb{T}$ acts trivially on $\mathfrak{S}$, i.e. $a_\vartheta(\phi) = \phi$ for all $\vartheta \in \mathbb{T}$ and all $\phi \in \mathfrak{S}$. This is called the spin 0 representation. We will now assume that the representation is non-trivial, and choose $\phi \in \mathfrak{S}$ and $\vartheta \in \mathbb{T}$ such that $a_\vartheta(\phi) \neq \phi$. Then the set $\{a_\vartheta(\phi) \mid \vartheta \in \mathbb{T}\}$ spans the convex set $\mathfrak{S}$. For a probability measure

$\nu \in M_+^1(\mathbb{T})$ put

$$m(\nu) = \int a_\vartheta(\phi)\, d\nu\vartheta \ .$$

Then m is a map $M_+^1(\mathbb{T}) \to \mathfrak{S}$. If ν is the measure δ_{θ_0} with all its mass in the point ϑ_0 then

$$\begin{aligned} m(\nu) &= m(\delta_{\theta_0}) \\ &= \int a_\vartheta(\phi)\, d\delta_{\theta_0}\vartheta \\ &= a_{\vartheta_0}(\phi) \ . \end{aligned}$$

For $n \in \mathbb{N}$ and $(a,b) \in \mathbb{R}^2$ consider the function

$$f(\vartheta) = 1 + a\cos(n\vartheta) + b\sin(n\vartheta) \ .$$

Then f can be written as

$$f(\vartheta) = 1 + \begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} \cos(n\vartheta) \\ \sin(n\vartheta) \end{pmatrix},$$

and we see that f is a positive function if and only if $a^2 + b^2 \leq 1$. This gives a map $m_n : \Delta \to \mathfrak{S}$. Now we consider f as a probability density of a probability measure μ on $\mathbb{T}$. Then

$$\begin{aligned} a_\zeta(m_n(\mu)) &= a_\zeta\left(\int a_\vartheta(\phi)\ (1 + a\cos(n\vartheta) + b\sin(n\vartheta))\, dU\vartheta\right) \\ &= \int a_\zeta(a_\vartheta(\phi))\ (1 + a\cos(n\vartheta) + b\sin(n\vartheta))\, dU\vartheta \\ &= \int a_{\zeta+\vartheta}(\phi)\ (1 + a\cos(n\vartheta) + b\sin(n\vartheta))\, dU\vartheta \\ &= \int a_\vartheta(\phi)\ (1 + a\cos(n(\vartheta-\zeta)) + b\sin(n(\vartheta-\zeta)))\, dU\vartheta \ . \end{aligned}$$

Then we use that

$$\begin{aligned} &a\cos(n(\vartheta-\zeta)) + b\sin(n(\vartheta-\zeta)) \\ &\qquad = a(\cos(n\vartheta)\cos(n\zeta) + \sin(n\vartheta)\sin(n\zeta)) \\ &\qquad + b(\cos(n\vartheta)\sin(n\zeta) - \sin(n\vartheta)\cos(n\zeta)) \\ &\qquad\quad = (a\cos(n\zeta) + b\sin(n\zeta))\cos(n\vartheta) \\ &\qquad\quad + (a\sin(n\zeta) - b\cos(n\zeta))\sin(n\vartheta) \\ &\qquad\qquad = \begin{pmatrix} \cos(n\zeta) & \sin(n\zeta) \\ \sin(n\zeta) & -\cos(n\zeta) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} \cos(n\vartheta) \\ \sin(n\vartheta) \end{pmatrix}. \end{aligned}$$

Therefore the action of ϑ on $\mathfrak{S}$ gives a rotation of Δ by an angle $n\zeta$. This action of $\mathbb{T}$ on Δ is obviously irreducible. Therefore either $m_n(\Delta)$ spans $\mathfrak{S}$ or $m_n(\Delta)$ is a point.

Assume that $m_n(\Delta)$ is a point for all $n \in \mathbb{N}$. Then all functions of the form $1 + a\cos(n\vartheta) + b\sin(n\vartheta)$ are mapped into the same point. Then also functions of the form

$$1 + \sum_{n=1}^{k} a_n \cos(n\vartheta) + b_n \sin(n\vartheta)$$

are mapped into the same point in $\mathfrak{S}$. These functions are (weak) dense in the set of probability measures on $\mathbb{T}$ and therefore all probability measures on $\mathbb{T}$ are mapped into a point. Especially the probability measures with all mass in one point are mapped into a point and we see that $\mathbb{T}$ acts trivially on $\mathfrak{S}$. This is called the spin 0 representation of $\mathbb{T}$.Assume that $m_n(\Delta)$ spans $\mathfrak{S}$. Then $\mathfrak{S}$ can be identified with a subset of $\mathbb{R}^2$ spanned by Δ. The group $\mathbb{T}$ acts on Δ as rotation and there is a unique extension of this action to an action as rotations in $\mathbb{R}^2$. Now $\mathfrak{S}$ is bounded and closed so there exists $\phi \in \mathfrak{S}$ such that $\|\phi\|_2$ is maximal. Rotations of this state gives a circle in $\mathbb{R}^2$. Then $\mathfrak{S}$ contains no states outside the circle and by convexity $\mathfrak{S}$ contains all states inside the circle. After a suitable multiplication $\mathfrak{S}$ can be identified with Δ. As we have seen the unit disc can be identified with the set of density matrices on a real 2 dimensional Hilbert space.

A quantum mechanic system characterized by the number n is said to have spin $n/2$. To recognize a spin $n/2$ system one should do the following. First one should find a measurement which is sensitive to rotations. If no such measurement exists, the system has spin 0. If a rotation sensitive measurement has been found one should observe the effect of rotations. If ζ_0 is the smallest rotation such that a rotation by the angle ζ_0 gives the same measurement results as no rotation then the system has spin π/ζ_0. The table lists some quantum particles and their spin. The particle with integer spin are called bosons and the other are called fermions. Roughly speaking matter is composed of fermions and forces are are carried by bosons.

We should also remark that

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} 1/2+a & b \\ b & 1/2-a \end{pmatrix} \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} = \begin{pmatrix} 1/2 + a\cos 2t - b\sin 2t & a\sin 2t + b\cos 2 \\ a\sin 2t + b\cos 2t & 1/2 - a\cos 2t + b\si \end{pmatrix}$$

so that a rotation of Δ by an angle $2t$ imbedded a matrix acting on a Hilbert space corresponds to a rotation by t of the Hilbert space. We see that for a spin $n/2$ particle a rotation by ζ gives a rotation in the Hilbert space by $\zeta n/2$. For bosons $\zeta(n/2)$ is well defined but for fermions $n/2$ is not an integer and there is no unique way to divide by 2 in $\mathbb{T}$. To solve this problem we introduce a double covering of $\mathbb{T} \to \mathbb{T}$ via multiplication by 2.

Note that the unitary matrix

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

has eigenvalues e^{it} and e^{-it} so it is more instructive to consider the representa-

Symbol	Name	Spin
H^0	*Higgs particle*	0
π^0, π^+, π^-	pion	0
e	electron	1/2
ν_e	neutrino	1/2
μ	muon	1/2
τ	tau	1/2
p	proton	1/2
n	neutron	1/2
u, d, s, c, b, t	*quarks*	1/2
γ	photon	1
W^+, W^-	W-boson	1
Z_0	Z-boson	1
g	*gluon*	1
Δ	delta baryon	3/2
Ω^-	omega particle	3/2
	graviton	2

Table 1: A selection of elementary particles. The Higgs particle, the gluon and the graviton are hypothetical. The quarks cannot exists independently but there is a lot of experimental evidens for their existence.

tion

$$\begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix} \begin{pmatrix} 1/2 & b+ic \\ b-ic & 1/2 \end{pmatrix} \begin{pmatrix} e^{-it} & 0 \\ 0 & e^{it} \end{pmatrix} = \begin{pmatrix} 1/2 & e^{2it}(b+ic) \\ e^{-2it}(b-ic) & 1/2 \end{pmatrix}.$$

Until now only planar rotations have been discussed. Often it is possible to rotate a particle in 3 dimensions but not always. For instance it is only obvious how to rotate a photon if the axis of rotation is parallel to its velocity vector. Irreducible representations of rotations in 3 dimensions are more complicated than in two dimensions, but we can show the state space of irreducible representations can be imbedded in the set of density matrices on a complex Hilbert space.

We shall see the relation to representations in 3 dimensions i.e. representations of $SO(3)$. The group $SU(2)$ acts on the Bloch sphere and therefore we get a group homomorphism $\pi : SU(2) \to SO(3)$. Let U_1 and U_2 be special unitary operators in $SU(2)$ leading to the same rotation of the Bloch sphere. Then $U_1 = \alpha U_2$. Taking the determinant gives $\alpha^2 = 1$. Therefore $\alpha = \pm 1$, and $U_1 = \pm U_2$. Therefore the map $\pi : SU(2) \to SO(3)$ is 2 to 1. Any projective representation $\rho : SO(3) \to U(n)$ gives a projective representation $\rho \circ \pi : SU(2) \to U(n)$. A representation $\tau : SU(2) \to U(n)$ gives a unitary representation of $SO(3)$ if and only if $\tau(-1) = 1$.

We introduce the matrices

$$\mathbf{i} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$
$$\mathbf{j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
$$\mathbf{k} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Then

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$$
$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k}$$
$$\mathbf{jk} = -\mathbf{kj} = \mathbf{i}$$
$$\mathbf{ki} = -\mathbf{ik} = \mathbf{j}\ .$$

The matrix $U = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ is has the form

$$\begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}$$

and the determinant $a^2 + b^2 + c^2 + d^2$. The adjoint is $U^* = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$. Therefore

$$\begin{aligned} UU^* &= U^*U \\ &= a^2 + b^2 + c^2 + d^2. \end{aligned}$$

Hence $U \in SU(2)$ if and only if $a^2 + b^2 + c^2 + d^2 = 1$. It is easy to check that all elements in $SU(2)$ are of this form. Therefore $SU(2)$ has the same topology as a sphere in 4 dimensions.

Example 16 *Now put $z_1 = a+ib$ and $z_2 = c+id$. Then $SU(2)$ can be identified with the vectors (z_1, z_2) with $|z_1|^2 + |z_2|^2 = 1$.The group $SU(2)$ act on the set $\left\{(z_1, z_2) \in \mathbb{C}^2 \mid |z_1|^2 + |z_2|^2 = 1\right\}$ via*

$$(z_1, z_2) \to (u_{11}z_1 + u_{21}z_2, u_{12}z_1 + u_{22}z_2)$$

where

$$U = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \in SU(2).$$

If $U = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ and $(z_1, z_2) = (1, 0)$ then (z_1, z_2) is mapped into $(u_{11}, u_{12}) = (a + ib, c + id)$.

Let m be a nonnegative integer. Let $\mathcal{H}_m$ be the linear space of homogeneous polynomials of degree m in two complex variables z_1 and z_2 provided with the scalar product

$$(p \mid q) = \int_{|z_1|^2+|z_2|^2=1} p(z_1, z_2)\, \overline{q(z_1, z_2)} d\mu(z_1, z_2)$$

where μ denotes the Haar measure on $SU(2)$, i.e. the uniform distribution on a sphere. Then $\mathcal{H}_m$ is a Hilbert space of dimension $m+1$. Left translation of a polynomial is given by

$$(T_m(u)\,p)(z_1,z_2) = p(u_{11}z_1+u_{21}z_2, u_{12}z_1+u_{22}z_2)$$

where

$$u = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \in SU(2).$$

We see that $\mathcal{H}_m$ is invariant under left translation so for each m we have a unitary representation. One can show that the representations are irreducible and that any L^2-function on $SU(2)$ can be written as a sum of homogeneous polynomials so any irreducible representation is isomorphic to the representation on $\mathcal{H}_i$ for some i.

The group $SU(2)$ has the same topology as the sphere in 4 dimensions which is simply connected. Therefore for any projective (finite dimensional) representation of $SU(2)$ is equal to a unitary representation times a complex function of modulus 1. Therefore we are interested in the unitary representations of $SU(2)$.

The unitary matrix

$$\begin{pmatrix} \exp(i\theta) & 0 \\ 0 & \exp(-i\theta) \end{pmatrix}$$

corresponds to a rotation by 2θ in $SO(2)$. Let P_k be the homogeneous polynomial $P_k(z_1,z_2) = z_1^{m-k}z_2^k$ $(k=0,1,...,m)$. Then

$$\begin{aligned} T_m(u_\theta)\,P_k(z_1,z_2) &= (\exp(i\theta)\,z_1)^{m-k}(\exp(-i\theta)\,z_2)^k \\ &= \exp(i(m-2k)\theta)\,P_k(z_1,z_2)\ . \end{aligned}$$

On $span(P_0,P_m)$ the unitary $T_m(u_\theta)$ has the matrix

$$\begin{pmatrix} \exp(im\theta) & 0 \\ 0 & \exp(-im\theta) \end{pmatrix}$$

so the restriction of the representation of $SU(2)$ to $span(P_0,P_m)$ is a spin $m/2$ representation of $SO(2)$. The m-dimensional representation of $SU(2)$ is therefore called the *spin* $m/2$ representation.

It is possible to show any irreducible unitary representation of $SU(2)$ is isomorphic to a spin $m/2$ representation for some nonnegative integer m. Let U_g denote a unitary representation of $SU(2)$. Then the representation on the density matrices is given by $S \to U_g S U_g^*$ which is an unitary transformation in the Hilbert space of matrices with the inner product

$$(S \mid T) = Tr(ST^*).$$

It can be considered as the tensor product of the representation U_g with it self. The character of an unitary representation is defied as

$$\chi(g) = Tr(U_g)$$

The character satisfies that if U and V are unitary representations then

$$\begin{aligned}\chi_{U\oplus V}(g) &= \chi(U_g) + \chi(V_g)\\ \chi_{U\otimes V}(g) &= \chi(U_g)\,\chi(V_g)\\ \chi_{U^*}(g) &= \overline{\chi_U(g)}\ .\end{aligned}$$

For the spin representations we have

$$\chi_{T_m}(g)\,\chi_{T_n}(g) = \chi_{T_{n-m}}(g) + \chi_{T_{n-m+2}}(g) + ... + \chi_{T_{n+m-2}}(g) + \chi_{T_{n+m}}(g)\,.$$

In particular

$$\begin{aligned}\chi_{T_m}(g)\,\chi_{T_m^*}(g) &= \chi_{T_m}(g)\,\overline{\chi_{T_m}(g)}\\ &= (\chi_{T_m}(g))^2\\ &= \sum_{j=0}^{m}\chi_{T_{2j}}(g)\ .\end{aligned}$$

Therefore the complex unitary representation T_m induces all orthogonal real representation of even order up to $2m$. Similarly

$$\chi_{T_m\oplus 1}(g)\,\chi_{(T_m\oplus 1)^*}(g) = \chi_{T_m}(g)\,\chi_{T_m^*}(g) + \chi_{T_m}(g) + \chi_{T_m^*}(g) + 1$$

so the compels unitary representation $T_m \oplus 1$ induces the the orthogonal real representation corresponding to T_m.

3.3 Superposition

Once again the simplest example with rotations in 2 dimensions will be treated. Consider a quantum mechanical system containing 2 independent preparations which can be rotated independently around a given axis. The group $\mathbb{T}^2$ acts on the set of preparations and we will assume that it also acts on the state space. We search irreducible representations of $\mathbb{T}^2$. The group $\mathbb{T}^2$ is commutative and connected so an irreducible representation is trivial or the state space is isomorphic to Δ, and a rotation by the angles α, β corresponds to a rotation of the unit disc by an angle $k\alpha + l\beta$ where $(k,l) \in \mathbb{Z}^2$. We will concentrate on the case $(k,l) = (1,-1)$. If one of the angles is fixed we get a spin-1/2 representation of the other. The standard interpretations is that one has made a preparation of 2 spin-1/2 particles. Now the system is equipped with a detector which can give the results *detected* or *not detected*. Like in the previous section the measurement can be written as a convex combination of a measurement not depending of the state and a measurement giving the result with certainty for some states. We will assume that the detector is of the last mentioned type.

Let ϕ be a pure state which with certainty gives the result detection. Then

$$\begin{aligned} P\left(\text{detection} \mid \phi \text{ rotated angles } \alpha \text{ and } \beta\right) &= \frac{\cos\left(\alpha-\beta\right)+1}{2} \\ &= \frac{\cos\left(\alpha\right)\cos\left(\beta\right)+\sin\left(\alpha\right)\sin\left(\beta\right)+1}{2} \\ &= \frac{\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}\cdot\begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix}+1}{2} \\ &= \frac{\left(\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}+\begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix}\right)^2}{4} . \end{aligned}$$

The standard interpretation is that a particle in a state given by the vector $\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$ interferes with another particle in state given by the vector $\begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix}$, or even worse if the states are not in phase, that the particles get extinct by each other. This kind of language usage has been a source to many paradoxes.

In general the superposition principle can be explained using representations of products of groups with them selves.

4 Algebras

4.1 The group algebra

Let G be a finite group with n elements and neutral element e. The a composition on $C\left(G\right)$ is defined by

$$\left(f*g\right)\left(x\right)=\sum_{y\in G} f\left(xy^{-1}\right)g\left(y\right) \ .$$

A convolution is defined by

$$f^{*}\left(x\right)=\overline{f\left(x^{-1}\right)}$$

where $\overline{\ }$ denotes complex conjugation. With this structure $C\left(G\right)$ is a so-called finite dimensional $*$-algebra called the group algebra of G. Thus symmetry groups leads to the matrix algebras in a natural way.

Define

$$\phi\left(f\right)=f\left(e\right)$$

and note that $\phi(1) = 1$. We also have

$$\begin{aligned}\phi(f * f^*) &= (f * f^*)(e) \\ &= \sum_{y \in G} f\left(ey^{-1}\right) f^*(y) \\ &= \sum_{y \in G} f\left(y^{-1}\right) \overline{f\left(y^{-1}\right)} \\ &= \sum_{y \in G} |f(y)|^2 \\ &\geq 0\end{aligned}$$

with equality if and only if $f = 0$. A function ϕ on an algebra satisfying these conditions is called a faithful state. Here one should note that $(f * f^*)(x)$ may be negative for some values of x.

Example 17 *The group $\mathbb{Z}_4 = \mathbb{Z}/4\mathbb{Z}$ is an example of a commutative group. For $n = 0, 1, 2, 3$ consider the function $f_n : j \to i^{nj}$, $j \in \mathbb{Z}_4$ where i is the imaginary unit. These functions forms a basis for $C(\mathbb{Z}_4)$. If $m \neq n$ then $f_m * f_n = 0$. Therefore each of the functions f_n generate a 1 dimensional sub-algebra of $C(\mathbb{Z}_4)$, and $C(\mathbb{Z}_4)$ is isomorphic to a sum of these algebras.*

There exists groups that have isomorphic group algebras although the groups are not isomorphic. Therefore part of the structure of the group is lost by forming the group algebra, but often what is lost is not important for the applications we have in mind and loosing irrelevant structure is what sometimes makes life easier.

4.2 $*-$Algebras

In order to describe quantum systems in more details we will introduce as an axiom that the state space of a quantum system geometrically has the same shape as the set of states of complex $*$-algebras. In this section we shall study certain sets of matrices and their algebraic properties. Both the Hilbert spaces and the algebras are over the field of complex numbers. Hilbert spaces and algebras over other fields than $\mathbb{C}$ will not be discussed, and it is important to realize that many of the results cannot be generalized to other fields.

4.3 Spectral theory

Let $\mathcal{H}$ be a finite dimensional complex Hilbert space. The linear maps $\mathcal{H} \to \mathcal{H}$ will be called *operators*. The set of operators on $\mathcal{H}$ will be denoted $\mathbb{B}(\mathcal{H})$. The characteristic polynomial of X is

$$P_{char}(\lambda) = \det(X - \lambda I)$$

Then $\lambda \in \mathbb{C}$ is an eigenvalue if and only if $P_{char}(\lambda) = 0$ where P_{char} is the characteristic polynomial of X. The set of eigenvalues of X is called the spectrum

of X and is denoted $Sp(X)$. A complex polynomial of degree d has at least one root and at most d roots. Therefore $Sp(X)$ is a finite non-empty set. Let P denote a polynomial given by

$$P(z) = a_n z^n + a_n z^{n-1} + ... + a_1 z + a_0.$$

Then $P(X)$ is defined by

$$P(X) = a_n X^n + a_n X^{n-1} + ... + a_1 X + a_0.$$

If X is self adjoined, i.e. $X = X^*$ then the eigenvectors are orthogonal, and using the eigenvectors as basis X can be written as

$$\begin{pmatrix} \lambda_1 & 0 & \cdots \\ 0 & \lambda_2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

with the eigenvalues in the diagonal. Then $P(X)$ can be written as

$$\begin{pmatrix} P(\lambda_1) & 0 & \cdots \\ 0 & P(\lambda_2) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

Let f be a complex function with domain $Sp(X)$ where X is self adjoined. Then $f(X)$ is defined as the linear map with matrix

$$\begin{pmatrix} f(\lambda_1) & 0 & \cdots \\ 0 & f(\lambda_2) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

In this way to any function with domain $Sp(X)$ an operator is associated. We see that $(f+g)(X) = f(X) + g(X)$ and $(f \cdot g)(X) = f(X) \cdot g(X)$.

4.4 $*$-algebras and their decomposition

Definition 18 *A finite dimensional $*$-algebra $\mathcal{A}$ over the complex numbers is a finite dimensional complex vector space over C equipped with a composition $\cdot : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ and and a map $^*: \mathcal{A} \to \mathcal{A}$ such that the following identities holds*

1. $X \cdot (Y \cdot Z) = (X \cdot Y) \cdot Z$ *for all* $X, Y, Z \in \mathcal{A}$.
2. $X \cdot (Y + Z) = X \cdot Y + X \cdot Z$ *for all* $X, Y, Z \in \mathcal{A}$.
3. $\lambda (X \cdot Y) = \lambda X \cdot Y = X \cdot \lambda Y$ *for all* $\lambda \in \mathbb{C}, X, Y \in \mathcal{A}$.
4. $(X^*)^* = X$ *for all* $X \in \mathcal{A}$.
5. $(X + Y)^* = X^* + Y^*$ *for all* $X, Y \in \mathcal{A}$.

6. $(X \cdot Y)^* = Y^* \cdot X^*$ *for all* $X, Y \in \mathcal{A}$.

7. $(\lambda X)^* = \bar{\lambda} X^*$ *for all* $\lambda \in \mathbb{C}, X \in \mathcal{A}$.

The elements in a $*$*-algebra are called* operators. *A* $*$*-algebra is commutative if* $X \cdot Y = Y \cdot X$ *for all* $X, Y \in \mathcal{A}$.

The algebra is said to have a *unit* (or to be *unital*) if there exists an operator E such that E is a neutral element with respect to multiplication. A standard argument proves that a $*$-algebra can have at most one unit. Unless directly stated the $*$-algebras in these notes are always assumed to be unital and the unit will be denoted $\mathbf{1}$.

Denote by $\mathbb{B}(\mathcal{H})$ the set of linear maps $\mathcal{H} \to \mathcal{H}$. It is easy to check that $\mathbb{B}(\mathcal{H})$ is a finite dimensional $*$-algebra. If $\mathcal{H}$ has dimension d then $\mathbb{B}(\mathcal{H})$ can be identified with the matrix algebra of $d \times d$ complex matrices. Let $C(U)$ denote the set of complex function $U \to \mathbb{C}$ where U is a finite set. Then $C(U)$ is a commutative $*$-algebra with the normal addition, multiplication and conjugation. The rest of this section is devoted to a classification of all finite $*$-algebras, and the result will be that $\mathbb{B}(\mathcal{H})$ is the most non-commutative algebra and any algebra is somewhere in between algebras like $\mathbb{B}(\mathcal{H})$ and $C(U)$.

For a self adjoined operator $X \in \mathbb{B}(\mathcal{H})$ the algebra $C(Sp(X))$ is the smallest $*$-algebra in $\mathbb{B}(\mathcal{H})$ containing X.

An operator X in a $*$-algebra is said to be self adjoint if $X = X^*$. An operator is said to be positive if there exists an operator Y such that $X = Y^*Y$. The positive operators are automatically self adjoint.

Definition 19 *An element* P *in the* $*$*-algebra* $\mathcal{A}$ *is called an orthogonal projection if* P *is self adjoined and* $P^2 = P$.

If P is an orthogonal projection then $\mathbf{1} - P$ is also a projection. The set of operators in $\mathcal{A}$ which commutes with P is a sub-algebra $\mathcal{B}$ of $\mathcal{A}$. Now

$$X \curvearrowright PXP + (1 - P) X (1 - P)$$

is a projection of $\mathcal{A}$ into $\mathcal{B}$. The set of operators P and $\mathbf{1} - P$ is an example of a resolution of the identity which will now be defined in full generality.

Definition 20 *Let* $X_1, X_2, ..., X_n$ *be a set of operators in a* $*$*-algebra. Then* $X_1, X_2, ..., X_n$ *is said to be a resolution of the identity if all operators are positive and*

$$\sum_{i=1}^{n} X_i = \mathbf{1}.$$

If $X_i X_j = 0$ *for* $i \neq j$ *the the resolution of the identity is said to be orthogonal.*

Let $P_1, P_2, ..., P_n$ be an orthogonal resolution of the identity. Then

$$\begin{aligned} X_i &= X_i \cdot \mathbf{1} \\ &= X_i \cdot \sum_{i=1}^{n} X_i \\ &= X_i^2 \end{aligned}$$

for all $i = 1, 2, ..., n$, and all P_i are projections. For all $X \in \mathcal{A}$ define $\mathbb{E}(X)$ by

$$\mathbb{E}(X) = \sum_{i=1}^{n} P_i X P_i \ .$$

Then $\mathbb{E}$ is a projection of $\mathcal{A}$ into the a sub-algebra $\mathcal{B}$ consisting of operators which commute with all P_i.

Let $\mathcal{A}$ and $\mathcal{B}$ be $*$-algebras. Then the sum of the algebras is denoted $\mathcal{A} \oplus \mathcal{B}$ where addition and multiplication is defined componentwise. Let P be a projection in $\mathcal{A}$ commuting with all operators in $\mathcal{A}$. Then $P\mathcal{A}$ and $(1-P)\mathcal{A}$ are sub-algebras of $\mathcal{A}$. and the algebra $\mathcal{A}$ is isomorphic to $P\mathcal{A} \oplus (1-P)\mathcal{A}$. A $*$-algebra is said to be simple if it can not be written as a non-trivial sum of $*$-algebras. We see that a simple algebra has no projections except 0 and $\mathbf{1}$.

Definition 21 *Let ϕ be a linear map $\mathcal{A} \to \mathbb{C}$. Then ϕ is called a* state *if $X \geq 0$ implies $\phi(X) \geq 0$, and $\phi(\mathbf{1}) = 1$. The extreme states are called pure states. If $\phi(X^*X) = 0$ implies $X = 0$ then ϕ is said to be* faithful.

Example 22 *Let U be a finite set and $C(U)$ the $*$-algebra of functions $U \to \mathbb{C}$. Then for any probability vector $(p_1, p_2, ..., p_n)$ on U a state is given by*

$$f \to \sum_{i=1}^{n} f(i)\, p_i$$

which is the mean value of the random variable f. The state is faithful if and only if $p_i > 0$ for $i = 1, 2, ..., n$.

Theorem 23 *Let ϕ be a state on the finite dimensional $*$-algebra $\mathbb{B}(\mathcal{H})$. Then there exists a positive operator $S_\phi \in \mathbb{B}(\mathcal{H})$ with $Tr(S_\phi) = 1$ such that*

$$\phi(X) = Tr(XS_\phi)$$

for all X in $\mathbb{B}(\mathcal{H})$. A state is pure if and only if there exists a vector $\vec{u} \in \mathcal{H}$ such that

$$\phi(X) = \langle \vec{u} \mid X\vec{u} \rangle .$$

The state ϕ is faithful if and only if 0 is not an eigenvalue of S_ϕ.

Proof. First we observe that the formula

$$(X \mid Y) = Tr\,(XY^*)$$

defines an inner product on $\mathbb{B}(\mathcal{H})$. Therefore there exists an operator S in $\mathbb{B}(\mathcal{H})$ such that

$$\begin{aligned}\phi(X) &= (X \mid S) \\ &= Tr\,(XS^*).\end{aligned}$$

For any vector $\vec{v}$ in $\mathcal{H}$ a linear map is defined by $\vec{u} \to \langle \vec{u} \mid \vec{v} \rangle \cdot \vec{v}$. The operator X of this map is positive. Then the operator XS^* gives a linear map $\vec{u} \to \langle S^*\vec{u} \mid \vec{v} \rangle \cdot \vec{v} = \langle \vec{u} \mid S\vec{v} \rangle \cdot \vec{v}$, and the trace of XS^* is $\langle \vec{v} \mid S\vec{v} \rangle$. Therefore $\langle \vec{v} \mid S\vec{v} \rangle \geq 0$ for all $\vec{v} \in \mathcal{H}$, and S is a positive operator.

A state is given by a positive operator S_ϕ which has trace one. Choosing a suitable basis it can be written as

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_d \end{pmatrix}$$

where $\lambda_i \geq 0$ and $\sum \lambda_i = 1$. Matrices of this form is a simplex and the extreme points are matrices where one of the eigenvalues is 1 and the others are 0. If ϕ is extreme choose $\vec{u}$ as an eigenvector of S_ϕ with eigenvalue 1. ∎

The operator S defined in the theorem is the density operator S_ϕ corresponding to the state ϕ.

Let X be an operator in $\mathcal{A}_1 \oplus \mathcal{A}_2$, and let P_1 and P_2 be projections on $\mathcal{A}_1$ and $\mathcal{A}_2$. Let ϕ be a state on $\mathcal{A}_1 \oplus \mathcal{A}_2$ such that $\phi(P_i) \neq 0$. Then

$$\begin{aligned}\phi(X) &= \phi(P_1X + P_2X) \\ &= \phi(P_1X) + \phi(P_2X) \\ &= \phi(P_1) \cdot \frac{\phi(P_1X)}{\phi(P_1)} + \phi(P_2) \cdot \frac{\phi(P_2X)}{\phi(P_2)}.\end{aligned}$$

Remark that

$$X \to \frac{\phi(P_iX)}{\phi(P_i)}$$

is a state on $\mathcal{A}_i$. In normal probability theory this is the well-known Bayes' formula. In our more general setup it tells that the states on a sum of algebras is a mixture of states on the individual algebras. To find the states on a $*$-algebra we have to write the algebra as a sum of algebras where the set of states is known.

A *representation* of a finite dimensional $*$-algebra $\mathcal{A}$ is a pair $(\pi, \mathcal{H})$ where $\mathcal{H}$ is a Hilbert space and $\pi : \mathcal{A} \to \mathbb{B}(\mathcal{H})$ is a $*$-homomorphism. The representation is said to be *cyclic* if there exist a (*cyclic*) vector $\vec{v}$ in $\mathcal{H}$ such that $\{\pi(X)\,\vec{v} \mid X \in \mathcal{A}\} = \mathcal{H}$. The following theorem is known as the Gelfand-Naimark-Segal construction.

Theorem 24 (Gelfand-Naimark-Segal construction) *Let $\mathcal{A}$ be a finite dimensional $*$-algebra with a faithful state ϕ. Then there exists a cyclic representation $(\pi_\phi, \mathcal{H}_\phi, \vec{v}_\phi)$ such that*

$$\phi(X) = (\pi_\phi(X)\,\vec{v}_\phi \mid \vec{v}_\phi) \ .$$

The $*$-homomorphism π_ϕ is injective.

Proof. Put $\mathcal{H}_\phi = \mathcal{A}$ equipped with the following inner product

$$(X \mid Y) = \phi(Y^* X) .$$

The $*$-homomorphism is given by

$$\pi_\phi(X)\,Y = XY.$$

Then

$$\begin{aligned}(\pi_\phi(X)\,Y \mid Z) &= (XY \mid Z)\\ &= \phi(Z^* XY)\\ &= \phi\left((X^* Z)^* Y\right)\\ &= (Y \mid X^* Z)\\ &= (Y \mid \pi_\phi(X^*)\,Z) \ .\end{aligned}$$

Put $\vec{v}_\phi = \pi_\phi(\mathbf{1})$. Then

$$\begin{aligned}(\pi_\phi(X)\,\vec{v}_\phi \mid \vec{v}_\phi) &= \phi(\mathbf{1}^* X \mathbf{1})\\ &= \phi(X) .\end{aligned}$$

The vector $\vec{v}_\phi$ is cyclic because

$$\begin{aligned}\pi_\phi(X)\,\vec{v}_\phi &= X\mathbf{1}\\ &= X \ .\end{aligned}$$

The $*$-homomorphism π_ϕ is injective because $\pi_\phi(X)\,\vec{v}_\phi = \pi_\phi(Y)\,\vec{v}_\phi$ if and only if $X = Y$. ■

A representation $\pi : \mathcal{A} \to \mathbb{B}(\mathcal{H})$ is said to be irreducible if the only subspaces that are invariant under the action of $\mathcal{A}$ are the trivial subspaces.

Theorem 25 *Any representation $\pi : \mathcal{A} \to \mathbb{B}(\mathcal{H})$ of a finite $*$-algebra is a sum of irreducible representations.*

Let K be an invariant subspace of H. Then for $u \in K$ and $v \in K^\perp$ we have

$$\begin{aligned}(u \mid \pi(X)\,v) &= \left(\pi(X)^*\,u \mid v\right)\\ &= (\pi(X^*)\,u \mid v) = 0\end{aligned}$$

because $X^* \in \mathcal{A}$ and K is invariant under the action of $\mathcal{A}$. Therefore $K^\perp$ is also invariant under the action of $\mathcal{A}$. Let P denote the projection of H on K q the projection of $K^\perp$. Then

$$\begin{aligned}\pi(X) &= (P+Q)\,\pi(X)\,(P+Q)\\ &= P\pi(X)\,P + P\pi(X)\,Q + Q\pi(X)\,P + Q\pi(X)\,Q\\ &= P\pi(X)\,P + Q\pi(X)\,Q.\end{aligned}$$

Now we note that $X \to P\pi(X)\,P$ and $X \to Q\pi(X)\,Q$ are representations. In this way we can decompose a representation until we have a sum of irreducible representations.

A representation is said to be transitive if any vector different from $\vec{0}$ is cyclic. It is straight forward to check that a representation is irreducible if and only if it is transitive.

Theorem 26 (Burnside's Theorem) *Any irreducible representation $\pi : \mathcal{A} \to \mathbb{B}(\mathcal{H})$ is surjective.*

The proof given here builds on the master thesis of Rune Johansen.

Proof. The proof is by induction on the dimension of $\mathcal{H}$. for $\dim(\mathcal{H}) = 1$ the result is obvious.

For the induction step assume that the theorem holds for any Hilbert spaces W with $\dim(W) \leq n$. Let $\dim(H) = n+1$. If $\pi(\mathcal{A}) = \mathbb{C}$ then π is not transitive. Let $X \in \pi(\mathcal{A})/\mathbb{C}$ be some matrix. Then X has an eigenvalue λ and define $F = X - \lambda$. Note that $F \neq 0$ and that F is not invertible. Thus the range of F is a subspace of $\mathcal{H}$ that shall be denoted W. An algebra $\mathcal{B}$ is defined by

$$\mathcal{B} = \left\{ FX_{|W} \mid A \in \pi(A) \right\}.$$

Let $\vec{x}$ be any vector in W. Then

$$\begin{aligned}\mathcal{B}x &= \{Bx \mid B \in \mathcal{B}\}\\ &= \left\{FA_{|W}\vec{x} \mid X \in \pi(A)\right\}\\ &= F\{A\vec{x} \mid X \in \pi(A)\}\\ &= F(\mathcal{H})\\ &= W.\end{aligned}$$

Thus $\mathcal{B}$ is transitive on W and according to the induction hypothesis $\mathcal{B} = \mathbb{B}(W)$. In particular $\mathcal{B}$ contains a one dimensional projection P on some vector $\vec{u}$. This is given by $P(\vec{x}) = (\vec{x} \mid \vec{u}) \cdot \vec{u}$. Now

$$\begin{aligned}\pi(A_1)\,P\pi(A_2)\,(\vec{x}) &= (\pi(A_2)\,\vec{x} \mid \vec{u}) \cdot \pi(A_1)\,\vec{u}\\ &= (\vec{x} \mid \pi(A_2^*)\,\vec{u}) \cdot \pi(A_1)\,\vec{u}\end{aligned}$$

Using transitivity of π we see that $\pi(A)$ contains all mappings of the for $x \to (x \mid v) \cdot w$ and therefore $\pi(A) = \mathbb{B}(\mathcal{H})$. ■

Theorem 27 *Let $\mathcal{A}$ be a finite dimensional unital $*$-algebra with a faithful state ϕ. Then $\mathcal{A}$ is isomorphic to a sum of matrix algebras. If A is commutative then $\mathcal{A}$ is isomorphic to $C(U)$.*

Example 28 *Let G be the set of permutations of the elements of $\{1,2,3\}$. This group has 6 elements so the group algebra is 6 dimensional. A matrix algebra of dimension less than or equal to 6 is either of dimension 1 or 4. If the group algebra was a sum of 1 dimensional matrix algebras then $C(G)$ and G would be commutative. Therefore the group algebra is a sum of 2 one dimensional algebras and one four dimensional algebra. A constant function $x \to 1$ commutes with any other element $f \in G$ because*

$$\begin{aligned}(f*1)(x) &= \sum_{y\in G} f\left(xy^{-1}\right)1 \\ &= \sum_{y\in G} 1f(y) \\ &= (1*f)(x) \ .\end{aligned}$$

Therefore the constant functions form a 1 dimensional sub-algebra of the group algebra. Similarly the function which is 1 on the even permutations and -1 on the odd permutations commutes with any other element. Therefore this function also generates a 1 dimensional sub-algebra of the group algebra. functions which are orthogonal to these 2 functions then forms a sub-algebra of the group algebra isomorphic to a matrix algebra of 2×2 matrices.

Using the decomposition of an a algebra into matrix algebras it follows that any state ψ on $\mathcal{A}$ is given by a positive operator $S_\psi \in \mathcal{A}$ with $Tr(S_\psi) = 1$ such that $\psi(X) = Tr(S_\psi X)$, where the trace denotes the sum of the traces on the matrix algebras. Until now we have seen density operators with respect to the trace. Similarly if $S \geq 0$ and $\phi(S) = 1$ then the map

$$X \to \phi(SX)$$

is a state. All states are of this form if and only if ϕ if faithful.

4.5 Measurements in $*$-algebras

From now on we will assume that the state space can be represented by the set of states on a $*$-algebra. There are several reasons to choose $*$-algebras. First of all they seem to include all examples in mechanics (classic, quantum mechanic or statistical mechanic). Secondly the category of $*$-algebras and homeomorphisms is nice in the sense that one can form sum, tensor product and limes inside the category. Further one can perform the construction crossed product inside the category, where symmetries on the algebra are expressed as unitary operators in an extended algebra. This illustrates that the commutative part of the algebra has its origin in a commutative algebra of random variables and the non commutative properties reflects symmetries. Our next goal is to

make representations of the set of measurements. According to the general definitions a measurement with values in U is given by an affine mapping from the state space into the set of probability distributions on U.

Let $X_{u_1}, X_{u_2}, ..., X_{u_n}$ be a resolution of the identity with $U = \{u_1, u_2, ..., u_n\}$. For $B \subseteq U$ put

$$M_B = \sum_{u \in B} X_u \ .$$

Then the map $B \to M_B$ is called a *positive operator valued measure* ($POVM$). POVMs corresponding to orthogonal resolutions of the identity are sometimes called *spectral measures* and it is possible to integrate with respect to such spectral measures.

Theorem 29 *Let $\mathcal{A}$ be a finite dimensional unital $*$-algebra with a faithful state ϕ. The relation*

$$\begin{aligned} \mu_\psi(B) &= \psi(M_B) \\ &= Tr(S_\psi M_B), \ B \subseteq U \end{aligned}$$

establish a bijective correspondence between measurements on the state space $\mathbb{S}(\mathcal{A})$ and POVMs M_B in $\mathcal{A}$.

Proof. Let M_B be a POVM. Then it is easy to check that

$$\mu_\psi(B) = \phi(S_\psi M_B), \ B \subseteq U$$

defines a measurement

For any $u \in U$ the probability $\mu_\phi(u)$ is a linear function of ϕ. Therefore there exists a positive operator $X_u \in \mathcal{A}$ such that

$$\begin{aligned} \mu_\psi(u) &= (S_\psi \mid X_u) \\ &= Tr(S_\psi X_u^*) \ . \end{aligned}$$

If ϕ is a pure state corresponding to the vector $\vec{v} \in \mathcal{H}$, then $\mu_\psi(u) = (\vec{v} \mid X_u \vec{v}) \geq 0$. Therefore $X_u \geq 0$. For $B \subseteq U$ we have

$$\begin{aligned} \mu_\psi(B) &= \sum_{u \in B} \mu_\psi(u) \\ &= \sum_{u \in B} \phi(S_\psi X_u) \\ &= \phi\left(S_\psi \sum_{u \in B} X_u\right) \\ &= \phi(S_\psi M_B) \ . \end{aligned}$$

■

Example 30 *In a commutative $*$-algebra $C(U)$ the simple measurements corresponds to partitions of U.*

Let μ be a simple measurement. Then $\phi \to Tr\left(S_\phi M_B\right)$ is extreme in the set functionals $\mathbb{S}\left(\mathcal{A}\right) \to \left[0;1\right]$. All operators M_B satisfy $0 \leq M_B \leq \mathbf{1}$, i.e. all eigenvalues of M_B are in $\left[0;1\right]$. The extreme operators satisfying $0 \leq M_B \leq \mathbf{1}$ can only have eigenvalues 0 and 1, and therefore M_B is a projection. Then the corresponding resolution of the identity is orthogonal.

Theorem 31 *There is a unique correspondence between simple measurements μ with values in $\mathbb{R}$ and self adjoined operators X_μ such that*

$$\sum_{u \in U} f\left(u\right) \mu_\phi\left(u\right) = \phi\left(f\left(X_\mu\right)\right)$$

Proof. Let M_u, $u \in \mathbb{R}$ be the orthogonal resolution of the identity corresponding to μ. Put $X_\mu = \sum u \cdot M_u$. Then X_μ is a self adjoined matrix and according to the spectral theorem the correspondence is uniquely determined. ■

A simple measurement with values in $\mathbb{R}$ is called an *observable*. The theorem tells that observables can be identified with self adjoined operators. Let X be an observable, If the state is ϕ then the mean of X is $\phi\left(X\right)$ and the variance of X is

$$Var\left(X\right) = \phi\left(\left(X - \phi\left(X\right)\right)^2\right) \ .$$

Let X and Y be self adjoined operators. Then the *commutator* of X and Y is given by $\left[X,Y\right] = XY - YX$. The following theorem is a version of *Heisenberg's uncertainty relation.*

Theorem 32 *Let X and Y be self adjoined operators in a $*$algebra $\mathcal{A}$ with state ϕ. 'Then*

$$Var\left(X\right) Var\left(Y\right) \geq \frac{1}{4} \phi\left(i\left[X,Y\right]\right)^2 .$$

Proof. First remark that $Var\left(X\right) = Var\left(X - \phi\left(X\right)\right)$ and

$$\begin{aligned}
\left[X - \phi\left(X\right), Y\right] &= \left(X - \phi\left(X\right)\right) Y - Y\left(X - \phi\left(X\right)\right) \\
&= XY - YX - \phi\left(X\right) Y + X\phi\left(Y\right) \\
&= XY - YX \\
&= \left[X,Y\right] \ .
\end{aligned}$$

Therefore X can be replaced by $X - \phi\left(X\right)$ in the formula we have to prove. Similarly Y can be replaced by $Y - \phi\left(Y\right)$. Then the algebra can be equipped with an inner product

$$\left(X \mid Y\right) = \phi\left(XY^*\right).$$

Then for all $c \in \mathbb{R}$ we have

$$\begin{aligned}
0 &\leq \left\|X - icY\right\|^2 \\
&= Var\left(X\right) + c^2 Var\left(Y\right) + c\phi\left(i\left[X,Y\right]\right).
\end{aligned}$$

Therefore the discriminant is positive and we have

$$\left(\phi\left(i\left[X,Y\right]\right)\right)^{2}-4Var\left(X\right)Var\left(Y\right)\geq 0$$

■

Orthogonal resolutions of the identity are sometimes called spectral measures and it is possible to integrate with respect to such spectral measures.

Theorem 33 *If M_B is an orthogonal resolution of the identity then $M_BM_C=0$ if $B\cap C=\varnothing$ and $M_B^2=M_B$. All extreme measurements are simple if and only if U has only 2 elements or the algebra is commutative.*

Proof. Using that M_U is extreme M_U is a projection and therefore $M_U^2=M_U$. Then $M_U\left(1-M_U\right)M_U=\left(M_U^2-M_U\right)M_U=0$ and for M_V we get $0\leq M_V\leq 1-M_U$ and further $M_UM_VM_U=0$. This proves that $M_UM_V^2M_U=M_UM_V\left(M_UM_V\right)^{*}=0$. Therefore $M_UM_V=0$. If U has 2 elements then a measurement is given by 2 operators M and M' where $M'=1-M$. Therefore the set of measurements is isomorphic to the set of operators between 0 and 1. The extreme measurements are exactly the projections and gives the orthogonal resolutions of the identity. We have to prove that if U has at least 3 elements and the algebra is not commutative then there exists an extreme measurements which is not simple. If this is true when U contains 3 elements then it is true also when U has more than 3 elements. Therefore, assume that $\mathfrak{U}$ has 3 elements. The algebra is commutative and therefore there exists a sub-algebra isomorphic to $\mathbb{B}\left(\mathbb{C}^2\right)$. The matrices will be imbedded in the algebra via this sub-algebra. Define

$$M_u=\frac{1}{3}\left(\begin{array}{cc}1 & \cos\left(\theta_u\right)+i\sin\left(\theta_u\right)\\ \cos\left(\theta_u\right)-i\sin\left(\theta_u\right) & 1\end{array}\right)$$

where θ_u takes the values 0^0, 120^0 and 240^0. Then M_u is a resolution of the identity. The matrix M_u has eigenvalues 0 and $2/3$. Assume that M_u' and M_u'' are 2 resolutions of the identity such that $M_u=\frac{1}{2}\left(M_u'+M_u''\right)$. Then $M_u'=\alpha_u\cdot M_u$ because M_u is $2/3$ of a projection. Then

$$\begin{aligned}1&=\sum M_u'\\ &=\sum\alpha_u\cdot M_u\\ &=\sum\frac{\alpha_u}{3}\left(\begin{array}{cc}1 & \cos\left(\theta_u\right)+i\sin\left(\theta_u\right)\\ \cos\left(\theta_u\right)-i\sin\left(\theta_u\right) & 1\end{array}\right)\\ &=\left(\begin{array}{cc}\sum\frac{\alpha_u}{3} & \sum\frac{\alpha_u}{3}\left(\cos\left(\theta_u\right)+i\sin\left(\theta_u\right)\right)\\ \sum\frac{\alpha_u}{3}\left(\cos\left(\theta_u\right)-i\sin\left(\theta_u\right)\right) & \sum\frac{\alpha_u}{3}\end{array}\right).\end{aligned}$$

This shows that $\alpha_u=1$ and therefore $M_u'=M_u$. ■

The theorem shows that in general it is not sufficient to consider simple measurements in quantum theory. Many interesting measurements are described by non-simple measurements.

Let the Hilbert space $\mathcal{H}'$ be imbedded as a subspace of the Hilbert space $\mathcal{H}$, and let P be the projection on $\mathcal{H}'$. If M_B is a measurement in $\mathfrak{A}\subseteq\mathbb{B}\left(\mathcal{H}\right)$ then PM_BP is a measurement in $\mathcal{H}'$.

Theorem 34 *For any measurement M_B in $\mathcal{A} \subseteq \mathbb{B}(\mathcal{H})$ there exists a Hilbert space $\mathcal{L}$ containing $\mathcal{H}$ and a simple measurement E_B in $\mathbb{B}(\mathcal{L})$ such that $M_B = PE_BP$ where P is the projection of $\mathcal{L}$ on $\mathcal{H}$.*

Proof. Let $\mathcal{L}$ be the set of mappings $U \to \mathcal{H}$. Put

$$\langle f \mid g \rangle = \sum_u \left(f(u) \mid M_{\{u\}}(g(u)) \right) .$$

Then $\mathcal{L}$ is a pre-Hilbert space. Let $\mathcal{L}$ denote the completion of $\mathcal{L}$ with respect to $\langle \cdot \mid \cdot \rangle$. The map $l : v \to (u \to v)$ is an isometry of $\mathcal{H}$ into $\mathcal{L}$, and $\mathcal{H}$ can be identified with a subspace of $\mathcal{L}$. Let M_u be the operator $f \to f \cdot 1_u$. Then E obviously is an orthogonal resolution of the identity. For $v \in \mathcal{H}$ we have

$$\left(v \mid M_{\{u\}} v \right) = \left\langle l(v) \mid E_{\{u\}}(l(v)) \right\rangle$$

which proves that $M_B = PE_BP$. ■

5 Group representations on a Hilbert space

Let G be a connected group acting on the state space of a sum of matrix algebras. Let ϕ be a pure state. Then $a_g(\phi)$ is a pure state for each g in G. Therefore the set $\{a_g(\phi) \mid g \in G\}$ is a connected set of pure states, and therefore $\{a_g(\phi) \mid g \in G\}$ are states on one of the matrix algebras in the sum. If G is a connected group with a irreducible action on the state space of a $*$-algebra with a faithful state then the $*$-algebra is isomorphic to a matrix algebra. The most important groups in physics are connected and therefore we will restrict our attention to state spaces of the form $\mathbb{B}^1_+(\mathcal{H})$ where $\mathcal{H}$ is a Hilbert space.

An operator V from the Hilbert space $\mathcal{H}$ into itself is called *anti-unitary* if it is conjugated linear and satisfies $(V\vec{u} \mid V\vec{v}) = \overline{(V\vec{v} \mid V\vec{u})}$. The following theorem is due to Wigner.

Theorem 35 *Any automorphism of the state space $\mathbb{B}^1_+(\mathcal{H})$ is of the form*

$$S \to VSV^*,$$

where S denotes the density operator of a state, and where V is a unitary or anti-unitary operator in the Hilbert space $\mathcal{H}$.

It is straight forward to check that VSV^* is a density operator when S is a density operator. We will not prove the converse in the general case, but in case $\dim(\mathcal{H}) = 2$ it is easy and instructive. If $\dim(\mathcal{H}) = 2$ then the state space has the shape of a ball in 3 real dimensions. Therefore an automorphism of the state space maps the ball into it self. If the ball is centered then an automorphism is given by an orthogonal map in 3 real dimensions. The an orthogonal map is given by a rotation around an axis and, if the orthogonal map reverse orientation, a reflection in a plane orthogonal to the axis. By

a suitable choice of basis the density operators can be identified with 2×2 density matrices such that the axis of rotation is identified with the matrices

$$\begin{pmatrix} \frac{1}{2} & ic \\ -ic & \frac{1}{2} \end{pmatrix}.$$

An orientation preserving rotation is given by

$$V = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

The reflection

$$\begin{pmatrix} \frac{1}{2}+a & b+ic \\ b-ic & \frac{1}{2}-a \end{pmatrix} \curvearrowright \begin{pmatrix} \frac{1}{2}+a & b-ic \\ b+ic & \frac{1}{2}+a \end{pmatrix}$$

is given by the anti-unitary operator

$$\begin{pmatrix} x+iy \\ z+iw \end{pmatrix} \curvearrowright \begin{pmatrix} x-iy \\ z-iw \end{pmatrix}.$$

Remark 36 *An automorphism given by $S \to VSV^*$ maps the pure state corresponding to the vector $\vec{v}$ into the pure state corresponding to $V\vec{v}$.*

It is important to note that the unitary or anti unitary operator is only unique up to a scalar factor of unit modulus. That is V can be multiplied by $\omega \in \mathbb{C}, |\omega| = 1$ without changing the state $VS\acute{V}^*$. Now let G act on the state space. Then for each $g \in G$ there exists a unitary or anti unitary operator V_g in $\mathcal{H}$ such that $a_g(S) = V_g S V_g^*$. Further

$$V_g V_h S V_h^* V_g^* = V_{gh} S V_{gh}^*,\ g, h \in G,$$

for all density operators S. Therefore there exists a complex function $g, h \curvearrowright \omega(g,h)$ with $|\omega(g,h)| = 1$ such that

$$V_g V_h = \omega(g,h) V_{gh}\ ,\ g, h \in G. \tag{2}$$

Assume that G is a topological group which is connected, and assume that $g \curvearrowright V_g$ is continuous. Then V_g is unitary because it is not possible to pass continuously from a unitary to a anti-unitary operator.

Definition 37 *A continuous map $G \to U(\mathcal{H})$ such that (2) is satisfied is called a* projective unitary representation *of the group G on the Hilbert space $\mathcal{H}$. If $\omega(g,h) = 1$ for all $g, h \in G$ then the representation is said to be* unitary. *A projective unitary representation is said to be* irreducible *if $\{0\}$ and $\mathcal{H}$ are the only sub spaces of $\mathcal{H}$ invariant under V_g for all $g \in G$.*

Remark that if a projective unitary representation gives a irreducible action of the group on the state space, then the unitary representation is irreducible, but the opposite is not necessarily the case.

Theorem 38 *Let V_θ be a projective unitary representation of $\mathbb{R}$ on a complex Hilbert space $\mathcal{H}$ of dimension d. Then there exists a unitary representation of $\mathbb{R}$ given by a self adjoined operator A such that*

$$V_\theta = \alpha_\theta \exp(i\theta A) \ ,$$

where $|\alpha_\theta| = 1$.

Proof. First observe that $\det(V_\theta)$ is a continuous function $\mathbb{R} \to \{z \in \mathbb{C} \mid |z| = 1\}$. Then there exists a uniquely determined function α_θ such that $\alpha_\theta^d = \det(V_\theta)$. Then

$$\det\left(\frac{V_\theta}{\alpha_\theta}\right) = \frac{\det V_\theta}{\alpha_\theta^d} = 1 \ .$$

Thus $\tilde{V}_\theta = \frac{V_\theta}{\alpha_\theta}$ is a projective unitary representation $G \to SU(\mathcal{H})$ with

$$\tilde{V}_\theta \tilde{V}_\zeta = \tilde{\omega}(\theta, \zeta) \tilde{V}_{\theta+\zeta}. \tag{3}$$

Taking the determinant on each side of (3) leads to

$$1 = (\tilde{\omega}(\theta, \zeta))^d .$$

The equation $z^d = 1$ has the d unit roots as solutions which is a discrete set. The map $(\theta, \zeta) \curvearrowright \tilde{\omega}(\theta, \zeta)$ from $\mathbb{R}^2$ to C is continuous and $\tilde{\omega}(0, 0) = 1$. Therefore $\tilde{\omega}(\theta, \zeta) = 1$ for all $(\theta, \zeta) \in \mathbb{R}^2$ and the representation $\tilde{V}_\theta$ is unitary.

Let $\vec{v}_1, \vec{v}_2, ..., \vec{v}_d$ be eigenvectors of the unitary operator $\tilde{V}_\theta$. Then $\vec{v}_1, \vec{v}_2, ..., \vec{v}_d$ are also eigenvectors of $\tilde{V}_{n\theta}$. This is proved by induction in n. For $n = 1$ this is obvious. Assume that $\vec{v}_1, \vec{v}_2, ..., \vec{v}_d$ are eigenvectors for $V_{n\theta}$. Then $\tilde{V}_{(n+1)\theta} = V_{n\theta+\theta} = (\omega(g, h))^{-1} V_{n\theta} V_\theta$ and $\vec{v}_1, \vec{v}_2, ..., \vec{v}_d$ are also eigenvectors for $\tilde{V}_{(n+1)\theta}$. Therefore if $\mathcal{L}$ is the set of eigenvectors corresponding to an eigenvalue λ of V_θ then $\mathcal{L}$ is invariant under $V_{\frac{m\theta}{n}}$. By continuity $\mathcal{L}$ invariant under the V_θ for all θ. Therefore $\mathcal{H}$ can be written as a sum of Hilbert spaces $\mathcal{L}_i$ such that V_θ has only one eigenvalue λ_θ^i on the invariant $\mathcal{L}_i$. Then $\theta \to \lambda_\theta^i$ is a unitary representation on $\mathbb{C}$.Therefore there exists a_i such that $\lambda_\theta^i = \exp(i\theta\alpha_i)$. Let P_i denote the projection of $\mathcal{H}$ on $\mathcal{L}_i$. Put $A = \sum a_i P_i$. Then

$$\begin{aligned} V_\theta &= \alpha_\theta \tilde{V}_\theta \\ &= \alpha_\theta \sum \lambda_\theta^i P_i \\ &= \alpha_\theta \sum \exp(i\theta a_i) P_i \\ &= \alpha_\theta \exp(i\theta A) \ . \end{aligned}$$

■

An important example of the group $\mathbb{R}$ acting on a state space is time translations. Assume that a (closed) quantum system at time 0 is in a state described by the density operator S_0. Then the theorem states that there exists a self adjoined operator H such the state at time t is described by the density operator

$$S_t = \exp(-iHt/\hbar) S_0 \exp(iHt/\hbar)$$

where $\hbar = h/2\pi$ and h is *Planck's constant* and A in the theorem is given by $-H/\hbar$. In this case the self adjoined operator H is called the *Hamiltonian* and the corresponding observable is called the *energy observable*. Taking the derivative of this equation we get

$$i\hbar \frac{dS_t}{dt} = [H, S_t]$$

where $[A, B]$ is the commutator of the operators A and B given by $AB - BA$. If S_0 is a pure state given by the vector ψ then the time evolution is given by the Schrödinger equation

$$i\hbar \frac{d\psi_t}{dt} = H\psi_t \ .$$

If the group of time translations and time reversions are considered the group is no longer connected and the time reversions should be described by anti unitary operators. Remark also that

$$\begin{aligned} E_t(H) &= Tr\left(\exp(-iHt/\hbar) S_0 \exp(iHt/\hbar) H\right) \\ &= Tr\left(\exp(-iHt/\hbar) S_0 H \exp(iHt/\hbar)\right) \\ &= Tr(S_0 H) \\ &= E_t(H) \end{aligned}$$

so the mean energy is constant.

Now, put

$$S_\theta = \exp(i\theta A) S \exp(-i\theta A)$$

where A is self adjoined, and let X be any self-adjoined operator. Then we get the *Mandelstam-Tamm inequality*

$$\begin{aligned} Var_\theta(X) Var_\theta(A) &\geq \frac{1}{4} \left(Tr\left(S_\theta i[X, A]\right)\right)^2 \\ &= \frac{1}{4} \left(\frac{d}{d\theta} E_\theta(X)\right)^2 . \end{aligned}$$

Remark that $Tr(S_\theta A)$

$$\begin{aligned} Var_\theta(A) &= Tr\left(S_\theta (A - E(A))^2\right) \\ &= Tr\left(\exp(i\theta A) S \exp(-i\theta A) (A - E(A))^2\right) \\ &= Tr\left(\exp(i\theta A) S (A - E(A))^2 \exp(-i\theta A)\right) \\ &= Tr\left(S (A - E(A))^2\right) \\ &= Var(A) . \end{aligned}$$

Now, consider the measurement corresponding to X as an estimator of θ. Then the estimator is said to be *unbiased* if $E_\theta(X) = \theta$. Therefore for an unbiased

estimator X of θ we have

$$Var_{\theta}(X) \geq \frac{1}{4Var(A)} .$$

This is a kind of non-commutative *Rao-Cramér inequality.*

Let U_g be a unitary representation of a group G on a Hilbert space $\mathcal{H}$, and let $\mathcal{A}$ be the group algebra of F. Then a $*$-homomorphism of $\mathcal{A}$ into $\mathbb{B}(\mathcal{H})$ is given by

$$f \to \sum_{g \in G} f(g) U(g)$$

where f in the group algebra is a function $G \to \mathbb{C}$. Therefore finding all unitary representations is equivalent to finding all $*$-homomorphism of $\mathcal{A}$ into $\mathbb{B}(\mathcal{H})$. Now it is useful to know that $\mathcal{A}$ is a sum of matrix algebras. Using this technique we just have to find find $*$-homomorphism of matrix algebras or equivalently find unitary representations of $SU(n)$ for different values of n.

Example 39 *Let $G = \mathbb{R}^2$ be the set of 1 dimensional space translations and changes of the velocity. Let $W_{x,v}$ denote the unitary operator corresponding to a spatial shift of x and velocity translation of v. Put $V_x = W_{x,0}$ and $U_v = W_{0,v}$. Then $x \to V_x$ and $v \to U_v$ are projective representations, and we can assume that they are unitary representations. Now $(x,v) = (x,0) + (0,v) = (0,v) + (x,0)$, but the corresponding unitary operators should give the same change of state*

$$S \to U_v V_x S V_x^* U_v^* = V_x U_v S U_v^* V_x^*$$

for any S. It follows that

$$U_v V_x = \exp(i\eta(x,v)) V_x U_v$$

where $(x,v) \to \eta(x,v)$ is a real continuous function. For $x = 0$ or $v = 0$ this implies $1 = \exp(i\eta(0,v)) = \exp(i\eta(x,0))$ so we can put $\eta(x,0) = \eta(0,v) = 0$ since $V_0 = U_0 = 0$. Then

$$\eta(x, v+v') = \eta(x,v) + \eta(x,v') \mod 2\pi.$$

The only continuous solution to this equation satisfying $\eta(x,0) = 0$ is $\eta(x,v) = \eta(x) \cdot v$. In the same way

$$\eta(x+x', v) = \eta(x,v) + \eta(x',v) \mod 2\pi .$$

Therefore $\eta(x,v) = mxv$ for some real constant m which is called the mass. *We thus get*

$$U_v V_x = \exp(imxv) V_x U_v .$$

This is called Weyl's canonical commutator relation (CCR). Since Wx, v is $V_x U_v$ up to an arbitrary factor of unit modulus we can choose

$$W_{x,v} = \exp(imxv/2) V_x U_v .$$

Assume that the representation $(x,v) \to W_{x,v}$ is irreducible. Then $\mu = 0$ implies $[V_x, U_v] = 0$ and the only possibility is the one dimensional representation $(x,v) \to \exp(i(\alpha x + \beta v))$ with $\alpha, \beta \in \mathbb{R}$. The case with $\mu > 0$ is more important. The constant μ is (proportional to) the mass of the quantum system (particle) associated with the representation. The group R^2 is simply connected. Therefore no finite dimensional projective representation can be associated with a quantum system with $\mu \neq 0$.

Assume that $U_v = \exp ivA$ and $V_x = \exp ixB$. The the derivative of Weyl's CCR with respect to x and v in $(x,v) = (0,0)$ is

$$iAiB = \frac{i\mu}{2} + iBiA$$

which is equivalent to

$$[A,B] = \frac{\mu}{2i}.$$

This is called Heisenberg's canonical commutator relation. Then

$$\begin{aligned} Var(A)\, Var(B) &\geq \frac{1}{4}\phi\left(i[A,B]\right)^2 \\ &= \frac{1}{4}\phi\left(\frac{\mu}{2}\right)^2 \\ &= \frac{\mu^2}{16} \end{aligned}$$

Let $\mathcal{H}$ be the infinite dimensional Hilbert space $L^2(\mathbb{R})$ equipped with the usual scalar product. Then unitary operators are defined by

$$\begin{aligned} V_x\psi(\xi) &= \psi(\xi - x) \\ U_v\psi(\xi) &= \exp(imv\xi)\,\psi(\xi) \end{aligned}$$

Then a projective representation is given by

$$\begin{aligned} W_{x,v}\psi(\xi) &= \exp(imxv/2)\, V_x U_v \psi(\xi) \\ &= \exp\left(imv\left(\xi - \frac{x}{2}\right)\right)\psi(\xi - x) \ . \end{aligned}$$

6 Crossed products

We have seen that groups and $*$-algebras are closely related. With the tools that we have available now we are able to make a new construction that emphasize this relation. Let $\mathcal{A}$ be a $*$-algebra represented on a Hilbert space $\mathbb{H}$. Let α be an action of a group G on $\mathcal{A}$. If $g \to U_g$ is a unitary representation of the group then an action α is given by

$$\alpha_g(X) = U_g X U_g^* \ .$$

If $U_g \in A$ for all g then the action α is said to be an inner action.

Proposition 40 *Let G be a group acting on A. Then the set of inner automorphisms a normal subgroup of G.*

Proof. Assume that $\alpha_{g_1}(X) = U_{g_1} X U_{g_1}^*$ and let α_{g_2} be another automorphism. Then

$$\begin{aligned}\left(\alpha_{g_2}\alpha_{g_1}\alpha_{g_2}^{-1}\right)(X) &= \alpha_{g_2}\left(U_{g_1}\alpha_{g_2}^{-1}(X)\,U_{g_1}^*\right)\\ &= \alpha_{g_2}(U_{g_1})\,\alpha_{g_2}\left(\alpha_{g_2}^{-1}(X)\right)\alpha_{g_2}\left(U_{g_1}^*\right)\\ &= \alpha_{g_2}(U_{g_1})\,X\,(\alpha_{g_2}(U_{g_1}))^*\end{aligned}$$

so $\alpha_{g_2}\alpha_{g_1}\alpha_{g_2}^{-1}$ is given by the unitary operator $\alpha_{g_2}(U_{g_1})$. ■

Theorem 41 *If $\mathcal{A}$ is a simple $*$-algebra with a faithful state then $\mathcal{A}$ is perfect, i.e. any automorphism is inner.*

Proof. A simple $*$-algebra has the form $\mathbb{B}(\mathbb{H})$ for some Hilbert space $\mathbb{H}$. Let $P_1, P_2, ..., P_n$ be orthogonal one dimensional projections in $\mathcal{A}$ such that $\sum P_j = 1$. Then

$$\sum \alpha(P_j) = \alpha\left(\sum P_j\right) = \alpha(1) = 1.$$

We know that $\alpha(P_j)$ is a projection different from 0 for all j and therefore $\alpha(P_j)$ must be an orthogonal one dimensional projection. Let v_j be the unit vector corresponding to P_j and w_j be the unit vector corresponding to $\alpha(P_j)$. Then there exists a unique unitary operator U such that $U(v_j) = w_j$. With these definitions

$$\begin{aligned}\alpha(X) &= \alpha\left(\sum_{j,k} P_j X P_k\right)\\ &= \alpha\left(\sum_{j,k}\left(v_j \otimes v_j^*\right) X\left(v_k \otimes v_k^*\right)\right)\\ &= \alpha\left(\sum_{j,k}\left(Xv_k \mid v_j\right)\left(v_j \otimes v_k^*\right)\right)\\ &= \sum_{j,k}\left(Xv_k \mid v_j\right)\left(w_j \otimes w_k^*\right)\\ &= \sum_{j,k}\left(XU^*w_k \mid U^*w_j\right)\left(w_j \otimes w_k^*\right)\\ &= \sum_{j,k}\left(UXU^*w_k \mid w_j\right)\left(w_j \otimes w_k^*\right)\\ &= UXU^*.\end{aligned}$$

■

We see that α is an inner automorphism of a simple $*$-algebra then the associated unitary operator U is uniquely determined modulo a factor $\lambda \in \mathbb{C}$ with $|\lambda| = 1$. Thus any action of a group on a simple $*$-algebra is given by a projective unitary repretentation of G.

An automorphism that is not inner is said to be outer. We shall now embed the algebra in a larger algebra such that the group action is given by a unitary representation that is inner. we shall assume that the algebra is represented on a Hilbert space $\mathbb{H}$. A new Hilbert space $L^2(G, \mathbb{H})$ is defined as the set of functions from G to $\mathbb{H}$. The group G is assumed to be compact so that it has a Haar measure μ and the inner product on $L^2(G, \mathbb{H})$ is defined by

$$(f \mid h) = \int_G (f(g) \mid h(g)) \ d\mu(g).$$

The algebra is embedded in the operators on $L^2(G, \mathbb{H})$ as follows. An element $g \in G$ is embedded in $\mathbb{B}\left(L^2(G, \mathbb{H})\right)$ as the unitary operator U_g

$$U_g(f)(g_0) = f\left(g^{-1} * g_0\right)$$

Let X be an operator. Then $\pi(X) \in \mathbb{B}\left(L^2(G, \mathbb{H})\right)$ is given by

$$\pi(X)(f)(g) = \alpha_{g^{-1}}(X)(f(g)).$$

One easily checks that $g \to U_g$ and π are embeddings. Next we see that

$$\begin{aligned} U_g \pi(X) U_g^*(f)(g_0) &= \pi(X) U_g^*(f)\left(g^{-1} * g_0\right) \\ &= \alpha_{g_0^{-1} * g}(X)\left(U_g^*(f)\left(g^{-1} * g_0\right)\right) \\ &= \alpha_{g_0^{-1} * g}(X)(f(g_0)) \\ &= \alpha_{g_0^{-1}}(\alpha_g(X))(f(g_0)) \\ &= \pi(\alpha_g(X))(f)(g_0) \ . \end{aligned}$$

Thus

$$U_g \pi(X) U_g^* = \pi(\alpha_g(X)).$$

Definition 42 *The algebra generated by the subalgebra $\pi(A)$ and the operators U_g is called a crossed product and is denoted $\mathcal{A} \ltimes_\alpha G$.*

Any element in $\mathcal{A} \ltimes_\alpha G$ can be written as

$$\sum_{g \in G} \pi(X_g) U_g$$

where $g \to X_g$ is a arbitrary function from G to $\mathcal{A}$. We see that

$$\pi(X) = \sum_{g \in G} \pi\left(\alpha_{g^{-1}}(X)\right) U_g$$

and

$$U_{g_0} = \sum_{g \in G} \pi\left(\delta_{g,g_0}\right) U_g \ .$$

We have

$$\begin{aligned}
\sum_{g_1 \in G} \pi\left(X_{g_1}\right) U_{g_1} \sum_{g_2 \in G} \pi\left(Y_{g_2}\right) U_{g_2} &= \sum_{g_1 \in G} \sum_{g_2 \in G} \pi\left(X_{g_1}\right) U_{g_1} \pi\left(Y_{g_2}\right) U_{g_1}^* U_{g_1} U_{g_2} \\
&= \sum_{g_1 \in G} \sum_{g_2 \in G} \pi\left(X_{g_1} \pi \alpha_{g_1}\left(Y_{g_2}\right)\right) U_{g_1 g_2} \\
&= \sum_{g \in G} \left(\sum_{g_1 g_2 = g} \pi\left(X_{g_1} \alpha_{g_1}\left(Y_{g_2}\right)\right) \right) U_g \\
&= \sum_{g \in G} \left(\sum_{g_1 \in G} \pi\left(X_{g_1} \alpha_{g_1}\left(Y_{g_1^{-1} g}\right)\right) \right) U_g .
\end{aligned}$$

Example 43 *If $A = \mathbb{C}$ then the Hilbert space can be chosen to be one dimensional and can be identified with $\mathbb{C}$. In this case $L^2(G, \mathbb{H})$ is simply functions on G. The elements of $\mathcal{A} \ltimes_\alpha G$ have the form*

$$\sum_{g \in G} f\left(g\right) U_g$$

where f is a complex function G. Then

$$\begin{aligned}
\sum_{g \in G} f\left(g\right) U_g \sum_{g \in G} h\left(g\right) U_g &= \sum_{g \in G} \left(\sum_{g_1 \in G} f\left(g_1\right) h\left(g_1^{-1} * g\right) \right) U_g \\
&= \sum_{g \in G} \left(f * h\right)\left(g\right) U_g \ .
\end{aligned}$$

Therefore $\mathbb{C} \ltimes_\alpha G$ is isomorphic to the group algebra of G.

If ϕ is a state on A then a state on $\mathcal{A} \ltimes_\alpha G$ is defined by

$$\phi\left(\sum_{g \in G} \pi\left(X_g\right) U_g \right) = \int \phi\left(\alpha_g\left(X_g\right)\right) \ d\mu\left(g\right).$$

With this definition

$$\begin{aligned}
\phi\left(\pi\left(X\right)\right) &= \phi\left(\sum_{g \in G} \pi\left(\alpha_{g^{-1}}\left(X\right)\right) U_g \right) \\
&= \int \phi\left(\alpha_g\left(\alpha_{g^{-1}}\left(X\right)\right)\right) \ d\mu\left(g\right) \\
&= \int \phi\left(X\right) \ d\mu\left(g\right) \\
&= \phi\left(X\right) \ .
\end{aligned}$$

Assume that $\rho : G \to U(\mathbb{H})$ is a unitary representation such that $\rho(g) \in \mathcal{A}$ for all $g \in G$ and such that $\alpha(X) = \rho(g) X \rho(g)^*$. Then $\mathcal{A} \ltimes G$ can be projected into $\mathcal{A}$ via

$$l : \sum_{g \in G} \pi(X_g) U_g \to \sum_{g \in G} X_g \rho(g)$$

We shall show that it is a homomorphism. Obviously it is linear.

$$\begin{aligned}
l\left(\sum_{g_1 \in G} \pi(X_{g_1}) U_{g_1}\right) l\left(\sum_{g_2 \in G} \pi(X_{g_2}) U_{g_2}\right) &= \sum_{g_1 \in G} X_{g_1} \rho(g_1) \sum_{g_2 \in G} X_{g_2} \rho(g_2) \\
&= \sum_{g_1 \in G} \sum_{g_2 \in G} X_{g_1} \rho(g_1) X_{g_2} \rho(g_1)^* \rho(g_1) \rho(g_2) \\
&= \sum_{g_1 \in G} \sum_{g_2 \in G} X_{g_1} \alpha_{g_1}(X_{g_2}) \rho(g_1 g_2) \\
&= \sum_{g \in G} \sum_{g_1 \in G} X_{g_1} \alpha_{g_1}\left(X_{g_1^{-1} g}\right) \rho(g) \\
&= l\left(\sum_{g \in G} \sum_{g_1 \in G} \pi\left(X_{g_1} \alpha_{g_1}\left(X_{g_1^{-1} g}\right)\right) U_g\right) \\
&= l\left(\sum_{g_1 \in G} \pi(X_{g_1}) U_{g_1} \sum_{g_2 \in G} \pi(X_{g_2}) U_{g_2}\right).
\end{aligned}$$

Note that we actively use that ρ is a unitary representation and not only a projective unitary representation. Next we shall show that $l \circ \pi$ is the identity. For this

$$\begin{aligned}
(l \circ \pi)(X) &= l(\pi(X) U_e) \\
&= X \rho(e) \\
&= X.
\end{aligned}$$

For $\mathcal{A} = \mathbb{C}$ we know that $\mathcal{A} \ltimes G$ is the group algebra which is $|G|$-dimensional so if G is not trivial then $\mathcal{A}$ will be a proper subalgebra of $\mathcal{A} \ltimes G$. This is a major disadvantage about crossed products: In general the crossed product is an algebra that is much bigger than it ought to be for our applications.

Let G be a group acting on $\mathcal{A}$ and let N be a normal subgroup of G. Then there also an action of N on G. It is interesting to compare $\mathcal{A} \ltimes N$ with $\mathcal{A} \ltimes G$. The quotient G/N acts on $\mathcal{A} \ltimes N$ via

$$\tilde{\alpha}_{g_0}\left(\sum_{g \in N} \pi(X_g) U_g\right) = \sum_{g \in N} \pi(\alpha_{g_0}(X_g)) U_{g_0 g g_0^{-1}}.$$

We have to check that this is a homomorphism. We have

$$\begin{aligned}
\tilde{\alpha}_{g_0}\left(\sum_{g_1\in N}\pi\left(X_{g_1}\right)U_{g_1}\right)\tilde{\alpha}_{g_0}\left(\sum_{g_2\in N}\pi\left(Y_{g_2}\right)U_{g_2}\right) &= \sum_{g_1\in N}\pi\left(\alpha_{g_0}\left(X_{g_1}\right)\right)U_{g_0g_1g_0^{-1}}\sum_{g_2\in N}\pi\left(\alpha_{g_0}\left(X_{g_2}\right)\right)U_{g_0g_2g_0^{-1}} \\
&= \sum_{g_1\in N}\sum_{g_2\in N}\pi\left(\alpha_{g_0}\left(X_{g_1}\right)\right)U_{g_0g_1g_0^{-1}}\pi\left(\alpha_{g_0}\left(X_{g_2}\right)\right)U^*_{g_0g_1g_0^{-1}}U_{g_0g_1g} \\
&= \sum_{g_1\in N}\sum_{g_2\in N}\pi\left(\alpha_{g_0}\left(X_{g_1}\right)\right)\pi\left(\alpha_{g_0g_1g_0^{-1}}\alpha_{g_0}\left(X_{g_2}\right)\right)U_{g_0g_1g_2g_0^{-1}} \\
&= \sum_{g_1\in N}\sum_{g_2\in N}\pi\left(\alpha_{g_0}\left(X_{g_1}\right)\alpha_{g_0g_1}\left(X_{g_2}\right)\right)U_{g_0g_1g_2g_0^{-1}} \\
&= \sum_{g\in N}\sum_{g_1\in N}\pi\left(\alpha_{g_0}\left(X_{g_1}\alpha_{g_1}\left(X_{g_1^{-1}g}\right)\right)\right)U_{g_0gg_0^{-1}} \\
&= \tilde{\alpha}_{g_0}\left(\sum_{g\in N}\sum_{g_1\in N}\pi\left(X_{g_1}\alpha_{g_1}\left(X_{g_1^{-1}g}\right)\right)U_{g_0gg_0^{-1}}\right).
\end{aligned}$$

We also have

$$\begin{aligned}
\left(\tilde{\alpha}_{g_1}\circ\tilde{\alpha}_{g_2}\right)\left(\sum_{g\in N}\pi\left(X_g\right)U_g\right) &= \tilde{\alpha}_{g_1}\left(\sum_{g\in N}\pi\left(\alpha_{g_2}\left(X_g\right)\right)U_{g_2gg_2^{-1}}\right) \\
&= \sum_{g\in N}\pi\left(\alpha_{g_1}\left(\alpha_{g_2}\left(X_g\right)\right)\right)U_{g_1g_2gg_2^{-1}g_1^{-1}} \\
&= \sum_{g\in N}\pi\left(\alpha_{g_1g_2}\left(X_g\right)\right)U_{g_1g_2g(g_1g_2)^{-1}} \\
&= \tilde{\alpha}_{g_1g_2}\left(\sum_{g\in N}\pi\left(X_g\right)U_g\right).
\end{aligned}$$

Next we shall see that $(\mathcal{A}\ltimes N)\ltimes G/N$ is isomorphic to $\mathcal{A}\ltimes G$.

7 Tensor products

7.1 Tensor products of Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{K}$ be finite dimensional vector spaces. Then a bilinear map f from $\mathcal{H}\times\mathcal{K}$ into the vector space $\mathcal{L}$ is a map which satisfies:

- For $\vec{v}\in\mathcal{K}$ the map $\vec{u}\to f\left(\vec{u},\vec{v}\right)$ is linear.
- For $u\in\mathcal{H}$ the map $\vec{v}\to f\left(\vec{u},\vec{v}\right)$ is linear.

Now we will define a new vector space called the *tensor product* of $\mathcal{H}$ and $\mathcal{K}$. Let $\left(\vec{u}_1,\vec{u}_2,...,\vec{u}_m\right)$ be an orthonormal basis of $\mathcal{H}$ and let $\left(\vec{v}_1,\vec{v}_2,...,\vec{v}_n\right)$ be

an orthonormal basis of $\mathcal{K}$. Then $\mathbb{C}^{nm}$ has a structure as a vector space of dimension nm. Choose an orthonormal basis of $\mathbb{C}^{nm}$ and label the nm basis vectors $\vec{w}_{i,j}, i = 1, 2, ..., m, j = 1, 2, ..., n$. Then the map $l : \mathcal{H} \times \mathcal{K} \to \mathbb{C}^{nm}$ defined by

$$l\left(\sum x_i \vec{u}_i, \sum y_j \vec{v}_j\right) = \sum_{i,j} x_i y_j \cdot \vec{w}_{i,j}$$

is bilinear. Let $\vec{u} \in H$ and $\vec{v} \in K$ then $l(\vec{u}, \vec{v})$ is called the *tensor product* of $\vec{u}$ and $\vec{v}$ and is denoted $\vec{u} \otimes \vec{v}$. We see that $\otimes$ is distributive in the sense that

$$(\vec{u} + \vec{w}) \otimes \vec{v} = \vec{u} \otimes \vec{v} + \vec{w} \otimes \vec{v}$$

and

$$\vec{v} \otimes (\vec{u} + \vec{w}) = \vec{v} \otimes \vec{u} + \vec{v} \otimes \vec{w}.$$

The tensor product of $\mathcal{H}$ and $\mathcal{K}$ and is denoted $\mathcal{H} \otimes \mathcal{K}$.

An equivalent way to define the tensor products is the following. Identify the vector space $\mathcal{H}$ with the set $C(U)$ for some finite set U and identify the vector space $\mathcal{K}$ with $C(V)$ for some finite space V. Then the tensor product of $\mathcal{H}$ and $\mathcal{K}$ is identified with $C(U \times V)$. The map l is then identified with the map $l(f, g) : (u, v) \to f(u) g(v)$.

Theorem 44 *Let $\mathcal{H}, \mathcal{K}$ and $\mathcal{L}$ be vector spaces, and let $f : \mathcal{H} \times \mathcal{K} \to \mathcal{L}$ be a bilinear map. Then there exists a unique linear map $g : \mathcal{H} \otimes \mathcal{K} \to \mathcal{L}$ such that*

$$f(\vec{u}, \vec{v}) = g(\vec{u} \otimes \vec{v}).$$

Assume that $\mathbb{M}$ is a vector space and $l : \mathcal{H} \times \mathcal{K} \to \mathbb{M}$ is a bilinear map such that for any bilinear map $f : \mathcal{H} \times \mathcal{K} \to \mathcal{L}$ there exists a linear map $g : \mathbb{M} \to \mathcal{L}$ such that

$$f(u, v) = g(l(u, v)).$$

Then there exists an injection map $g : \mathcal{H} \otimes \mathcal{K}. \to \mathbb{M}$ such that $l(\vec{u}, \vec{v}) = g(\vec{u} \otimes \vec{v})$.

Proof. Let $f : \mathcal{H} \times \mathcal{K} \to \mathcal{L}$ be a bilinear map. Let $(\vec{u}_1, \vec{u}_2, ..., \vec{u}_m)$ be an orthonormal basis of $\mathcal{H}$ and let $(\vec{v}_1, \vec{v}_2, ..., \vec{v}_n)$ be an orthonormal basis of $\mathcal{K}$. If $\vec{u} = \sum x_i \vec{u}_i$ and $\vec{v} = \sum y_j \vec{v}_j$ then

$$f(\vec{u}, \vec{v}) = \sum_{i,j} x_i y_j \cdot f(\vec{u}_i, \vec{v}_j).$$

Therefore the linear map g is uniquely determined by

$$g(\vec{u}_i \otimes \vec{v}_j) = f(\vec{u}_i, \vec{v}_j).$$

Let M be a vector space with the properties stated. Then there exists a $g : \mathcal{H} \otimes \mathcal{K} \to \mathbb{M}$ such that

$$l(\vec{u}, \vec{v}) = g(\vec{u} \otimes \vec{v}),$$

so we just have to prove that g is injective. Let $\vec{w}$ be a vector in $\mathcal{H}\otimes\mathcal{K}$. Then the map $f:\mathcal{H}\otimes\mathcal{K}\rightarrow\mathbb{C}$ given by $f\left(\vec{u},\vec{v}\right)=\left(\vec{u}\otimes\vec{v}\mid\vec{w}\right)$ is bilinear and there exists a linear map $h:\mathbb{M}\rightarrow\mathbb{C}$ such that

$$\begin{aligned}\left(\vec{u}\otimes\vec{v}\mid\vec{w}\right) &= f\left(\vec{u},\vec{v}\right)\\ &= h\left(l\left(\vec{u},\vec{v}\right)\right)\\ &= h\left(g\left(\vec{u}\otimes\vec{v}\right)\right).\end{aligned}$$

If $g\left(\vec{w}\right)=0$ and $\vec{w}=\sum_i\vec{u}^i\otimes\vec{v}^i$ then

$$\begin{aligned}\left(\vec{w}\mid\vec{w}\right) &= \sum\left(\vec{u}^i\otimes\vec{v}^i\mid\vec{w}\right)\\ &= \sum h\left(g\left(\vec{u}^i\otimes\vec{v}^i\right)\right)\\ &= \sum h\left(g\left(\vec{w}\right)\right)\\ &= 0\end{aligned}$$

and $\vec{w}=0$. Therefore the map $g:\mathcal{H}\otimes\mathcal{K}\rightarrow\mathbb{M}$ is injective. ■

Let $\oplus$ denote addition of vector spaces. Then the tensor product is distributive in the sense that

$$\mathcal{H}\otimes\left(\mathcal{K}\oplus\mathcal{L}\right)=\left(\mathcal{H}\otimes\mathcal{K}\right)\oplus\left(\mathcal{H}\otimes\mathcal{L}\right)$$

and

$$\left(\mathcal{K}\oplus\mathcal{L}\right)\otimes\mathcal{H}=\left(\mathcal{K}\otimes\mathcal{H}\right)\oplus\left(\mathcal{L}\otimes\mathcal{H}\right).$$

7.2 Tensor products of $*$-algebras

The construction of tensor products can be extended to $*$-algebras. Let $\mathcal{A}$ and $\mathcal{B}$ be finite dimensional $*$-algebras each of them being unital and with a faithful state. Then $\mathcal{A}$ and $\mathcal{B}$ both have a structure as a vector space. Therefore it is possible to define $\mathcal{A}\otimes\mathcal{B}$ as the vector space tensor product of $\mathcal{A}$ and $\mathcal{B}$. The tensor product is organized as an algebra via the product given by

$$\left(X\otimes Y\right)\left(V\otimes W\right)=XV\otimes YW.$$

If $\mathcal{A}$ is represented on $\mathcal{H}$ via $\pi:\mathcal{A}\rightarrow\mathbb{B}\left(\mathcal{H}\right)$ and $\mathcal{B}$ is represented on $\mathcal{K}$ via $\rho:\mathcal{B}\rightarrow\mathbb{B}\left(\mathcal{K}\right)$ then $\mathcal{A}\otimes\mathcal{B}$ is represented on $\mathcal{H}\otimes\mathcal{K}$ via

$$\left(\pi\otimes\rho\right)\left(X\otimes Y\right)\left(\vec{u}\otimes\vec{v}\right)=\pi\left(X\right)\left(\vec{u}\right)\otimes\rho\left(Y\right)\left(\vec{v}\right).$$

The special case where $\mathcal{A}=\mathbb{B}\left(\mathcal{H}\right)$ and $\mathcal{B}=\mathbb{B}\left(\mathcal{K}\right)$ gives a "natural" isomorphism

$$\mathbb{B}\left(\mathcal{H}\right)\otimes\mathbb{B}\left(\mathcal{K}\right)\rightarrow\mathbb{B}\left(\mathcal{H}\otimes\mathcal{K}\right).$$

It works as follows. For a pair of operators $\left(X,Y\right)\in\mathbb{B}\left(\mathcal{H}\right)\times\mathbb{B}\left(\mathcal{K}\right)$ a bilinear map from $\mathcal{H}\times\mathcal{K}$ into $\mathcal{H}\otimes\mathcal{K}$ is given by

$$\left(\vec{u},\vec{v}\right)\rightarrow X\left(\vec{u}\right)\otimes Y\left(\vec{v}\right).$$

This bilinear map is given by an operator $Z : \mathcal{H} \otimes \mathcal{K} \to \mathcal{H} \otimes \mathcal{K}$. Thus the map which maps $(X, Y) \in \mathbb{B}(\mathcal{H}) \otimes \mathbb{B}(\mathcal{K})$ into $Z \in \mathbb{B}(\mathcal{H} \otimes \mathcal{K})$ is bilinear and is given by a linear map $\mathbb{B}(\mathcal{H}) \otimes \mathbb{B}(\mathcal{K}) \to \mathbb{B}(\mathcal{H} \otimes \mathcal{K})$. It is obviously injective and therefore it must be an isomorphism because $\mathbb{B}(\mathcal{H}) \otimes \mathbb{B}(\mathcal{K})$ and $\mathbb{B}(\mathcal{H} \otimes \mathcal{K})$ have the same dimension.

8 Partial measurements

A measurement has a state as input and a probability distribution as output. A partial measurement has a state as input but the output consist of two parts: a probability distribution and a state. The output state can then serve as input for a new measurement or a new partial measurement. The formal definition of a partial measurement is quite abstract, and is given as follows. As usual subsets of a set are identified with their indicator functions.

Definition 45 *Let $\mathcal{A}$ and $\mathcal{B}$ be finite dimensional $*$-algebras embedded in the set of operators on Hilbert spaces. Let U be a finite set. Then a partial measurement on $\mathcal{A}$ with values in U and output states in $\mathcal{B}$ is given by a* positive map valued measure *(PMVM) $\mathcal{E}$ mapping subsets of U into positive linear maps from $\mathcal{A}$ into $\mathcal{B}$ such that*

$$\mathcal{E}(1_{A \cup B}) = \mathcal{E}(1_A) + \mathcal{E}(1_B)$$

for A and B disjoined subsets of U and $\mathcal{E}(1)$ maps density operators into density operators.

The interpretation is as follows. If the input state is given by the density operator S then the probability of measuring a result in $A \subseteq U$ is

$$Tr(\mathcal{E}(1_A)(S)),$$

and if a measurement in A is observed then our knowledge of the output state is described by the density operator

$$\frac{\mathcal{E}(1_A)(S)}{Tr(\mathcal{E}(1_A)(S))}.$$

If $A = \{u_1, u_2, ..., u_n\}$ then by linearity

$$\mathcal{E}(1_A) = \sum_{i=1}^{n} \mathcal{E}(1_{u_i}).$$

Therefore $\mathcal{E}$ is determined by its values on the elements in U. The partial measurement can formally be extended to admit functions $U \to C$ as input.

$$\mathcal{E}(f) = \sum_{u \in U} f(u)\, \mathcal{E}(1_u)$$

Sometimes partial measurements are called *instruments*. It is possible to compose two partial measurements by using the output state of the first partial measurement as input of the second measurement. The formal definition of the composed measurement is as follows.

Definition 46 *Let $\mathcal{E}_1$ be a partial measurement with input states in $\mathcal{A}$, output states in $\mathcal{B}$ and results in the finite set U, and let $\mathcal{E}_2$ be a measurement with input states in $\mathcal{B}$ and output states in $\mathcal{C}$ and results in the finite set V. The a partial measurement with input states in $\mathcal{A}$, output states in $\mathcal{C}$ and results in $V \times U$ is defined by the formula*

$$\left(\mathcal{E}_2 \circ \mathcal{E}_1\right)(f) = \sum_{(v,u)\in V\times U} f(v,u)\,\mathcal{E}_2(1_v)\,\mathcal{E}_1(1_u) \ .$$

Remark 47 *This definition relies heavily on the condition that U and V are finite sets. The definition of par*

Theorem 48 *If $\mathcal{E}$ is an instrument then there exists a uniquely defined measurement $\mathcal{M}_{\mathcal{E}}$ such that*

$$\frac{\mathcal{E}(1_B)(\phi)}{\mathcal{M}_{\mathcal{E}1_B}(\phi)} \in \mathfrak{S}$$

for $1_B \in \mathcal{B}(\mathfrak{U})$.

Proof. For any measurable function f there exists a reel number $c_f(\phi)$ such that

$$\frac{\mathcal{E}(f)(\phi)}{c_{\mathcal{E}f}(\phi)} \in \mathfrak{S} \ .$$

Using the properties of $\mathcal{E}$ we see that $\mathcal{M}_{\mathcal{E}1_B}(\phi) = c_{\mathcal{E}1_B}(\phi)$ is a measurement. ■

It is possible to compose instruments:

Theorem 49 *Let $\mathcal{E}^i$ be an instrument with values in $\mathfrak{U}^i$ for $i = 1, 2$. If $\mathfrak{U}^i$ are finite then there exists an instrument $\mathcal{E}^2 \times \mathcal{E}^1$ with values in $\mathfrak{U}^2 \times \mathfrak{U}^1$ such that*

$$\mathcal{E}^2 \times \mathcal{E}^1(f_2 \times f_1)(S) = \mathcal{E}^2(f_2)\,\mathcal{E}^1(f_1)(S)$$

for all states S and all $f_i \in \mathcal{B}\left(\mathfrak{U}^i\right)$.

Proof. Using that $\mathfrak{U}^i$ is finite we can define

$$\mathcal{E}^2 \times \mathcal{E}^1(f)(S) = \sum_{(e_2,e_1)\in\mathfrak{U}^2\times\mathfrak{U}^1} f(e_2,e_1)\cdot\mathcal{E}^2(e_2)\,\mathcal{E}^1(e_1)(S) \ .$$

■

Similarly it is possible to compose an instrument with a measurement.

Definition 50 *An instrument is said to be* repetitive *if $\mathcal{E}(B)\,\mathcal{E}(B) = \mathcal{E}(B)$ for all measurable sets B.*

Definition 51 *The mapping* $\mathcal{E}(1):\mathfrak{S}\to\mathfrak{S}$ *will be called* the reduction of the state space.

Theorem 52 *If an instrument* $\mathcal{E}$ *with values in* $\mathfrak{U}$ *is repetitive the corresponding measurement* $\mathcal{M}_{\mathcal{E}}$ *is simple on the reduced state space* $\mathcal{E}(1)(\mathfrak{S})$.

Proof. Without loss of generality we may assume that $\mathfrak{U}$ contains 2 elements and that $\mathcal{E}(1)(\mathfrak{S})=\mathfrak{S}$. We have to prove that $\mathcal{M}(1),\mathcal{M}(1_{u_1}),\mathcal{M}(1_{u_2})$ and $\mathcal{M}(0)$ are extreme functionals: $\mathfrak{S}\to[0;1]$. We have $\mathcal{M}(1)=id$, $\mathcal{M}(0)=0$ and $\mathcal{M}(1_{u_2})=1-\mathcal{M}(1_{u_1})$ and therefore it is sufficient to show that $\mathcal{M}(1_{u_1})$ is extreme. Assume $\mathcal{M}(1_{u_1})=\frac{1}{2}\Phi_1+\frac{1}{2}\Phi_2$. If $\mathcal{E}(1_{u_1})(\phi)=0$ then $\mathcal{M}_\phi(1_{u_1})=0$ and $\Phi_i(\phi)=0$. If $\mathcal{E}(1_{u_1})(\phi)=\phi$ then $\mathcal{M}_\phi(1_{u_1})=1$ and $\Phi_i(\phi)=1$. In general we have

$$\phi=\mathcal{M}_{\phi}(1_{u_1})\frac{\mathcal{E}(1_{u_1})(\phi)}{\mathcal{M}_{\phi}(1_{u_1})}+\mathcal{M}_{\phi}(1_{u_2})\frac{\mathcal{E}(1_{u_2})(\phi)}{\mathcal{M}_{\phi}(1_{u_2})}$$

which proves that

$$\Phi_i(\phi)=\mathcal{M}_{\phi}(1_{u_1})\cdot 1+\mathcal{M}_{\phi}(1_{u_2})\cdot 0=\mathcal{M}_{\phi}(1_{u_1})\ .$$

■

Theorem 53 *Let* $\mathcal{M}$ *be a simple measurement on* $\mathfrak{S}$. *Then there exists at most 1 repetitive instrument*$\mathcal{E}_{\mathcal{M}}$ *such that* $\mathcal{M}$ *is the measurement corresponding to* $\mathcal{E}_{\mathcal{M}}$ *and such that* $\mathcal{E}_{\mathcal{M}}(1)$ *is the identity on* $\mathfrak{S}$.

Proof. Let $\mathcal{E}$ be a repetitive instrument such that $\mathcal{M}$ is the corresponding measurement and such that $\mathcal{E}(1)$ is the identity on $\mathfrak{S}$. Let B be a measurable subset of $\mathfrak{U}$ and let $\phi\in\mathfrak{S}$ be an extreme state. Then

$$\phi=\mathcal{M}_{\phi}(1_B)\frac{\mathcal{E}(1_B)(\phi)}{\mathcal{M}_{\phi}(1_B)}+\mathcal{M}_{\phi}(1_B)\frac{\mathcal{E}(1_B)(\phi)}{\mathcal{M}_{\phi}(1_B)}$$

which proves that

$$\phi=\frac{\mathcal{E}(1_B)(\phi)}{\mathcal{M}_{\phi}(1_B)}$$

and therefore

$$\mathcal{E}(1_B)(\phi)=\mathcal{M}_{\phi}(1_B)\cdot\phi$$

so that $\mathcal{E}(1_B)$ is uniquely determined on the extreme points in $\mathfrak{S}$ and therefore on the whole of $\mathfrak{S}$. ■

Theorem 54 *Let* $\mathcal{E}$ *be a repetitive instrument. Then* $\mathcal{E}$ *is uniquely determined by its reduction* $\mathcal{E}(1)$ *and its measurement* $\mathcal{M}_{\mathcal{E}}$ *via the formula*

$$\mathcal{E}=\mathcal{E}_{\mathcal{M}_{\mathcal{E}}}\circ\mathcal{E}(1)\ .$$

Proof. This is an immediate consequence of the 2 previous theorems. ■

Although $\mathcal{E}$ is repetitive the corresponding measurement need not be simple on all of $\mathfrak{S}$. To see this let $\mathfrak{S}$ be the convex hull of the extreme points ϕ_1, ϕ_2 and ϕ_3. Let an instrument $\mathcal{E}$ with values in $\mathfrak{U} = \{\mathfrak{u}_1, \mathfrak{u}_2\}$ given by

$$\begin{array}{cc} \mathcal{E}(1_{u_1})(\phi_1) = \phi_1 & \mathcal{E}(1_{u_2})(\phi_1) = 0 \\ \mathcal{E}(1_{u_1})(\phi_2) = 0 & \mathcal{E}(1_{u_2})(\phi_2) = \phi_2 \\ \mathcal{E}(1_{u_1})(\phi_3) = \frac{1}{2}\phi_1 + \frac{1}{2}\phi_2 & \mathcal{E}(1_{u_2})(\phi_3) = \frac{1}{2}\phi_1 + \frac{1}{2}\phi_2 \end{array} .$$

Then $\mathcal{E}(1)(\phi_1) = \phi_1$, $\mathcal{E}(1)(\phi_2) = \phi_2$ and $\mathcal{E}(1)(\phi_1) = \frac{1}{2}\phi_1 + \frac{1}{2}\phi_2$ which shows that $\mathcal{E}(1) = \frac{1}{2}\mathcal{E}_2(1) + \frac{1}{2}\mathcal{E}_2(1)$ where

$$\mathcal{E}_i(1)(\phi_1) = \phi_1 \quad \mathcal{E}_i(1)(\phi_2) = \phi_2 \quad \mathcal{E}_i(1)(\phi_3) = \phi_i \quad .$$

This shows that $\mathcal{M} = \frac{1}{2}\mathcal{M}_1 + \frac{1}{2}\mathcal{M}_1$, and therefore that $\mathcal{M}$ is not extreme and therefore not simple.

As we have seen earlier a repetitive instrument is completely determined by its measurement and its reduction. The following theorems gives a complete description of these.

Theorem 55 *Let $\mathcal{M}$ be a simple measurement given by the orthogonal resolution of the identity $\{M_B, B \subseteq \mathfrak{U}\}$. If the instrument $\mathcal{E}_{\mathcal{M}}$ exists then it is given by*

$$\mathcal{E}_{\mathcal{M}}(f)(S) = \left(\int f\, dM\right) \mathcal{E}_{\mathcal{M}}(1)\, S \tag{4}$$

and M_B is element in the commutator to $\mathcal{E}_{\mathcal{M}}(1)(\mathcal{W})$.

Proof. First we show that M_B are central.

$$tr\left(M_B \frac{\mathcal{E}_{\mathcal{M}}(1_B)(S)}{tr\left(\mathcal{E}_{\mathcal{M}}(1_B)(S)\right)}\right) = 1 ,$$

which shows that M_B commutes with $\mathcal{E}_{\mathcal{M}}(1_B)(S)$. Therefore M_B commutes with $\mathcal{E}_{\mathcal{M}}(1)(S) = \mathcal{E}_{\mathcal{M}}(1_B)(S) + \mathcal{E}_{\mathcal{M}}(1_{\complement B})(S)$, and M_B commutes with all density operators in $\mathcal{E}_{\mathcal{M}}(1)(\mathcal{W})$. This proves that (4) defines an instrument so it is sufficient to remark that $\mathcal{E}_{\mathcal{M}}(1) = id$, and that $\mathcal{M}_{\mathcal{E}_{\mathcal{M}}} = \mathcal{M}$, which is obvious. The general result is obtained using that a von Neumann algebra is the direct limes of its finite sub-algebras. ■

Theorem 56 *Let $\mathcal{E}$ be a repetitive instrument on a von Neumann algebra W. Then $\mathcal{E}(1)$ is given by a conditional expectation $\mathbb{E} : W \to W'$ where W' is a sub-algebra of W, and $\mathcal{E}(1)(\phi) = \phi \circ \mathbb{E}$.*

Proof. A proof appears in Davies (1976) and only needs minor changes. ■

A conditional expectation is typically of the form

$$\mathbb{E}(A) = \sum P_i A P_i$$

where P_i is the measurement corresponding to the instrument.

Theorem 57 *(Holevo 1986) For any measurement M in a Hilbert space $\mathcal{H}$ there exists a Hilbert space $\mathcal{K}$ with pure state S_r and a simple measurement E in $\mathcal{H} \otimes \mathcal{K}$ such that*

$$\mu^E_{S \otimes S_r}(D) = \mu^M_S(D)$$

for any state S on $\mathcal{H}$.

Theorem 58 *For any measurement M_B in $\mathfrak{A} \subseteq \mathbb{B}(\mathcal{H})$ there exists a Hilbert space $\mathcal{L}$ containing $\mathcal{H}$ and a simple measurement E_B in $\mathbb{B}(\mathcal{L})$ such that $M_B = PE_BP$ where P is the projection of $\mathcal{L}$ on $\mathcal{H}$.*

Proof. Let $\mathcal{L}$ be the set of mappings $\mathfrak{U} \to \mathcal{H}$. Put

$$\langle f \mid g \rangle = \sum_u \left(f(u) \mid M_{\{u\}}(g(u)) \right) .$$

Then $\mathcal{L}$ is a Hilbert space. Let denote the completion of $\mathcal{L}$ with respect to $\langle \cdot \mid \cdot \rangle$. The map $l : v \to (u \to v)$ is an isometry of $\mathcal{H}$ into $\mathcal{L}$, and $\mathcal{H}$ can be identified with a subspace of $\mathcal{L}$. Let M_u be the operator $f \to f \cdot 1_u$. Then E obviously is an orthogonal resolution of the identity. For $v \in \mathcal{H}$ we have

$$\left(v \mid M_{\{u\}} v \right) = \left\langle l(v) \mid E_{\{u\}}(l(v)) \right\rangle$$

which proves that $M_B = PE_BP$. The general result can be obtained by going to the limit.

Let E be a resolution of the identity in a Hilbert space $\mathcal{L}$ containing $\mathcal{H}$ such that $M = PEP$. Then $\mathcal{L}$ is isomorphic with $\mathcal{H} \otimes L^2(\mathfrak{U})$, and there exists an isomorphism such that $\mathcal{H} = \mathcal{H} \otimes 1_u$. Put $S_0 = |1_u\rangle\langle 1_u|$. For $X = E(D)$ we have $Tr(S \otimes S_0) X = Tr(SPXP) = Tr(SM(D))$. ■

9 Entropy

Let $P = (p_1, p_2, ..., p_n)$ be a probability vector. Then the entropy of P is given by

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i .$$

Here we have used the convention $0 \log 0 = 0$. The function $P \to H(P)$ is positive, concave and continuous because each of the functions $p_i \to -p_i \log p_i$ are concave and continuous. Here we will use the natural logarithm. Then $H\left(\frac{1}{2}, \frac{1}{2}\right) = \log 2$ and we say that the entropy of an experiment with the possible outcomes 0 and 1 each with probability $\frac{1}{2}$ is one *bit*. Using that $H(P)$ is concave and symmetric in its arguments we see that $H(P)$ is minimal on the deterministic distributions and maximal with value n on the uniform distribution. A sequence of length d of independent zeros and ones each with probability $\frac{1}{2}$ has entropy $-2^d \cdot \frac{1}{2^d} \log \frac{1}{2^d} = d$. According to Shannon's first coding theorem the value of a random variable with entropy H can be encoded in a sequence of zeros and ones of (mean) length approximately equal to H.

Now let a state be given by a density matrix $S = \sum_{i=1}^{d} \lambda_i P_i$ where λ_i are eigenvalues and P_i are projections to corresponding eigenvectors. Then $\sum_{i=1}^{d} \lambda_i = 1$ and $\lambda_i \geq 0$. Now P_i is a resolution of the identity and the corresponding measurement maps S into the probability vector $(\lambda_1, \lambda_2, ..., \lambda_d)$. Then

$$\begin{aligned} H(\lambda_1, \lambda_2, ..., \lambda_d) &= -\sum_{i=1}^{d} \lambda_i \log \lambda_i \\ &= -Tr(S \log S) \ . \end{aligned}$$

This is called the *entropy* of S and is denoted $H(S)$ We see that the entropy satisfy

- $H(S) \geq 0$ with equality if and only if S is a pure state.
- $H(S) \leq d$ with equality if and only if $S = \frac{\mathbf{1}}{d}$.

Later we shall see how to encode a quantum state with entropy $H(S)$ into approximately $H(S)$ qubits.

The entropy only depend on the eigenvalues of the state therefore the entropy is invariant under a unitary transformation.

$$\begin{aligned} H(USU^*) &= -Tr(USU^* \log(USU^*)) \\ &= -Tr(US \log(S) U^*) \\ &= -Tr(S \log S) \\ &= H(S) \end{aligned}$$

Especially the entropy is invariant under time shifts. This seems to contradict the second law of thermodynamics. We also introduce the information divergence (often called relative entropy) by the equation

$$D(S \parallel T) = Tr(S(\log S - \log T)).$$

Proposition 59 *Let S and T be densities of states, and let $\alpha, \beta \geq 0$ be numbers with $\alpha + \beta = 1$ Then*

$$\begin{aligned} H(\alpha S + \beta T) = &\alpha H(S) + \beta H(T) \\ &+ \alpha D(S \parallel \alpha S + \beta T) + \beta D(T \parallel \alpha S + \beta T) \ . \end{aligned}$$

Proof. The proof is a simple exercise in the definitions and is left to the reader. ■

Theorem 60 *For density operators S and T the following inequality holds*

$$D(S \parallel T) \geq \frac{1}{2} Tr\left((S - T)^2\right).$$

Proof. Put $\eta(t) = -t\log t, t \in]0;1]$. Then

$$\begin{aligned}\eta'(t) &= -\log t - 1\\ \eta''(t) &= -\frac{1}{t} \leq -1\ .\end{aligned}$$

Then a Taylor expansion gives

$$\begin{aligned}\eta(x) &= \eta(y) + (x-y)\eta'(y) + \frac{1}{2}(x-y)^2\eta''(\theta)\\ &\leq \eta(y) + (x-y)\eta'(y) - \frac{1}{2}(x-y)^2\end{aligned}$$

For some θ between x and y. Therefore

$$\begin{aligned}0 &\leq -\eta(x) + \eta(y) + (x-y)\eta'(y) - \frac{1}{2}(x-y)^2\\ &= x\log x - x\log y + y - x - \frac{1}{2}(x-y)^2\ .\end{aligned}$$

Let $S = \sum \lambda_i P_i$ and $T = \sum \kappa_j Q_j$ be spectral decompositions of the operators S and T. Then

$$\begin{aligned}0 &\leq \left(\lambda_i\log\lambda_i - \lambda_i\log\kappa_j + \kappa_j - \lambda_i - \frac{1}{2}(\kappa_j-\lambda_i)^2\right) Tr\left(P_iQ_j\right)\\ &= Tr\left(\left(\lambda_i\log\lambda_i - \lambda_i\log\kappa_j + \kappa_j - \lambda_i - \frac{1}{2}(\kappa_j-\lambda_i)^2\right)P_iQ_j\right)\ .\end{aligned}$$

Summing over i and j gives

$$\begin{aligned}0 &\leq Tr\left(\sum_{i,j}\begin{pmatrix}\lambda_i\log\lambda_i P_iQ_j - \lambda_i\log\kappa_j P_iQ_j\\ +\kappa_j P_iQ_j - \lambda_i P_iQ_j - \frac{1}{2}(\kappa_j-\lambda_i)^2 P_iQ_j\end{pmatrix}\right)\\ &= Tr\begin{pmatrix}\sum_i \lambda_i\log\lambda_i P_i - \sum_{i,j}\lambda_i P_i\log\kappa_j Q_j\\ +\sum_i\kappa_i - \sum_j\lambda_j - \sum_{i,j}\frac{1}{2}(\kappa_j-\lambda_i)^2 P_iQ_j\end{pmatrix}\\ &= Tr\left(S\log S - S\log T - \frac{1}{2}(S-T)^2\right)\end{aligned}$$

and the result follows. ■

The theorems shows that divergence is non-negative and therefore that the entropy is concave.

10 List of notation

$\mathcal{A}, \mathcal{B}$	$*$-Algebras.
a_g	The action of the group element g on a state space.
$\mathbb{B}(\mathcal{H})$	Matrices/bounded operators on the Hilbert space $\mathcal{H}$.
$\mathbb{B}_+^1(\mathcal{H})$	Density matrices/operators on the Hilbert space $\mathcal{H}$.
C, K	Convex sets.
d	Dimension.
$\det(X)$	The determinant of X.
δ_x	Dirac measure, i.e. probability measure with all weight in x.
Δ	The unit disc $\left\{(a,b) \in R^2 \mid a^2 + b^2 \leq 1\right\}$.
ϕ, ψ	Abstract states or states on a $*$-algebra.
G	A group.
$\mathcal{H}, \mathcal{K}$	Hilbert spaces.
$M_+^1(U)$	The set of probability vectors (or probability measures) on U.
$\|\|\|_{tot}$	The norm total variation.
$\|\|\|_2$	The 2 norm in a Hilbert space.
$\mathbb{N}$	The set of natural numbers.
$(\cdot \mid \cdot), \langle \cdot \mid \cdot \rangle, \cdot$	Inner products in Hilbert spaces.
$\times$	Product set or product measure.
$Sp(X)$	The spectrum of X.
Tr	Trace. The sum of the diagonal element of a matrix.
$P_{char}(X)$	The characteristic polynomial of X.
$\mathcal{P}$	Set of preparations.
$\mathbb{R}$	Set of real numbers.
$\mathbb{S}(\mathcal{A})$	States on the algebra $\mathcal{A}$.
S, T	Density matrices or density operators.
$\mathfrak{S}$	State space.
$\otimes$	Tensor product of Hilbert spaces or vectors or algebras.
$\mathbb{T}$	$\mathbb{R}/2\pi\mathbb{Z}$ can be identified with $[0; 2\pi[$.
$\vec{u}, \vec{v}, \vec{w}$	Vectors in a real or complex Hilbert space.
$\mathbb{Z}$	The set of integers.

References