Preamble

This is the draft of a chapter of our larger work Assumptions of physics.

Contents

Contents 1

1 Quantum mechanics 3

1.1 Notes . . . 3

Proposed nomenclature . . . 3

Sketch of results . . . 3

Conditions . . . 4

Proposed index . . . 5

Formulations of quantum mechanics . . . 7

The postulates of quantum mechanics . . . 7

Bloch sphere review . . . 9

The problem with the current formulations . . . 10

1.2 States and projective space . . . 11

Vector vs projective space . . . 11

Quantum state representations . . . 14

Quantum ensembles representations . . . 15

Two linearities: superposition and statistical mixing . . . 17

The gap between the two linearities . . . 19

Basis and gauge choice . . . 20

Probability from expectations . . . 21

Born rule and angles in projective space . . . 22

Entropy and the Born rule . . . 23

1.3 Schroedinger equation and unitary evolution . . . 24

Time evolution is unitarity . . . 24

Time evolution is preservation of the inner product . . . 25

Time evolution is preservation of the norm . . . 25
Time evolution is preservation of the Born rule . . . 26
States and states change are perpendicular on the complex plane . . . 27
Time evolution is preservation of bases orthonormality . . . 27
Physical conditions . . . 28
Determinism and Reversibility . . . 28
Preservation of probability distributions . . . 28
Preservation of information entropy . . . 28
Quasi-static processes . . . 29
1.4 Projection and measurements . . . 30
1.5 Observables . . . 30
Observables are self-adjoint operators (textbook definition) . . . 30
Observables are generators of unitary transformations . . . 30
WORK IN PROGRESS: Observables have real spectrum . . . 31
WORK IN PROGRESS: Observables are linear functionals of observables . . . 32
TODO list . . . 32
1.6 next . . . 32
1.7 Problems with infinite dimensional spaces . . . 33
Equivalence of Hilbert spaces . . . 34
Domain of operators . . . 35
Continuous spectra . . . 37
Schwartz spaces . . . 38
Probability on a continuum . . . 38
Self-adjoint vs Hermitian . . . 38
Wavefunctions and equivalent classes . . . 38
Schauder basis . . . 38
Physical states and well-defined moments . . . 39
Example for pure states . . . 40
Operator domain and the second moment . . . 40
1.8 WARNINGS . . . 41
1.9 Open problem: Extremality of quantum ensembles . . . 41
1.10 Open problem: Spectral theorem on Fréchet spaces . . . 41
1.11 Open problem: Ensembles for Schwartz spaces . . . 42

Bibliography **43**

Chapter 1

Quantum mechanics

1.1 Notes

In this section we will review the standard formulations of quantum mechanics in terms of projective Hilbert spaces.

Proposed nomenclature

- State vector $|\psi\rangle \in \mathcal{H}$
- ? Wave function $\psi(x)$ $\psi(s_z)$
- State $\psi \in P(\mathcal{H})$
- Ensemble $\rho \in D(\mathcal{H})$
- ? Space of observables, space of bounded operators, space of unitaries, ...

Sketch of results

How to break up the different assumptions for quantum states.

QP Space of pure states by themselves tells us nothing. Complex projective spaces are symplectic spaces, and are symplectic manifolds in finite dimension. We can write Hamiltonians and do classical mechanics with probability distributions over the states.

QE Convex space of density operators tells you more, but not everything. We know they form a convex subset of a vector space. We know which ensembles allow multiple decompositions and we can define observables as statistical quantities (i.e. linear functionals of the ensembles). However, we do not know what the entropy is. For example, take the Bloch ball with the standard entropoy. Take an inner ball and discard the rest. Now, the entropy difference between pure states and maximally mixed state is not 1, therefore a variable cannot extract one bit of information. This means that opposite states are not orthogonal (i.e. you could mistake up for down with a single measurement).

BR Conjecture. If we add orthogonal decomposability, we should be able to recover the Born rule, and the Born rule tells us ensembles are orthogonally decomposable. One should be able to show that all "effects" (i.e. linear functionals of ensembles that are either zero or one in the edges orthogonal to the gradients) have orthogonal eigenspaces of either zero or one, and therefore can extract one bit of information. That should be enough to define the entropy, find the Born rule, etc...

Conditions

Condition QST (Quantum states)**.** *The state space of the system is represented by the projective space of a complex inner product space $\mathcal{H}$. The ensemble space is represented by the density operators. The conditional probability $p(\psi|\phi)$ is given by the Born rule.*

Condition QP-RAY (Quantum states are rays)**.** *The state space of the system is represented by the projective space given by the rays of a complex vector space $\mathcal{H}$. That is, a state $\psi \in P(\mathcal{H})$ is a ray $\psi = \{c|\psi\rangle \mid c \in \mathbb{C}\}$ of a complex projective space.*

Condition QE-DENS (Quantum ensembles are density operators)**.** *The ensemble space of the system is represented by the set of density operators acting on a complex inner product space $\mathcal{H}$. That is, an ensemble $\rho \in D(\mathcal{H})$ is a positive semi-definite trace one self-adjoint operator $\rho : \mathcal{H} \to \mathcal{H}$.*

Condition QE-PROJ (Quantum pure ensembles are projectors)**.** *A pure ensemble is represented by a one dimensional projector of a complex inner product space $\mathcal{H}$. That is, a pure ensemble $\rho_\psi = \mathbf{1}_\psi \in P(\mathcal{H})$ is a projector $\mathbf{1}_\psi = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}$.*

Condition QE-IDEM (Quantum pure ensembles are idempotent)**.** *A pure ensemble is represented by an idempotent density operator. That is, $\rho_\psi \in P(\mathcal{H}) \subset D(\mathcal{H})$ if and only if $\rho_\psi = \rho_\psi^2$.*

Condition QE-EXTR (Quantum pure ensembles are extreme points)**.** *A pure ensemble is represented by a density operator that cannot be expressed as a non-trivial convex combination of other density operators. That is, $\rho_\psi \in P(\mathcal{H}) \subset D(\mathcal{H})$ if and only if $\rho_\psi = p\rho_1 + (1-p)\rho_2$ implies $\rho_1 = \rho_2 = \rho_\psi$.*

Conjecture 1.1. *The expectation of a statistical variable is the trace of some self-adjoint operator. That is, let $E[X \mid \cdot] : D(\mathcal{H}) \to \mathbb{R}$ be an affine continuous functional over density operators, then there exists a self-adjoint operator $X : \mathcal{H} \to \mathcal{H}$ such that $E[X \mid \rho] = \mathrm{tr}(X\rho)$.*

Condition MUTEX-QM (Mutual exclusivity conditions)**.** *Mutually exclusive states have zero conditional probability and maximize the entropy during mixing. That is, let $\psi_i \in P(\mathcal{H})$ be a countable sequence of states. Then $p(\psi_i|\psi_j) = 0$ for all $i \neq j$ if and only if $S(\sum_i p_i\psi_i) = -\sum_i p_i \log p_i$ for all p_i such that $\sum_i p_i = 1$.*

Condition BR-BORN (Conditional expectation from Born rule)**.** *The conditional probability $p : P(\mathcal{H}) \times P(\mathcal{H}) \to [0,1]$ is given by the Born rule. That is, $p(\psi|\phi) = \frac{\langle\phi|\psi\rangle\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}$.*

Condition BR-ANG (Conditional expectation from projective angle)**.** *The conditional probability $p : P(\mathcal{H}) \times P(\mathcal{H}) \to [0,1]$ is given by the cosine square of the projective angle. That is, $p(\psi|\phi) = \cos^2 \theta_{\psi\phi}$.*

Condition BR-IND (Conditional expectation is the expectation of projectors)**.** *The conditional probability $p : P(\mathcal{H}) \times P(\mathcal{H}) \to [0,1]$ is given by the expectation of the corresponding projector. That is, $p(\psi|\phi) = E_{\rho_\phi}[\mathbf{1}_\psi] = \mathrm{tr}(\mathbf{1}_\psi\rho_\phi)$.*

Condition BR-INPR (Conditional expectation is the projectors' inner product)**.** *The conditional probability $p : P(\mathcal{H}) \times P(\mathcal{H}) \to [0,1]$ is given by the inner product between projectors. That is, $p(\psi|\phi) = \mathrm{tr}(\mathbf{1}_\psi\mathbf{1}_\phi)$.*

Condition BR-MUTEX (Orthogonality is mutual exclusivity)**.** *Two states are mutually exclusive if and only if they are orthogonal. That is, ψ and ϕ are mutually exclusive if and only if $\langle\psi|\phi\rangle = 0$.*

Condition BR-ENT (Von Neumann entropy)**.** *The entropy $S : D(\mathcal{H}) \to \mathbb{R}$ for an ensemble $\rho \in D(\mathcal{H})$ is given by the Shannon entropy. That is, $S(\rho) = -\operatorname{tr}(\rho \log \rho)$.*

Condition DR-SCEQ (Schroedinger equation)**.** *The evolution follows the equation $\imath\hbar\frac{d}{dt}\psi(t) = H\psi(t)$ where H is a self-adjoint operator.*

Condition DR-UNIT (Unitary evolution)**.** *The evolution is unitary: $U_{dt}^{\dagger}U_{dt} = U_{dt}U_{dt}^{\dagger} = 1$.*

Condition DR-INN (Inner product preservation)**.** *The evolution preserves the inner product : $\langle U_{dt}\phi|U_{dt}\psi\rangle = \langle\phi|\psi\rangle$.*

Condition DR-NORM (Linear and norm preservation)**.** *The evolution is linear and preserves the norm: $\langle\psi(t)|\psi(t)\rangle = \langle\psi(t+dt)|\psi(t+dt)\rangle$.*

Condition DR-UBOR (Unitary Born rule)**.** *The square of the inner product between to states infinitesimally close in time is one: $|\langle\psi(t)|\psi(t+dt)\rangle|^2 = 1$.*

Condition DR-PERP (Perpendicular change)**.** *The state change is perpendicular to the original state in the complex plane: $\langle\psi(t)|d\psi(dt)\rangle$ is imaginary.*

Condition DR-OBAS (Orthogonality preservation)**.** *The evolution is linear and maps orthonormal basis to orthonormal basis: $\langle U_{dt}e_i|U_{dt}e_j\rangle = \langle e_i|e_j\rangle = \delta_{ij}$.*

Condition DR-EV (Determinism and reversibility)**.** *The evolution is deterministic and reversible*

Condition DR-PROB (Probability preservation)**.** *The evolution preserves probability distributions.*

Condition DR-MUTEX (Mutual exclusivity preservation)**.** *The evolution preserves mutual exclusivity between ensembles.*

Condition DR-INFO (Information preservation)**.** *The evolution preserves information entropy.*

Condition DR-PSEQ (Quasi-static process)**.** *Time evolution is a quasi-static process*

Condition DR-MSEQ (Series of measurements)**.** *The evolution is an infinite sequence of reversible measurements.*

Proposed index

1. Review postulates
2. States
    a) BT: a set of ensemble that can be mixed and have an entropy (discovered? Probably...)
    b) Pure states

QP-RAY A pure quantum state is represented by a ray
QP-EXTR A pure quantum state is represented by a density matrix with only one non-zero eigenvector
QP-PROJ A pure quantum state is represented by a one-dimensional projector
i. QP-RAY $\iff$ QP-EXTR $\iff$ QP-PROJ over BT

c) Ensembles

QE-DENS A quantum ensemble is represented density operator
QE $\implies$ QP Recover pure states as extreme points
QP $\not\implies$ QE We could imagine ensembles as probability distributions
? Superposition is multiple decomposition (Comparing linear structure of ensembles vs vectors)
? The linearity of statistical mixing is a physical requirement.
? The linearity of the Hilbert space is a mathematical convenience.
? The choice of a basis in the Hilbert space is a choice of a maximal set of orthogonal states (not necessarily mutually exclusive) plus a choice of gauge (i.e. a phase for each element of the basis).
? Pure states are not enough
? Decompose state requirements into 3
? Inner product is about geometry, shouldn't be about process

d) Born rule

BR-BORN Probability of measuring final states given initial states given by the born rule
BR-PROB The probability of measuring a final state given an initial state is well-defined.
BR-ENS Probability given by trace of product of pure states
BR-ANG Probability given by cosine of angles on Bloch sphere
BR-ENT Entropy is von Neumann entropy
BR-ME Orthogonality is mutual exclusivity
BR-EXP Expectations are well defined (?)
QST QE + BR
QP+BR $\implies$ QE Recover indistinguishable mixtures
QE $\not\implies$ BR We could have no orthogonal elements
i. Inner product as overlap between ensembles
ii. Projectors as indicator functions/support (and quantum logic?)

3. Observables

a) not all quantities are "Observables". For example, the x-y angle on a Bloch sphere is a measureable quantity, which we can also use to track state evolution, but it is not an observable (it is a parameter to identify a pure state); the temperature is a measurable quantity, but it is not an observable (it is a parameter to identify a mixed state from a family of mixed states); the four velocity for a relativistic spin is a measurable quantity, it has an operator associated with it (i.e. the gamma matrices) but the possible values are NOT the spectra of that operator (i.e. the only eigenstates are +c and -c); the entropy is well defined for all ensembles, but entropy of a mixture is not the average entropy

OBS-SELF Observables are represented by self-adjoint operators
b) Not all hermitian operators are self-adjoint

OBS-GEN Self-adjoint operators are exactly generators of one-parameter families of unitary operators

OBS-REAL Self-adjoint operators are exactly those with real spectrum

OBS-EXP Observables are those quantities that admit an expectation (i.e. they are real linear functionals of density operators)

QE-OBS The set of random variables over probability distributions that are mapped univocally under multiple decompositions fully identify the space of quantum ensembles (i.e. we can either start with the space of quantum ensembles and rule out those random variables that are not the same over equivalence classes of probability distributions, or we can start with the "nice" random variables, and find the ensembles as equivalence classes of probability distributions).

4. Measurements
5. Time evolution

DR-SCEQ The equation

DR-UNIT Unitary evolution

DR-INN Preservation of inner product

DR-NORM Preservation of norm

DR-UBOR Square of inner product infinitesimally close

DR-PERP Perpendicular change

DR-OBAS Preserves orthogonal basis

DR-EV Deterministic and reversible evolution

DR-PROB Probability distribution preserved

DR-INFO The evolution preserves information entropy

DR-PSEQ The evolution is quasi-static process

DR-MSEQ The evolution is an infinite sequence of reversible measurements.

PM-DEQ Projection measurements are process with equilibria that commute with a det/rev process.

- Unphysicality of non-Hermitian Hamiltonians

6. Composite system * Show that system composition of quantum systems that is itself a quantum system gives the tensor product * Review the arguments in the paper * see if using a Segre embedding would make our life easier
7. Fermion and Bosons as indistinguishibility of components (requires tensor product)

In QST, infinite dimension

- A pure state is in the domain of an observable if and only if the expectation of the second moment is finite

Formulations of quantum mechanics

The postulates of quantum mechanics

Let us first review the standard Hilbert space formulation of quantum mechanics. The goal is to have a general understanding of the math and how it naively maps to the physics. We will also define the notation that we are going to use throughout this chapter.

The first postulate of quantum mechanics characterizes how states are represented in the mathematical framework.

Postulate ST (State postulate)**.** *The state space of a quantum system is represented by the projective space* $\mathrm{P}(\mathcal{H})$ *of a complex Hilbert space* $\mathcal{H}$*. A state is represented by a ray of the Hilbert space* $\mathcal{H}$*. An ensemble is represented by a density operator* $\rho : \mathcal{H} \to \mathcal{H}$*, that is a positive semi-definite trace one self-adjoint operator.*

In terms of notation, we will mostly follow the convention most used by physicists, with some addition or modification to make the connection to the math more clear. A state $\psi \in \mathrm{P}(\mathcal{H})$ is a ray, which is an equivalence class $\psi = \{c|\psi\rangle \mid c \in \mathbb{C} \setminus 0\}$ for some vector $|\psi\rangle \in \mathcal{H}$. A wavefunction $\psi(x^i)$ is the representation of a vector $|\psi\rangle$, where x^i represent the possible values (i.e. the joint spectrum) of a complete set of commuting observables. Note that, over a continuous spectrum, such representation is not unique as it can differ on a set of measure zero. We will note $\mathrm{D}(\mathcal{H})$ for the set of all mixed states, and will consider $\mathrm{P}(\mathcal{H})$ a subset of $\mathrm{D}(\mathcal{H})$ since every ray is associated to a unique density operator. If we want to specify the density operator for a particular state, we will write ρ_ψ.

Mathematically, a complex Hilbert space $\mathcal{H}$ has three defining properties. It is a vector space, meaning that it is closed under addition $+ : \mathcal{H} \times \mathcal{H} \to \mathcal{H}$ and scalar multiplication $\cdot : \mathbb{C} \times \mathcal{H} \to \mathcal{H}$. It has an inner product, meaning that it is equipped with a map $\langle \cdot \mid \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ that is positive-definite, linear in the second argument and conjugate symmetric. It is a complete metric space in the sense that every Cauchy sequence converges, which can be characterized in the following way: given a sequence of vectors $|\psi_i\rangle$, if the series given by their norm $\sum_i \sqrt{\langle\psi_i|\psi_i\rangle}$ converges, then the series given by the vectors $\sum_i |\psi_i\rangle$ converges as well.

The second postulate of quantum mechanics characterizes observables.

Postulate OBS (Observable postulate)**.** *An observable is represented by a self-adjoint operator* $O : \mathcal{D}(O) \subseteq \mathcal{H} \to \mathcal{H}$*. Given a state* $|\psi\rangle$*, the expectation is given by* $\frac{\langle\psi|O|\psi\rangle}{\langle\psi|\psi\rangle}$*, and, given an ensemble* ρ*, the expectation is given by* $\mathrm{tr}(O\rho)$*.*

The expectations are guaranteed to be real numbers precisely because O is self-adjoint. Also note that, given the completeness of $\mathcal{H}$, a self-adjoint operator is defined over the whole space if and only if it is bounded. In physics, we are also interested in unbounded observables, like position, momentum, energy, number of particles and so on. In this case, we require that the domain $\mathcal{D}(O)$ is dense in $\mathcal{H}$.

The third postulate characterizes the effect of projective measurements.

Postulate PROJ (Projection measurement postulate)**.** *A projection measurement is defined by a set* $M = \{\mathbf{1}_i\}$ *of projectors, representing the measurement outcomes, such that* $\sum_i \mathbf{1}_i = I$*. If* $|\psi\rangle \in \mathcal{H}$ *represents the state before the measurement interaction, the ensemble after the interaction will be* $\rho_M = \sum_i \mathbf{1}_i \mathbf{1}_\psi \mathbf{1}_i$ *and the state after the measurement of the* i*-th outcome is* $|\phi_i\rangle = \frac{\mathbf{1}_i|\psi\rangle}{\sqrt{\langle\psi|\mathbf{1}_i|\psi\rangle}}$ *with probability* $\frac{\langle\psi|\mathbf{1}_i|\psi\rangle}{\langle\psi|\psi\rangle}$*. If* ρ *is the ensemble before the measurement interaction, the ensemble after the interaction will be* $\rho_M = \sum_i \mathbf{1}_i \rho \mathbf{1}_i$ *and the state after the measurement of the* i*-th outcome is* $\hat{\rho}_i = \frac{\mathbf{1}_i \rho \mathbf{1}_i}{\mathrm{tr}(\mathbf{1}_i \rho)}$ *with probability* $\mathrm{tr}(\mathbf{1}_i \rho)$*.*

Note that we use $\mathbf{1}_V$ for a projector onto subspace V. Projectors in quantum mechanics play a similar role to indicators in classical probability, so this notation will allow us to make the parallel more evident, and show exactly what works the same and what works differently. Given a state ψ, the corresponding projector is $\mathbf{1}_\psi = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}$.

The projection postulate implicitly defines the Born rule. Given a state ϕ, the observable that returns one if the state is ϕ and zero if not is exactly $O_\phi = \mathbf{1}_\phi$. This means that the probability of measuring ϕ given ψ is $p(\phi|\psi) = \frac{\langle\psi|\mathbf{1}_\phi|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\langle\psi|\phi\rangle\langle\phi|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}$. Note the expression gives the same result for all elements of the same ray. Moreover, since the inner product over continuous spectrum is an integral, two wavefunctions that differ over a set of measure zero will always give the same result. In this sense, the mathematics is consistent as all physical outcomes depend only on the rays.

The fourth postulate characterizes time evolution.

Postulate UNIT (Time evolution postulate)**.** *Time evolution of a state vector $|\psi\rangle$ is given by $d_t|\psi(t)\rangle = \frac{H}{\imath\hbar}|\psi(t)\rangle$ where H is an observable, typically called Hamiltonian. The time evolution of an ensemble ρ is given by $d_t\rho = \frac{[H,\rho]}{\imath\hbar}$.*

Note that we will tend to group $\imath\hbar$ in the denominators as it makes the pattern given by the Lie algebraic structure and its connection to classical mechanics more clear. For example:

$$\begin{aligned}
\imath\hbar d_t|\psi(t)\rangle = H|\psi(t)\rangle &\implies d_t|\psi(t)\rangle = \frac{H}{\imath\hbar}|\psi(t)\rangle \\
e^{-\frac{\imath}{\hbar}Ht} &\implies e^{\frac{Ht}{\imath\hbar}} \\
d_tO = \frac{\imath}{\hbar}[H,O] &\implies d_tO = \frac{[O,H]}{\imath\hbar} \\
d_t\rho = -\frac{\imath}{\hbar}[H,\rho] &\implies d_t\rho = \frac{[H,\rho]}{\imath\hbar} \\
[X,P] = \imath\hbar &\implies \frac{[X,P]}{\imath\hbar} = 1 \\
&\dots
\end{aligned} \tag{1.2}$$

Not only the notation is more consistent across equations, it brings out the connection between the Poisson bracket and commutator divided by $\imath\hbar$. In particular, the order/sign difference between the evolution of an observable O and a mixed state ρ is exactly the same in classical mechanics between a quantity and the probability density.

The last postulate characterizes composite systems.

Postulate COMP (Composition postulate)**.** *Given two quantum systems associated with the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, their composite system is associated with the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$. An observable O_1 for the first system becomes $O_1 \otimes I_2$ in the composite, and an observable O_2 for the second system becomes $I_1 \otimes O_2$, where I_1 and I_2 are the identities in the respective spaces.*

Bloch sphere review

The state space of a two-state quantum system (i.e. qubit) is of fundamental importance as it allows us to visualize and therefore better understand almost all features of quantum mechanics. The space of pure states is isomorphic to a unit 2-sphere, with opposite points corresponding to orthogonal states, meaning that we can understand each pure state as a unit vector along a direction of three dimensional space. Since this is also the space of a spin

1/2 system, we will often label states as spatial directions. Let us quickly review the main features.

Let $|z^+\rangle$ and $|z^-\rangle$ be two orthonormal states, representing, for example, the two possible directions of spin along the vertical direction. Every other normalized state can be expressed as $\rho|z^+\rangle + \sqrt{1-\rho^2}e^{\imath\varphi}|z^-\rangle$ with $\rho \in [0,1]$ and $\varphi \in [0, 2\pi)$. Normalization, in fact, imposes that the square of the norm of the coefficient sum to one. The arbitrariness of the phase allows us to impose the first coefficient to be real. Note that, for any $p \in (0,1)$, we can choose an arbitrary φ, meaning that we are choosing a point on a circle. At the endpoints $p \in \{0,1\}$, however, there is no longer a choice of φ. The topology, then, is the one of the sphere. We can therefore draw a sphere with $|z^+\rangle$ and $|z^-\rangle$ respectively at the top and at the bottom. The choice of ρ will select a horizontal plane, whose intersection with the sphere will be a circle. The choice of φ will pick a point on that circle. That is, ρ and φ can be understood as a sort of "cylindrical coordinates over a sphere." We can also choose spherical coordinates (θ, φ) such that $\rho = \cos\theta$.

TODO: show that ensemble parametrization in spherical coordinates is not affine (meaning the mixture of the ensembles does not give the mixture of the coordinates)

More remarkably, the interior points of the sphere represent mixed states. Therefore the whole space, the Bloch ball, represents the full space of ensembles for a two-state quantum systems. Given two ensembles ρ_1 and ρ_2, the statistical mixture $\rho = p\rho_1 + (1-p)\rho_2$ will be the point in between the two. The coefficient for each endpoint is higher the closer the mixture is to that point, meaning it is proportional to the distance to the other endpoint. That is,

$$\rho = \frac{\overline{\rho\rho_2}}{\overline{\rho_1\rho_2}}\rho_1 + \frac{\overline{\rho_1\rho}}{\overline{\rho_1\rho_2}}\rho_2. \tag{1.3}$$

The problem with the current formulations

The goal of Reverse Physics is finding exactly which mathematical condition correspond to which physical premise. We will want, for example, to understand which physical assumptions lead to the complex projective space or to unitary evolution. This will be a lot hard for quantum mechanics than for classical mechanics given the current status of the theory. First of all, there is no general consensus as to what quantum states represent, and therefore we will need to find a minimal set of operational characterizations on which every interpretation must agree. Second, the mathematics itself does not cleanly separate what elements should be considered physically meaningful and what simply mathematical artifacts. But most of all, the mathematical formulation bundles many aspects together, making it difficult to implement a careful analysis. Therefore, we will need to first disentangle the standard formulation into a set of distinct conditions, so that we can better analyze their logical relationships and find alternative but equivalent premises.

Unlike in classical mechanics, then, where we could essentially find the conditions organically, we need to set some conditions upfront. The choice will only be clear in further sections, when we'll take each condition and analyze it further. Let us go through them with a minimal motivation.

Condition QP-RAY (Quantum states are rays)**.** *The state space of the system is represented by a complex projective space $P(\mathcal{H})$. That is, a state $\psi \in P(\mathcal{H})$ is a ray $\psi = \{c|\psi\rangle \mid c \in \mathbb{C} \setminus \{0\}\}$ of some complex vector space $\mathcal{H}$.*

Note that the vector space itself is not the primary object, but simply a way to characterize the projective space. Also note that $\mathcal{H}$ is just a complex vector space, with no inner product and not necessarily complete. None of these are necessary to characterize the complex projective space.

Condition QE-DENS (Quantum ensembles are density operators)**.** *The ensemble space of the system is represented by the set of density operators acting on a complex inner product space $\mathcal{H}$. That is, an ensemble $\rho \in D(\mathcal{H})$ is a positive semi-definite trace one self-adjoint operator $\rho : \mathcal{H} \to \mathcal{H}$.*

The characterization of ensembles is separate from that of pure states, as we will need to understand whether one implies the other. Note that $\mathcal{H}$ is equipped with an inner product as the definition of trace one operators requires an inner product.

Condition BR-BORN (Conditional expectation from Born rule)**.** *The conditional probability $p(\cdot \mid \cdot) : P(\mathcal{H}) \times P(\mathcal{H}) \to [0,1]$ is given by the Born rule. That is, $p(\psi|\phi) = \frac{\langle\phi|\psi\rangle\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}$.*

The Born rule links the inner product to the conditional probability.

Condition QM-OBS (Observables are self-adjoint operators)**.** *The space of observables is given by the continuous self-adoint operators defined over the whole state space. Given a state vector $|\psi\rangle$, the expectation is given by $\frac{\langle\psi|O|\psi\rangle}{\langle\psi|\psi\rangle}$, and, given an ensemble ρ, the expectation is given by $\mathrm{tr}(O\rho)$.*

In contrast with postulate OBS, we require self-adjoint to be continuous and defined on the whole space. As we will see later, these seem the correct physical requirements: the expectation must exist for all states (whole space), be a real numbers (self-adjoint) and lead to small changes for small state changes. Since these cannot

We'll proceed to analyze each one, its relationships with the others and other possible equivalent characterizations. We will see that they form a sequence of increasingly stronger conditions, where each is implied by the next but not vice-versa.

1.2 States and projective space

In this section we will analyze the state postulate to understand what each mathematical structure represent. The complex vector space structure by itself defines pure states, but nothing else. The inner product structure is needed to define ensembles and statistical mixing, but does not require the Born rule. The Born rule links the inner product with probability, linking orthogonality with mutual exclusivity. (TODO: revise summary once the section is finished)

Vector vs projective space

Note that QP-RAY asserts that the state space is given by the projective space, and the role of the vector space is only to identify the correct projective space. We should therefore verify that physics only depends on the ray, and, since the ray is an equivalence class, not the choice of representative. Given that all physical outcomes are given by the Born rule

$$p(\phi|\psi) = \frac{\langle\phi|\psi\rangle\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}, \tag{1.4}$$

we just need to verify that the Born rule is indeed invariant under the change of representative $|\psi\rangle$ to $c|\psi\rangle$ with $c \in \mathbb{C} \setminus \{0\}$:

$$p(\phi|c\psi) = \frac{\langle\phi|c\psi\rangle\langle c\psi|\phi\rangle}{\langle c\psi|c\psi\rangle\langle\phi|\phi\rangle} = \frac{cc^*}{c^*c}\frac{\langle\phi|\psi\rangle\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle} = \frac{\langle\phi|\psi\rangle\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}. \tag{1.5}$$

Moreover, on the continuum the inner product between two wave functions is given by:

$$\langle\phi|\psi\rangle = \int_X \phi^*(x)\psi(x)dx. \tag{1.6}$$

Note that the integral does not vary under a change of $\phi(x)$ or $\psi(x)$ over a set of measure zero. For example, if we changed $\psi(x)$ such that it returned zero over the rationals, the inner product wouldn't change, therefore the physical state does not change either. This is not something peculiar to quantum mechanics, as this happens in classical probability as well. If we have a probability density $\rho(x,p)$, a change over a measure zero set will not change the expectation of any variable.[1] The representation in terms of wave-functions, then, is redundant and we choose those that, for example, are normalized and, to the extent possible, are continuous.

Conversely, $p(\psi|\phi) = 1$ if and only if $\psi = \phi$. Therefore, two different rays are physically distinguishable. This means that

This means that the ray representation is faithful

Theorem 1.7. *Condition QP-RAY provides a*

In physics, one typically assumes that the system is associated with a Hilbert space, and then

But if the choice of a particular wave-function is redundant, and the linearity of the vector space is, ultimately, addition and scalar multiplication over wave-functions, what part of this linearity is physically faithful and what part is physically redundant?

To test whether the linearity of the Hilbert space is a physical constraint or is a choice of representation, we look for a non-linear transformation that leaves the Born rule invariant. Since the Born rule determines the physics, and the Born rule depends on the modulus of the inner product but not its phase, a non-linear transformation that preserves the modulus of the inner product does not change the physics.

Note that, trivially, an anti-linear map does not preserve the inner product, it conjugates it, but they are already widely known to preserve the Born rule. This is, effectively, exchanging the bras with the kets in the representation. Since the addition is not affected, the non-linearity is minor. We are looking for something that breaks linearity in a more significant way.

Consider a two-state system and let $|z^+\rangle$ and $|z^-\rangle$ be two orthonormal state vectors. Every vector in the Hilbert space can be expressed as a linear combination $|\psi\rangle = c_+|z^+\rangle + c_-|z^-\rangle$. Now consider the map $m : \mathcal{H} \to \mathcal{H}$ defined by $m(0) \mapsto 0$ and, on non-zero vectors, by

$$m(c_+|z^+\rangle + c_-|z^-\rangle) \mapsto e^{\imath\theta\frac{|c_+|^2}{|c_+|^2+|c_-|^2}}(c_+|z^+\rangle + c_-|z^-\rangle), \tag{1.8}$$

where $\theta \in (0, 2\pi)$ is an arbitrary constant. Note that $\frac{|c_+|^2}{|c_+|^2+|c_-|^2} = p(z^+|\psi)$ is the probability of measuring z^+ given the state ψ. On the Bloch ball, if we put z^- at the bottom and z^+ at the

[1] Mathematically, the probability density function (PDF) is the Radon-Nikodym derivative of the cumulative distribution function (CDF) with respect to the Lebesgue measure, which is defined up to a measure zero set.

top, the expression gives us the normalized vertical height from the bottom to the top of the sphere as seen in Fig. 1.1. The map is therefore introducing a phase shift that depends on the vertical position. Since it is introducing only a total phase shift, it will map a vector to a physically equivalent vector. The physics will not change, and we do not expect the Born rule to change. However, since the shift is not linear, we will not expect the inner product to be conserved since only linear maps can preserve the inner product.

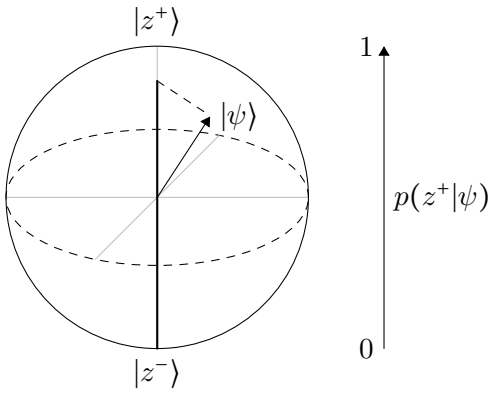


Figure 1.1: Probability is vertical height on the Bloch sphere.

Let's verify all of this in the math. First, m maps a ray to another ray. That is:

$$\begin{aligned} m(k(c_+|z^+\rangle + c_-|z^-\rangle)) &= m(kc_+|z^+\rangle + kc_-|z^-\rangle) = e^{\imath\theta\frac{|kc_+|^2}{|kc_+|^2+|kc_-|^2}}(kc_+|z^+\rangle + kc_-|z^-\rangle) \\ &= ke^{\imath\theta\frac{|k|^2|c_+|^2}{|k|^2(|c_+|^2+|c_-|^2)}}(c_+|z^+\rangle + c_-|z^-\rangle) = ke^{\imath\theta\frac{|c_+|^2}{|c_+|^2+|c_-|^2}}(c_+|z^+\rangle + c_-|z^-\rangle) \\ &= k\,m(c_+|z^+\rangle + c_-|z^-\rangle) \end{aligned} \tag{1.9}$$

In fact, we can see that each ray is mapped to the same ray. We can verify that the map is not linear. In fact:

$$\begin{aligned} m(|z^+\rangle) &= e^{\imath\theta}|z^+\rangle \\ m(|z^-\rangle) &= |z^-\rangle \\ m(|z^+\rangle + |z^-\rangle) &= e^{\frac{\imath\theta}{2}}(|z^+\rangle + |z^-\rangle) \neq m(|z^+\rangle) + m(|z^-\rangle) = e^{\imath\theta}|z^+\rangle + |z^-\rangle. \end{aligned} \tag{1.10}$$

The inner product becomes:

$$\langle m(v)|m(w)\rangle = e^{-\imath\theta\frac{|v_+|^2}{|v_+|^2+|v_-|^2}}e^{\imath\theta\frac{|w_+|^2}{|w_+|^2+|w_-|^2}}\langle v|w\rangle, \tag{1.11}$$

which is manifestly not conserved. However, for the Born rule we have:

$$\begin{aligned}
\langle m(v)|m(v)\rangle &= e^{-\imath\theta\frac{|v_+|^2}{|v_+|^2+|v_-|^2}} e^{\imath\theta\frac{|v_+|^2}{|v_+|^2+|v_-|^2}} \langle v|v\rangle = \langle v|v\rangle \\
p(m(v)|m(w)) &= \frac{\langle m(v)|m(w)\rangle\langle m(w)|m(v)\rangle}{\langle m(v)|m(v)\rangle\langle m(w)|m(w)\rangle} \\
&= \frac{e^{-\imath\theta\frac{|v_+|^2}{|v_+|^2+|v_-|^2}} e^{\imath\theta\frac{|w_+|^2}{|w_+|^2+|w_-|^2}} \langle v|w\rangle e^{-\imath\theta\frac{|w_+|^2}{|w_+|^2+|w_-|^2}} e^{\imath\theta\frac{|v_+|^2}{|v_+|^2+|v_-|^2}} \langle w|v\rangle}{\langle v|v\rangle\langle w|w\rangle} \\
&= \frac{\langle v|w\rangle\langle w|v\rangle}{\langle v|v\rangle\langle w|w\rangle} = p(v|w).
\end{aligned} \tag{1.12}$$

As expected, the Born rule is conserved while the inner product is not.

We should also note that m is not just a bijective map, but also a diffeomorphism. This means that every other relationship in quantum mechanics can be expressed in this non-linear space through m. This includes superposition relationships, expectation of observables and time evolution. These relationships would become considerably more complicated for likely no benefit. Therefore, the fact that we use an underlying linear vector space is, technically, not a physical requirement: it's mathematical convenience.

Theorem 1.13 (Linearity is a choice)**.** *The linearity of the Hilbert space in quantum mechanics is not constrained by the physics. That is, there exist non-linear transformations that preserve the Born rule, making a non-linear representation of quantum states possible.*

However, the physics allows such convenient representation. Why is that the case?

Quantum state representations

A quantum state, then, is not represented by a single vector, but an equivalence class of vectors that give the same results under the Born rule. On the continuum, with wavefunctions $\psi(x)$, there is an additional degeneracy. Mathematically, when we write the vector $|\psi\rangle$ we are actually talking about the set of all functions that are equivalent to the wave function $\psi(x)$.

To recap, a quantum state is represented by a ray, which is an equivalence class of vectors, which is an equivalence class of wavefunctions.

Since every ray is a one-dimensional subspace of the vector space, we can associate each state with the projector $\mathbf{1}_\psi = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}$. Note that the projector is invariant under the choice of representative vector. In fact,

$$\rho_{c\psi} = \frac{|c\psi\rangle\langle c\psi|}{\langle c\psi|\psi c\rangle} = \frac{c|\psi\rangle\langle\psi|c^*}{|c|^2\langle\psi|\psi\rangle} = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle} = \rho_\psi. \tag{1.14}$$

Therefore, given any representative $|\psi\rangle$ of ψ, we get the same projector $\mathbf{1}_\psi$. Conversely, note that given any vector $\phi \in \mathcal{H}$ we have

$$\mathbf{1}_\psi|\phi\rangle = \frac{|\psi\rangle\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle} = c\frac{|\psi\rangle}{\langle\psi|\psi\rangle}. \tag{1.15}$$

That is, the projector applied to any vector returns the component of that vector along ψ. Therefore, the projector applied to the whole vector space returns the one-dimensional

subspace, the ray. This means that

> The state ψ of a quantum system is represented by a one dimensional projector $\mathbf{1}_\psi = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}$ of a complex inner product space $\mathcal{H}$. (QP-PROJ)

is equivalent to **??**.

A one-dimensional projector is also a density operator, as it has only one non-zero eigenvalue which is one. Therefore it is a positive definite Hermitian operator with trace-one. Moreover, projectors have special properties among all the density operators. First of all, they are the only density operators for which $\rho = \rho^2$, since 0 and 1 are the only eigenvalues that, when squared, are equal to themselves. Therefore

> The state of a quantum system is represented by a density operator ρ_ψ such that $\rho_\psi = \rho_\psi^2$ (QP-IDEM)

is equivalent to QP-PROJ.

A way to characterize pure states among ensembles is that pure states cannot be expressed as a non-trivial combination of other ensemble. That is, if ρ_ψ represents a pure state, then the statistical mixture $\rho_\psi = p_1\rho_1 + p_2\rho_2$ implies that either $\rho_1 = \rho_2 = \rho_\psi$ or one of the two addends is equal to ρ_ψ and the corresponding coefficient is one. Mathematically, we say that pure states are the extreme points of the convex set formed by all ensembles. Therefore

> The state of a quantum system is represented by a density operator ρ_ϕ that cannot be expressed as non-trivial sum of other density operators. (QP-EXTR)

is equivalent to QP-PROJ.

Having different equivalent ways to characterize quantum states will be useful as different results are easier to see under different characterization and they help build a broader intuition.

Quantum ensembles representations

Let us know proceed to statistical ensemble with the following

> A statistical ensemble for a quantum system is represented by a density operator ρ, a positive semi-definite, self-adjoint operator of trace one. A statistical mixture is represented by a convex combination $p_1\rho_1 + p_2\rho_2$ of density operators. (QE-DENS)

As we saw before, we can characterize pure states as those ensembles that cannot be expressed as a non-trivial statistical mixture. Therefore condition QE-DENS implies QP-EXTR. The converse, however, is not true.

Suppose we assume **??**. For a two-dimensional quantum system, the state space is the Bloch sphere. A classical directional DOF is also represented by a two-sphere, with the space of ensemble given by the space of probability measures over the sphere. What is different in quantum mechanics is that the same ensemble can be expressed as different mixtures in terms of pure states. For example, the maximally mixed state can be recovered by a uniform distribution over the whole sphere or by a uniform distribution over any two opposite points. Therefore, knowing the state space is not enough to know which statistical mixtures actually represent the same ensemble. Therefore, **??** and QP-EXTR do not imply QE-DENS.

Since condition QE-DENS is stronger than **??**, we should understand what is the mathematical structure implied by the first that is not necessary on the second. Note that, to construct the projective space, we do not need the inner product structure. The vector space structure is sufficient. Therefore, consider the map $m : \mathcal{H} \to \mathcal{H}$ on the Bloch sphere defined by

$$\begin{aligned} m(c^+|z^+\rangle + c^-|z^-\rangle) &= c^+|z^+\rangle + c^-|x^+\rangle \\ m &= |z^+\rangle\langle z^+| + |x^+\rangle\langle z^-| = |z^+\rangle\langle z^+| + \left(\frac{\sqrt{2}}{2}|z^+\rangle + \frac{\sqrt{2}}{2}|z^-\rangle\right)\langle z^-| \\ &= |z^+\rangle\langle z^+| + \frac{\sqrt{2}}{2}|z^+\rangle\langle z^-| + \frac{\sqrt{2}}{2}|z^-\rangle\langle z^-| \\ m(|\psi\rangle) &= M|\psi\rangle = \begin{bmatrix} 1 & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} \end{bmatrix}|\psi\rangle \end{aligned} \tag{1.16}$$

This is a linear map, and therefore it maps rays to rays. It will also map projectors to projectors and therefore will preserve **??** and all its equivalent conditions.

However, it does not map ensembles to ensemble, density operators to density operators, as the trace is not preserved.

$$\begin{aligned} m(\rho) &= M\rho M^\dagger \\ M^\dagger M &= \begin{bmatrix} 1 & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}\begin{bmatrix} 1 & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} 1 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 1 \end{bmatrix} \\ \rho_{x^+} M^\dagger M &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 1 \end{bmatrix} = \begin{bmatrix} \frac{1+\sqrt{2}}{2} & \frac{1+\sqrt{2}}{2} \\ \frac{1+\sqrt{2}}{2} & \frac{1+\sqrt{2}}{2} \end{bmatrix} \\ \mathrm{tr}[m(\rho_{x^+})] &= \mathrm{tr}[M\rho_{x^+}M^\dagger] = \mathrm{tr}[\rho_{x^+}M^\dagger M] = 1 + \sqrt{2} \end{aligned} \tag{1.17}$$

If m were unitary, instead, we would have $MM^\dagger = I$ and therefore the trace would be preserved. Therefore, unitary maps, and therefore preservation of the inner product, is sufficient to map density operator to density operators and preserving their convex structure (i.e. the output of the mixture is the mixture of the outputs). Therefore, not all maps that map rays to rays map ensembles to ensembles.

The converse is also not true: mapping ensembles to ensembles does not necessarily mean mapping pure states to pure states. If we have an open quantum system, for example, it may increase entropy and therefore map pure states to mixed states. We should, therefore, find out which maps map ensembles to ensembles and also distinct pure states to distinct pure states.

On the Bloch ball, such a map would be an affine map (i.e. $m(p_1\rho_1 + p_2\rho_2) = p_1 m(\rho_1) + p_2 m(\rho_2)$) that maps the sphere to itself. An affine map can be expressed as a linear map together with a translation. Translations would necessarily move points out of the ball. Linear maps are restricted to rotations, stretch/shrink and reflections. Stretch/shrink operations would also move points out of the ball. Therefore the only transformation possible are rotations and reflections.

While we have fixed the transformations on the projective space, this is not enough to fix the transformations in the vector space. In fact, since each ray is an equivalence class of vector, there are multiple way to associate elements within the same equivalence classes. However,

since QE-DENS requires addition to correspond to statistical mixing,the map on the vector space will have to be linear. Rotations on the projective space can always be represented as unitary operations in the vector space, since the square of the inner product represents the cosine of an angle in projective space. Reflections on the projective space can be represented as anti-unitary operations, which are essentially switching the bras with kets. While these are usually not considered, in principle they are still valid changes of representation.

The conclusion is that **??** does not require the inner product while QE-DENS does. The inner product structure has more to do with statistical mixing of ensembles than superposition of pure states.

Two linearities: superposition and statistical mixing

A linear combination, a superposition, $\sum c_i|\psi_i\rangle$ will give us another state, but is somewhat unclear what this operation means. First of all, we have the problem that the superposition is physically unique up to a total phase, so only the phase differences are physically significant. Moreover, the meaning of the norm of the coefficient c_i is also ill defined. The overall intuition one gets is that $|c_i|^2$ corresponds to probability, but this does not actually work. In fact, we have

$$\begin{aligned}
|\phi\rangle &= c_1|\psi_1\rangle + c_2|\psi_2\rangle \\
\langle\phi|\phi\rangle &= |c_1|^2\langle\psi_1|\psi_1\rangle + c_1^*c_2\langle\psi_1|\psi_2\rangle + c_2^*c_1\langle\psi_2|\psi_1\rangle + |c_2|^2\langle\psi_2|\psi_2\rangle \\
\langle\psi_1|\phi\rangle &= c_1\langle\psi_1|\psi_1\rangle + c_2\langle\psi_1|\psi_2\rangle \\
p(\phi|\psi_1) &= \frac{\langle\phi|\psi_1\rangle\langle\psi_1|\phi\rangle}{\langle\psi_1|\psi_1\rangle\langle\phi|\phi\rangle} \\
&= \frac{(c_1^*\langle\psi_1|\psi_1\rangle + c_2^*\langle\psi_2|\psi_1\rangle)(c_1\langle\psi_1|\psi_1\rangle + c_2\langle\psi_1|\psi_2\rangle)}{\langle\psi_1|\psi_1\rangle(|c_1|^2\langle\psi_1|\psi_1\rangle + c_1^*c_2\langle\psi_1|\psi_2\rangle + c_2^*c_1\langle\psi_2|\psi_1\rangle + |c_2|^2\langle\psi_2|\psi_2\rangle)}.
\end{aligned} \tag{1.18}$$

In the special case where ψ_1 and ψ_2 are orthonormal, the above simplifies to

$$p(\phi|\psi_1) = \frac{|c_1|^2}{|c_1|^2 + |c_2|^2} \tag{1.19}$$

and therefore, if $|c_1|^2 + |c_2|^2 = 1$, we recover the probability. But this is a special case which does not work in general. And building our understanding on special cases that do not work in general is not a recipe for success.

We want to stress that superpositions play no fundamental role in classifying states: any state can be expressed as a linear combination of any other state. In a spin 1/2 system, for example, spin left is a linear combination of spin up and spin down with $|c_1|^2 = |c_2|^2 = \frac{1}{2}$. But spin up is also a linear combination of spin left and spin right with $|c_1|^2 = |c_2|^2 = \frac{1}{2}$. Any spin 1/2 state, in fact, can always be expressed as a linear combination of any two distinct states, not necessarily orthogonal. The reason is that any two linearly independent vectors of a two dimensional space will span the entire space.

We also want to stress that the linearity of superpositions is different from the linearity of statistical mixture. The first linearity acts on pure states only, uses complex coefficient, returns a pure state and is specific to quantum mechanics. The second linearity acts on all states, mixed or pure, uses real coefficient, returns mixed states and is present in both classical and quantum mechanics. In fact, it must be present in any physical theory, since

all measurements are statistical in nature. If we pick two pure states ψ and ϕ on the Bloch sphere for a spin 1/2 system, they will be represented by two points on the surface. The set of all superposition $c_\psi|\psi\rangle + c_\phi|\phi\rangle$ is the whole surface. The set of all statistical mixtures $p|\psi\rangle\langle\psi| + (1-p)|\phi\rangle\langle\phi|$, instead, is the line segment that contains the two points. Therefore, it is not true that superpositions are "a kind of" statistical mixture.

The two linearities, however, are related and understanding their relationship is key to understand why superpositions are possible in quantum mechanics but not in classical mechanics. Mathematically, if $|\phi\rangle$ is a superposition of $|\psi_1\rangle$ and $|\psi_2\rangle$, then we can find a statistical mixture ρ of $|\psi_1\rangle\langle\psi_1|$ and $|\psi_2\rangle\langle\psi_2|$ such that it can also be expressed as a mixture of $|\phi\rangle\langle\phi|$ and another state $|\hat{\phi}\rangle\langle\hat{\phi}|$. Let $|\phi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle$. We define $|\hat{\phi}\rangle = c_1|\psi_1\rangle - c_2|\psi_2\rangle$. We have:

$$\begin{aligned}
|\phi\rangle\langle\phi| &= c_1^* c_1|\psi_1\rangle\langle\psi_1| + c_1^* c_2|\psi_1\rangle\langle\psi_2| + c_2^* c_1|\psi_2\rangle\langle\psi_1| + c_2^* c_2|\psi_2\rangle\langle\psi_2| \\
|\hat{\phi}\rangle\langle\hat{\phi}| &= c_1^* c_1|\psi_1\rangle\langle\psi_1| - c_1^* c_2|\psi_1\rangle\langle\psi_2| - c_2^* c_1|\psi_2\rangle\langle\psi_1| + c_2^* c_2|\psi_2\rangle\langle\psi_2| \\
|\phi\rangle\langle\phi| + |\hat{\phi}\rangle\langle\hat{\phi}| &= 2|c_1|^2|\psi_1\rangle\langle\psi_1| + 2|c_2|^2|\psi_2\rangle\langle\psi_2|
\end{aligned} \tag{1.20}$$

$$\begin{aligned}
p_\phi &= \frac{\langle\phi|\phi\rangle}{\langle\phi|\phi\rangle + \langle\hat{\phi}|\hat{\phi}\rangle} \qquad p_{\hat{\phi}} = \frac{\langle\hat{\phi}|\hat{\phi}\rangle}{\langle\phi|\phi\rangle + \langle\hat{\phi}|\hat{\phi}\rangle} \\
p_1 &= \frac{2|c_1|^2\langle\psi_1|\psi_1\rangle|}{\langle\phi|\phi\rangle + \langle\hat{\phi}|\hat{\phi}\rangle} \quad p_2 = \frac{2|c_2|^2\langle\psi_2|\psi_2\rangle|}{\langle\phi|\phi\rangle + \langle\hat{\phi}|\hat{\phi}\rangle}
\end{aligned} \tag{1.21}$$

$$\begin{aligned}
\rho = p_\phi\rho_\phi + p_{\hat{\phi}}\rho_{\hat{\phi}} &= \frac{\langle\phi|\phi\rangle}{\langle\phi|\phi\rangle + \langle\hat{\phi}|\hat{\phi}\rangle}\frac{|\phi\rangle\langle\phi|}{\langle\phi|\phi\rangle} + \frac{\langle\hat{\phi}|\hat{\phi}\rangle}{\langle\phi|\phi\rangle + \langle\hat{\phi}|\hat{\phi}\rangle}\frac{|\hat{\phi}\rangle\langle\hat{\phi}|}{\langle\hat{\phi}|\hat{\phi}\rangle} \\
&= \frac{1}{\langle\phi|\phi\rangle + \langle\hat{\phi}|\hat{\phi}\rangle}\left(|\phi\rangle\langle\phi| + |\hat{\phi}\rangle\langle\hat{\phi}|\right) \\
&= \frac{1}{\langle\phi|\phi\rangle + \langle\hat{\phi}|\hat{\phi}\rangle}\left(2|c_1|^2|\psi_1\rangle\langle\psi_1| + 2|c_2|^2|\psi_2\rangle\langle\psi_2|\right) \\
&= \frac{2|c_1|^2\langle\psi_1|\psi_1\rangle|}{\langle\phi|\phi\rangle + \langle\hat{\phi}|\hat{\phi}\rangle}\frac{|\psi_1\rangle\langle\psi_1|}{\langle\psi_1|\psi_1\rangle} + \frac{2|c_2|^2\langle\psi_2|\psi_2\rangle|}{\langle\phi|\phi\rangle + \langle\hat{\phi}|\hat{\phi}\rangle}\frac{|\psi_2\rangle\langle\psi_2|}{\langle\psi_2|\psi_2\rangle} \\
&= p_1\rho_{\psi_1} + p_2\rho_{\psi_2}
\end{aligned} \tag{1.22}$$

Note that $p_\phi + p_{\hat{\phi}} = 1$, which means ρ is a statistical mixture with the given probability. This also means that $p_1 + p_2 = 1$ as well. We therefore have a mixed state that can be expressed either as a mixture of ψ_1 and ψ_2 or of ϕ and $\hat{\phi}$.

The converse is also true. Suppose that $\rho = p_\phi\rho_\phi + p_{\hat{\phi}}\rho_{\hat{\phi}} = p_1\rho_{\psi_1} + p_2\rho_{\psi_2}$. Since ρ is the mixture of two pure states, it is supported by a two dimensional subspace. Since ρ is a mixture of ρ_{ψ_1} and ρ_{ψ_2}, the subspace is the span of ψ_1 and ψ_2. Since ϕ must be in that subspace as well, it must be a superposition of ψ_1 and ψ_2. This means that ϕ is a superposition of ψ_1 and ψ_2 if and only if there is mixed state ρ that can be expressed as a mixture of both ψ_1 and ψ_2 as well as a mixture of ϕ and some other state. This lead to the following:

Insight 1.23. *The existence of superpositions of pure states in quantum mechanics is equivalent to the existence of multiple decomposition of statistical mixtures in terms of pure states.*

This links the abstract operation of quantum superposition to the more concrete physical property of statistical mixing.

This explains why there are no superpositions in classical mechanics. Every classical ensemble has a unique decomposition in terms of pure states, where the pure states are limits of distributions all concentrated at a point in phase space. Since classical mechanics does not allow multiple decompositions, it does not allow superpositions either.

Note that the previous insight does not tell us that if a theory allows multiple decompositions then it must be equivalent to quantum mechanics. For example, if the state space of statistical mixture is a disc instead of a ball, we would still have multiple decompositions but no complex superpositions. It does tell us, however, that what superpositions are allowed is fixed by what mixtures are allowed, which means that the pure states are rays of a complex Hilbert space if and only if the space of statistical mixtures is isomorphic to the quantum one.

The above insight also tells us why quantum mechanics is a linear theory. If we have a physical process $\mathcal{P}$ that takes as input a statistical mixture, we expect that the output of the statistical mixture is the statistical mixture of the individual outputs. That is, $\mathcal{P}(p\rho_1 + (1-p)\rho_2) = p\mathcal{P}(\rho_1) + (1-p)\mathcal{P}(\rho_2)$. But if the process is linear with respect to statistical mixing, and the linearity of statistical mixing can be converted to the linearity of superpositions, then the process must be also linear with respect to superpositions. Similarly, if we have an observable O and a statistical mixture, we expect that the expectation of the statistical mixture is the average of the expectations on the statistical components. That is $E[p\rho_1 + (1-p)\rho_2] = pE[\rho_1] + (1-p)E[\rho_2]$. In quantum mechanics, this means that $\operatorname{tr}(O(p\rho_1 + (1-p)\rho_2)) = p\operatorname{tr}(O\rho_1) + (1-p)\operatorname{tr}(O\rho_2)$. That is, the linearity of the observables is exactly the linearity of expectation values.

The gap between the two linearities

Note, however, that while the linearity of statistical mixture is a physically necessary requirement, the linearity of superposition is a choice dictated by mathematical convenience. As we saw, it is the Born rule that captures the physics, not the inner product. Therefore, a non-linear transformation that preserves the inner product is still physically valid. For example, consider the Bloch ball. Taken two orthogonal states ψ and ϕ, every other state can be expressed as a linear combination $c_\psi|\psi\rangle + c_\phi|\phi\rangle$. Now consider the map $m : \mathcal{H} \to \mathcal{H}$ defined by

$$m(c_\psi|\psi\rangle + c_\phi|\phi\rangle) \mapsto e^{\imath\theta \frac{|c_\phi|}{\sqrt{|c_\psi|^2+|c_\phi|^2}}}(c_\psi|\psi\rangle + c_\phi|\phi\rangle), \tag{1.24}$$

where $\theta \in (0, 2\pi)$ is an arbitrary constant. The map is introducing a phase shift that is proportional to the cosine of the angle along the vertical direction. Since it is introducing only a total phase shift, it will map a vector to a physically equivalent vector. The physics will not change, and we do not expect the Born rule to change. However, since the shift is not linear, we will not expect the inner product to me conserved since only linear maps can preserve the inner product.

Let's verify all of this in the math. First, m maps a ray to another ray. That is:

$$\begin{aligned} m(k(c_\psi|\psi\rangle + c_\phi|\phi\rangle)) &= m(kc_\psi|\psi\rangle + kc_\phi|\phi\rangle) = e^{\imath\theta\frac{|kc_\phi|}{\sqrt{|kc_\psi|^2+|kc_\phi|^2}}}(kc_\psi|\psi\rangle + kc_\phi|\phi\rangle) \\ &= ke^{\imath\theta\frac{|k||c_\phi|}{|k|\sqrt{|c_\psi|^2+|c_\phi|^2}}}(c_\psi|\psi\rangle + c_\phi|\phi\rangle) = ke^{\imath\theta\frac{|c_\phi|}{\sqrt{|c_\psi|^2+|c_\phi|^2}}}\left(c_\psi|\psi\rangle + c_\phi|\phi\rangle\right) \\ &= k\, m(c_\psi|\psi\rangle + c_\phi|\phi\rangle) \end{aligned} \tag{1.25}$$

Moreover, we can see that each ray is mapped to the same ray. We can verify that the map in not linear. In fact:

$$\begin{aligned} m(|\psi\rangle) &= |\psi\rangle \\ m(|\phi\rangle) &= e^{\imath\theta}|\phi\rangle \\ m(|\psi\rangle + |\phi\rangle) &= e^{\frac{\imath\theta}{\sqrt{2}}}(|\psi\rangle + |\phi\rangle) \neq |\psi\rangle + e^{\imath\theta}|\phi\rangle. \end{aligned} \tag{1.26}$$

The inner product becomes:

$$\langle m(v)|m(w)\rangle = e^{-\imath\theta\frac{|v_\phi|}{\sqrt{|v_\psi|^2+|v_\phi|^2}}} e^{\imath\theta\frac{|w_\phi|}{\sqrt{|w_\psi|^2+|w_\phi|^2}}}\langle v|w\rangle, \tag{1.27}$$

which is manifestly not conserved. However, for the Born rule we have:

$$\begin{aligned} \langle m(v)|m(v)\rangle &= e^{-\imath\theta\frac{|v_\phi|}{\sqrt{|v_\psi|^2+|v_\phi|^2}}} e^{\imath\theta\frac{|v_\phi|}{\sqrt{|v\psi|^2+|v_\phi|^2}}}\langle v|v\rangle = \langle v|v\rangle \\ p(m(v)|m(w)) &= \frac{\langle m(v)|m(w)\rangle\langle m(w)|m(v)\rangle}{\langle m(v)|m(v)\rangle\langle m(w)|m(w)\rangle} \\ &= \frac{e^{-\imath\theta\frac{|v_\phi|}{\sqrt{|v_\psi|^2+|v_\phi|^2}}} e^{\imath\theta\frac{|w_\phi|}{\sqrt{|w_\psi|^2+|w_\phi|^2}}}\langle v|w\rangle e^{-\imath\theta\frac{|w_\phi|}{\sqrt{|w_\psi|^2+|w_\phi|^2}}} e^{\imath\theta\frac{|v_\phi|}{\sqrt{|v_\psi|^2+|v_\phi|^2}}}\langle w|v\rangle}{\langle v|v\rangle\langle w|w\rangle} \\ &= \frac{\langle v|w\rangle\langle w|v\rangle}{\langle v|v\rangle\langle w|w\rangle} = p(v|w). \end{aligned} \tag{1.28}$$

As expected, the Born rule is conserved while the inner product is not.

This tells us that, technically, the linearity of the underlying vector space is not a physical requirement and, in principle, we could represent pure states with a space where superposition as "slightly" non-linear. That is, the fact that the space of pure states is a complex projective space is a physical requirement if we want to represent a set of states that provides the statistical mixtures given by quantum mechanics. However, using the underlying complex inner product space is not a physical requirement: it is mathematical convenience.

Insight 1.29. *The linearity of statistical mixing is a physical requirement. The linearity of the Hilbert space is a mathematical convenience.*

Basis and gauge choice

Once we decide to use the vector space, the representation of states in terms of vector is also defined up to an arbitrary choice. In a spin 1/2 system, for example, we typically define $|x^+\rangle = \frac{1}{\sqrt{2}}|z^+\rangle + \frac{1}{\sqrt{2}}|z^-\rangle$. But note the state corresponding to spin down is the whole ray, so

the above expression hides that fact that $|z^-\rangle$ hides a choice of arbitrary phase. We could, in fact, redefine $|z^-\rangle$ with a quarter of a turn phase shift, and have $|x^+\rangle = \frac{1}{\sqrt{2}}|z^+\rangle + \imath\frac{1}{\sqrt{2}}|z^-\rangle$. In other words, the choice of representation of the physics in terms of the Hilbert space is up to an arbitrary choice of phase for elements of the basis.

For continuous observable, this choice is even more important, as it is not a matter of basis, but of operators. Suppose we have a wave function $\psi(x)$. The corresponding position and momentum operators will be $X = x$ and $P = -\imath\hbar\partial_x$. Now choose an arbitrary continuous function $\theta(x)$. This corresponds to an arbitrary choice of phase at each point. If we set $\phi(x) = e^{\imath\theta(x)}\psi(x)$, the probability distribution over position will not change as the norm at each point remains the same. If set $P_\phi = -\imath\hbar\partial_x - \hbar\partial_x\theta(x)$, we have

$$\{X, P_\phi\} = \{X, P - \hbar\partial_x\theta(x)\} = \{X, P\} - \{X, \hbar\partial_x\theta(x)\} = \{X, P\} \tag{1.30}$$

since X will commute with any function of x. This means that X and P_θ have the same commutation relationships than X and P. Also note that

$$\begin{aligned}
\langle\phi|P_\phi|\phi\rangle &= \int_X e^{-\imath\theta}\psi(x)(-\imath\hbar\partial_x - \hbar\partial_x\theta(x))(e^{\imath\theta}\psi(x))dx \\
&= \int_X e^{-\imath\theta}\psi(x)(-\imath\hbar e^{\imath\theta}\imath\partial_x\theta(x)\psi(x) - e^{\imath\theta}\imath\hbar\partial_x\psi(x) - \hbar\partial_x\theta(x)e^{\imath\theta}\psi(x))dx \\
&= \int_X e^{-\imath\theta}\psi(x)(-e^{\imath\theta}\imath\hbar\partial_x\psi(x))dx \\
&= \int_X \psi(x)(-\imath\hbar\partial_x)\psi(x)dx = \langle\psi|P_\psi|\psi\rangle.
\end{aligned} \tag{1.31}$$

In other words, we can redefine both the phase of the wave function and the momentum operator so that the physics remains unchanged. This corresponds to the change of gauge we had in classical particle mechanics, where we could redefine momentum $p \to p + \partial f(x)$ while leaving the Poisson brackets, and therefore the space-time trajectories, unchanged.

Insight 1.32. *The choice of a basis in the Hilbert space is a choice of a maximal set of orthogonal states plus a choice of gauge (i.e. a phase for each element of the basis).*

To sum up, fixing the Hilbert space and its physical representation is equivalent to fixing the space of ensembles, including how they mix, requiring a linear representation in terms of the Hilbert space and choosing an arbitrary phase for each element of the spectrum of a complete set of observables.

Probability from expectations

Another interesting equivalence is the one between probability, subspaces and expectations. This equivalence exists also in classical probability, therefore we should understand it there fit. The standard way to define a probability space is by giving three objects. First, a sample space Ω that represents all the possible cases. In classical mechanics, it's phase space, where each point represent all the values of position and momentum. Second, an algebra of events Σ_Ω which represents all possible conditions we could test for. In classical mechanics, this is the collection of Borel sets, subsets of phase space that can be generated from the open regions using complements and countable unions. Third, a probability measure $p : \Sigma_\Omega \to [0, 1]$ that associated a probability to each event. From this one can define random variables $X : \Omega \to \mathbb{R}$

and their expectation. That is, we start with a notion of probability and define a notion of observables on top.

We could proceed the other way. The key insight is that for every event $A \in \Sigma_\Omega$ we can define the indicator function $\mathbf{1}_A : \Omega \to \{0, 1\}$ that returns one if the given point is in A or zero otherwise. This is a random variable and its expectation is exactly the probability of the event A. That is, $E[\mathbf{1}_A] = p(A)$. Suppose, therefore, that we are given Ω, Σ_Ω but not the probability measure. However, we are given the expectation operator $E : M(\Omega, \mathbb{R})$ that, given a function $f : \Omega \to \mathbb{R}$[2], $E[f]$ returns it's expectation. We can then define the probability measure as $p(A) = E[\mathbf{1}_A]$.

A similar perspective can be achieved in quantum mechanics, and it is the basis of some other axiomatic approaches and quantum reconstruction approaches, like quantum logic. As we are using the same notation for indicator functions and projectors, it should not be surprising that projectors play the role of indicator function in quantum mechanics. A projector P is an idempotent linear operator, meaning that $P^2 = P$. This also means that $E[P^2] = E[P]$. This is true for an indicator function I as well: since the only values I takes are zeros and one, $I^2 = I$, and $E[I^2] = E[I]$. On the other hand, the spectrum of a projector (i.e. the possible eigenvalues) consists only of zero and one, just like an indicator function. Therefore, if we can reverse the definitions for probability in quantum mechanics as well.

The standard construction starts with a set of pure states $\mathbf{P}(\mathcal{H})$, defines the Born rule $p(\phi|\psi)$ in terms of the inner product, and then the expectation based on the Born rule. We can start with a set of pure states $\mathbf{P}(\mathcal{H})$, where $\mathcal{H}$ is taken as a vector space forgetting its inner product, the space of all possible observables $O(\mathcal{H})$ and the expectation operator $E : O(\mathcal{H}) \to \mathbb{R}$. For every pure state $\phi \in \mathbf{P}(\mathcal{H})$ we can find the corresponding projector $\mathbf{1}_\phi$ and define $p(\phi) = E[\mathbf{1}_\phi]$. This allows to reconstruct the mixed state, much like we reconstructed the probability measure in classical mechanics. To reconstruct the Born rule, for each pure $\psi \in \mathbf{P}(\mathcal{H})$ we need an E_ψ that gives us the expectations of all the variables if we prepared ψ. Then we can set $p(\phi|\psi) = E_\psi[\mathbf{1}_\phi]$.

Therefore condition

Expectations are defined for all observables and for all ensembles (BR-EXP)

is an alternative characterization of the Born rule.

Born rule and angles in projective space

As we saw, the Born rule is essentially a relationship between rays, points in the projective space. We want to understand what exactly is it describing. Since it is a relationship between two pure states, let us take two rays $\psi, \phi \in P(\mathcal{H})$. Regardless of the dimensionality of $\mathcal{H}$, ψ and ϕ will span a two-dimensional subspace, a Bloch sphere. In fact, considering the circle that contains ψ, ϕ and its opposite $\phi_\perp$ will suffice.

Since ϕ and $\phi_\perp$ are opposite in the Bloch sphere, and therefore orthogonal, we can write

[2]The function will need to be measureable in the mathematical sense: the inverse image of a Borel set is a Borel set.

$|\psi\rangle = \rho|\phi\rangle + \sqrt{1-\rho^2}e^{i\theta}|\phi_\perp\rangle$. We have

$$\begin{aligned} p(\psi|\phi) &= \frac{\langle\phi|\psi\rangle\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle} = \rho^2 = \cos^2\theta_{\psi\phi} \\ \theta_{\psi\phi} &= \arccos\sqrt{\frac{\langle\phi|\psi\rangle\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}}. \end{aligned} \tag{1.33}$$

In other words, the Born rule defines the geometry of the projective space by defining the angles between points. Note that, since we are on a unit sphere, the angle can be understood also as a distance along the maximal circle that connects the two. The maximal circles are the geodesics of the sphere and therefore the angle is the distance between the points. Therefore $\theta_{\psi\phi}$ is literally the distance function that define the geometry of the space.[3] We have that

The angles between pure states are well defined. (BR-ANG)

is an alternative characterization of the Born rule.

Entropy and the Born rule

TODO: preamble.

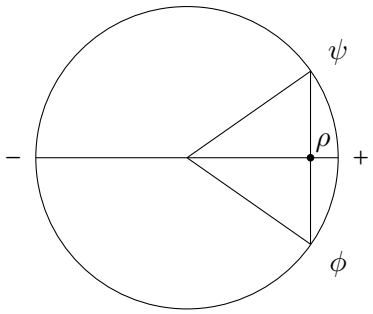


Figure 1.2: Geometric diagonalization of an equal mixture of two states.

To calculate the entropy of a mixture of two pure states, we can use the understanding that we built in terms of ensembles and their multiple decomposition. Let ψ and ϕ be two pure states. Their equal mixture $\rho = \frac{1}{2}\rho_\psi + \frac{1}{2}\rho_\phi$ will be the midpoint between them. To calculate the entropy, we need to diagonalize ρ and calculate the Shannon entropy from the eigenvalues. To diagonalize ρ we need to express it as a mixture of two orthogonal + and −. But + and − will be orthogonal only if they are opposite points on the Bloch sphere. Therefore, ρ is diagonalized by + and − if and only if it lies on the axis identified by + and − as we see in figure 1.2. Geometrical, we will have $\rho = \frac{\overline{-\rho}}{\overline{-+}}\rho_+ + \frac{\overline{\rho+}}{\overline{-+}}\rho_-$. Using simple trigonometry and recalling the relationship between the Born rule and angles, we have:

$$\begin{aligned} \cos\theta_{\psi+} &= \cos\frac{\theta_{\psi\phi}}{2} = \sqrt{p(\phi|\psi)} \\ \rho &= \frac{\overline{-\rho}}{\overline{-+}}\rho_+ + \frac{\overline{\rho+}}{\overline{-+}}\rho_- = \frac{1+\cos\theta_{\psi+}}{2}\rho_+ + \frac{1-\cos\theta_{\psi+}}{2}\rho_- \\ &= \frac{1+\sqrt{p(\phi|\psi)}}{2}\rho_+ + \frac{1-\sqrt{p(\phi|\psi)}}{2}\rho_-. \end{aligned} \tag{1.34}$$

[3] The angle is the distance as calculated using the Fubini-Study metric.

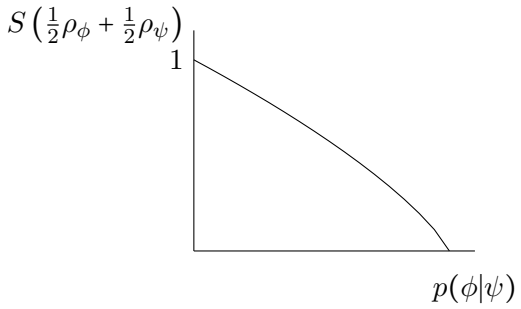


Figure 1.3: Entropy of a mixture of two pure states as a function of their Born rule.

The entropy is therefore

$$S(\rho) = -\operatorname{tr}(\rho \log \rho) = -\frac{1+\sqrt{p(\phi|\psi)}}{2}\log\frac{1+\sqrt{p(\phi|\psi)}}{2} - \frac{1-\sqrt{p(\phi|\psi)}}{2}\log\frac{1-\sqrt{p(\phi|\psi)}}{2}. \quad (1.35)$$

The relationship is plotted in figure 1.3 and, as one can see, it is strictly concave. The relationship is, therefore, invertible: if we are given the entropy of all pairs of state we would be able to recover the Born rule. In other words

The entropy is well defined for all ensembles (BR-ENT)

is an alternative characterization of the Born rule.

States are rays ¡-¿ states are pure ensembles.
Probability as expectations of indicator functions
Projections as indicator functions
Orthogonality as mutual exclusivity
Inner product as overlap between pure ensembles

1.3 Schroedinger equation and unitary evolution

We start with the Schroedinger equation

The evolution follows the equation $\imath\hbar\frac{d}{dt}\psi(t) = H\psi(t)$ where H is a self-adjoint operator (DR-SCEQ)

Time evolution is unitarity

The Schroedinger equation describes time evolution as a relationship between the change of the state and the Hamiltonian of the system. Here we want to characterize time evolution as a map U_{dt} that takes the state at time t and returns the state at time $t + dt$. We have

$$\begin{aligned}\psi(t+dt) &= \psi(t) + d\psi(t) = \psi(t) + \frac{d}{dt}\psi(t)dt = \psi(t) + \frac{Hdt}{\imath\hbar}\psi(t)\\ &= \left(1 + \frac{Hdt}{\imath\hbar}\right)\psi(t) = U_{dt}\psi(t).\end{aligned} \quad (1.36)$$

Since H is a self-adjoint operator, we have

$$\begin{aligned} U_{dt}^\dagger U_{dt} &= \left(1+\frac{H dt}{\imath\hbar}\right)^\dagger \left(1+\frac{H dt}{\imath\hbar}\right) = \left(1+\frac{H^\dagger dt}{(-\imath)\hbar}\right)\left(1+\frac{H dt}{\imath\hbar}\right) = \left(1-\frac{H dt}{\imath\hbar}\right)\left(1+\frac{H dt}{\imath\hbar}\right) \\ &= 1+\left(\frac{H dt}{\imath\hbar}\right)^2 = 1+O(dt^2). \end{aligned} \tag{1.37}$$

This means that the infinitesimal time evolution operator U_{dt} is unitary.

We can proceed in the opposite way. Given an infinitesimal unitary time evolution operator, we can recover the change of the state in time.

$$\begin{aligned} \frac{d\psi(t)}{dt} &= \frac{\psi(t+dt)-\psi(t)}{dt} = \frac{U_{dt}\psi(t)-\psi(t)}{dt} = \frac{U_{dt}-1}{dt}\psi(t) = A\psi(t) \\ U_{dt} &= 1+Adt \end{aligned} \tag{1.38}$$

From the unitarity we have

$$\begin{aligned} U_{dt}^\dagger U_{dt} &= (1+Adt)^\dagger (1+Adt) = \left(1+A^\dagger dt\right)(1+Adt) = 1+\left(A+A^\dagger\right)dt + A^\dagger A dt^2 \\ &= 1+\left(A+A^\dagger\right)dt + O(dt^2) = 1 \\ A &= -A^\dagger \end{aligned} \tag{1.39}$$

Therefore A is skew-adjoint, and we can set $A = \frac{H}{\imath\hbar}$ where H is a self-adjoint operator. Therefore $\imath\hbar\frac{d}{dt}\psi(t) = \imath\hbar A\psi(t) = H\psi(t)$. The evolution follows the Schroedinger equation. To sum it up

Time evolution is unitary: $U_{dt}^\dagger U_{dt} = U_{dt}U_{dt}^\dagger = 1$ (DR-UNIT)

is equivalent to DR-SCEQ.

TODO: make sure to discuss the invertibility condition!

Time evolution is preservation of the inner product

Note that if the evolution is unitary, then the inner product is conserved. In fact

$$\langle U_{dt}\phi | U_{dt}\psi\rangle = \langle \phi | U_{dt}^\dagger U_{dt} | \psi\rangle = \langle\phi|\psi\rangle. \tag{1.40}$$

The argument works in reverse as well: an infinitesimal time evolution operator that conserves the inner product is unitary. Therefore

Time evolution preserves the inner product : $\langle U_{dt}\phi | U_{dt}\psi\rangle = \langle\phi|\psi\rangle$ (DR-INN)

is equivalent to DR-UNIT.

Time evolution is preservation of the norm

In particular, unitary evolution will preserve the norm of vectors, since $|\psi|^2 = \langle\psi|\psi\rangle$. Note that the inner product can be expressed, through the polarization identity, in terms of the norm.

$$\langle\phi|\psi\rangle = \frac{1}{4}\left(|\phi+\psi|^2 - |\phi-\psi|^2 - \imath|\phi+\imath\psi|^2 + \imath|\phi-\imath\psi|^2\right) \tag{1.41}$$

If U_{dt} is a linear transformation that preserves the norm, that is $|U_{dt}\psi|^2 = |\psi|^2$, then, we have

$$\begin{aligned}\langle U_{dt}\phi|U_{dt}\psi\rangle &= \frac{1}{4}\left(|U_{dt}\phi + U_{dt}\psi|^2 - |U_{dt}\phi - U_{dt}\psi|^2 - \imath|U_{dt}\phi + \imath U_{dt}\psi|^2 + \imath|U_{dt}\phi - \imath U_{dt}\psi|^2\right)\\ &= \frac{1}{4}\left(|U_{dt}(\phi+\psi)|^2 - |U_{dt}(\phi-\psi)|^2 - \imath|U_{dt}(\phi+\imath\psi)|^2 + \imath|U_{dt}(\phi-\imath\psi)|^2\right)\\ &= \frac{1}{4}\left(|\phi+\psi|^2 - |\phi-\psi|^2 - \imath|\phi+\imath\psi|^2 + \imath|\phi-\imath\psi|^2\right) = \langle\phi|\psi\rangle\end{aligned} \tag{1.42}$$

Therefore condition

> Time evolution is linear and preserves the norm: $\langle\psi(t)|\psi(t)\rangle = \langle\psi(t+dt)|\psi(t+dt)\rangle$ (DR-NORM)

is equivalent to condition DR-INN.

Time evolution is preservation of the Born rule

We now look at the square of the inner product between two states infinitesimally close in time. We have

$$\begin{aligned}|\langle\psi(t)|\psi(t+dt)\rangle|^2 &= \langle\psi(t)|U_{dt}\psi(t)\rangle\langle U_{dt}\psi(t)|\psi(t)\rangle\\ &= \langle\psi(t)|1 + \frac{Hdt}{\imath\hbar}|\psi(t)\rangle\langle\psi(t)|\left(1+\frac{Hdt}{\imath\hbar}\right)^\dagger|\psi(t)\rangle\\ &= \langle\psi(t)|1 + \frac{Hdt}{\imath\hbar}|\psi(t)\rangle\langle\psi(t)|1 - \frac{Hdt}{\imath\hbar}|\psi(t)\rangle\\ &= \langle\psi(t)|\psi(t)\rangle\langle\psi(t)|\psi(t)\rangle + \langle\psi(t)|\psi(t)\rangle\langle\psi(t)| - \frac{Hdt}{\imath\hbar}|\psi(t)\rangle\\ &+ \langle\psi(t)|\frac{Hdt}{\imath\hbar}|\psi(t)\rangle\langle\psi(t)|\psi(t)\rangle + O(dt^2)\\ &= |\langle\psi(t)|\psi(t)\rangle|^2 + O(dt^2).\end{aligned} \tag{1.43}$$

In particular, if the vector is normalized, we have

$$|\langle\psi(t)|\psi(t+dt)\rangle|^2 = 1 + O(dt^2). \tag{1.44}$$

That is, the square of the inner product between two infinitesimally close states is one.

For the converse, let's assume U_{dt} is such that $|\langle\psi(t)|\psi(t+dt)\rangle|^2 = |\langle\psi(t)|\psi(t)\rangle|^2$. As we saw before, since U_{dt} is infinitesimal, we can write $U_{dt} = 1 + Adt$. We have:

$$\begin{aligned}|\langle\psi(t)|\psi(t)\rangle|^2 &= |\langle\psi(t)|\psi(t+dt)\rangle|^2 = \langle\psi(t)|U_{dt}\psi(t)\rangle\langle U_{dt}\psi(t)|\psi(t)\rangle\\ &= \langle\psi(t)|U_{dt}|\psi(t)\rangle\langle\psi(t)|U_{dt}^\dagger|\psi(t)\rangle\\ &= \langle\psi(t)|1 + Adt|\psi(t)\rangle\langle\psi(t)|1 + A^\dagger dt|\psi(t)\rangle\\ &= |\langle\psi(t)|\psi(t)\rangle|^2 + \langle\psi(t)|\psi(t)\rangle\langle\psi(t)|(A+A^\dagger)dt|\psi(t)\rangle + O(dt^2).\end{aligned} \tag{1.45}$$

This means that $A = A^\dagger$ and therefore $U_{dt} = 1 + Adt$ is unitary. Therefore condition

> The square of the inner product between to states infinitesimally close in time is one: $|\langle\psi(t)|\psi(t+dt)\rangle|^2 = 1$ (DR-UBOR)

is yet another equivalent condition to unitary evolution.

States and states change are perpendicular on the complex plane

The above two conditions make sense if we understand what happens geometrically. A unitary evolution is effectively a rotation in the Hilbert space, that is why the norm is conserved. An infinitesimal unitary evolution, then, is an infinitesimal rotation so the change is tangent to the circle, and therefore perpendicular with respect to the original vector. This is why the projection of the new vector onto the old one is equal to the norm.

However, perpendicular does not mean orthogonal in this case. In the complex plane, a multiplication by $\imath$ gives us a perpendicular vector that is not orthogonal with respect to the inner product. To see this

$$\begin{aligned}\langle\psi(t+dt)|\psi(t+dt)\rangle &= \langle\psi(t)+d\psi|\psi(t)+d\psi\rangle \\ &= \langle\psi(t)|\psi(t)\rangle + \langle\psi(t)|d\psi\rangle + \langle d\psi|\psi(t)\rangle + \langle d\psi|d\psi\rangle \\ &= \langle\psi(t)|\psi(t)\rangle + \langle\psi(t)|Adt|\psi(t)\rangle + \langle\psi(t)|A^\dagger dt|\psi(t)\rangle + O(dt^2) \\ \langle\psi(t+dt)|\psi(t+dt)\rangle - \langle\psi(t)|\psi(t)\rangle &= \langle\psi(t)|(A+A^\dagger)dt|\psi(t)\rangle + O(dt^2)\end{aligned} \tag{1.46}$$

Note how $A+A^\dagger$ gives us the change in the norm of ψ during the evolution. The norm does not change if and and only if A is skew-adjoint. Since $\langle\psi(t)|d\psi(t)\rangle = \langle\psi(t)|A|\psi(t)\rangle \neq 0$, the change in not orthogonal in the Hilber space. However, the quantity is imaginary so the change is perpendicular in the complex plane of every dimension. If we consider the triangle formed by $\psi(t+dt)$, $\psi(t)$ and $d\psi$, the triangle is perpendicular if and only if A is skew-adjoint. Therefore the condition

> The change is perpendicular to the original state: $\langle\psi(t)|d\psi(dt)\rangle$ is imaginary (DR-PERP)

is yet another equivalent condition to DR-SCEQ.

Time evolution is preservation of bases orthonormality

Another way to characterize unitary evolution is through what happens to an orthonormal basis $|e_i\rangle$. Given that a unitary evolution preserves the inner product, we have $\langle U_{dt}e_i|U_{dt}e_j\rangle = \langle e_i|e_j\rangle = \delta_{ij}$. Therefore the unitary evolution maps an orthonormal basis to another orthonormal basis. The converse is also true, if U_{dt} is linear and maps an orthonormal basis to another orthonormal basis, then the inner product is preserves. To see this, we can simply expand any vector in terms of the basis vector. That is

$$\begin{aligned}\langle\psi(t)|\phi(t)\rangle &= \langle c_i e_i|d_j e_j\rangle = c_i^* d_j\langle e_i|e_j\rangle = c_i^* d_j\langle U_{dt}e_i|U_{dt}e_j\rangle \\ &= \langle U_{dt}c_i e_i|U_{dt}d_i e_j\rangle = \langle U_{dt}\psi(t)|U_{dt}\phi(t)\rangle = \langle\psi(t+dt)|\phi(t+dt)\rangle\end{aligned} \tag{1.47}$$

Therefore condition

> Time evolution is linear and maps orthonormal basis to orthonormal basis: $\langle U_{dt}e_i|U_{dt}e_j\rangle = \langle e_i|e_j\rangle = \delta_{ij}$ (DR-OBAS)

The above condition can be relaxed to just preserving pairs of orthonormal vectors. Take any two vectors ψ and ϕ, not necessarily orthogonal. We can write $\phi_\perp = \psi - \frac{\langle\phi|\psi\rangle}{\langle\phi|\phi\rangle}\phi$ and $\psi_\perp = \phi - \frac{\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle}\psi$. Note that $\langle\phi|\phi_\perp\rangle = 0$ and $\langle\psi|\psi_\perp\rangle = 0$. We have

$$\begin{aligned}\langle U_{dt}\phi|U_{dt}\psi\rangle &= \langle U_{dt}\phi|U_{dt}(\phi_\perp + \frac{\langle\phi|\psi\rangle}{\langle\phi|\phi\rangle}\phi)\rangle = \langle U_{dt}\phi|U_{dt}\phi_\perp\rangle + \frac{\langle\phi|\psi\rangle}{\langle\phi|\phi\rangle}\langle U_{dt}\phi|U_{dt}\phi\rangle \\ &= 0 + \frac{\langle\phi|\psi\rangle}{\langle\phi|\phi\rangle}\langle\phi|\phi\rangle = \langle\phi|\psi\rangle,\end{aligned} \tag{1.48}$$

which means the inner product is preserved.

Physical conditions

In classical mechanics we saw that Hamiltonian evolution was equivalent to determinism and reversibility. The same applies to quantum mechanics, as we will see by looking at different but equivalent conditions.

Determinism and Reversibility

Suppose we start with a pure state $\rho(t) = |\psi(t)\rangle\langle\psi(t)|$, meaning that the ψ is prepared with 100% probability. If U_{dt} is unitary, $\rho(t+\Delta t) = U_{dt}|\psi(t)\rangle\langle\psi(t)|U_{dt}^\dagger = |U_{dt}\psi(t)\rangle\langle U_{dt}\psi(t)| = |\psi(t+\Delta t)\rangle\langle\psi(t+\Delta t)|$, which means the final state is also a pure state. There is a one-to-one correspondence between initial and finial state, and therefore the evolution is deterministic and reversible. Conversely, if the evolution is deterministic and reversible, if we start with a pure state $\rho(t) = |\psi(t)\rangle\langle\psi(t)|$, then the final state will have to be $\rho(t+\Delta t) = |\psi(t+\Delta t)\rangle\langle\psi(t+\Delta t)|$. Given that both $\rho(t)$ and $\rho(t+\Delta t)$ are trace one operators, the norm of both $\psi(t)$ and $\psi(t+\Delta t)$ must be unitary. That is, a deterministic and reversible evolution must preserve the norm. Which means that

The evolution is deterministic and reversible (DR-EV)

is equivalent to DR-SCEQ.

Preservation of probability distributions

The same argument can be developed on probability distributions, similarly to what we have seen in classical mechanics. If we start with a probability distribution, the final probability distribution must be the same as all the probability for one case must be mapped and only mapped to a single other case. That is, suppose that we start with a mixed state $\rho(t) = p_i|e_i(t)\rangle\langle e_i(t)|$, meaning that it can be understood as a classical mixture of a set of orthogonal states. If we have a deterministic and reversibile evolution, the final state must be $\rho(t+\Delta t) = p_i|U_{dt}e_i(t)\rangle\langle U_{\Delta t}e_i(t)|$. This is the case if and only if time evolution maps an orthonormal basis to an orthonormal basis. This means the evolution is unitary and

The evolution preserves probability distributions (DR-EV)

is another equivalent condition.

Preservation of information entropy

When looking at classical mechanics, we saw that determinism and reversibility could be expressed as conservation of information entropy. This is true for quantum mechanics as well. Let $\rho(t) = p_i|e_i(t)\rangle\langle e_i(t)|$. If U_{dt} is unitary, we have $\rho(t+\Delta t) = p_i|U_{dt}e_i(t)\rangle\langle U_{\Delta t}e_i(t)|$. Since the transformed orthonormal basis is still an orthonormal basis, the entropy in both cases is given by $-\sum_i p_i \log p_i$, which means it is conserved.

Conversely, if an evolution preserves entropy then pure states must be mapped to pure states because all pure states and only pure states have zero entropy. Moreover, we saw that the square of the inner product characterizes the entropy of the mixture of a pair of states, which must be conserved if entropy is to be conserved. In particular, orthogonal states must remain orthogonal since they maximize entropy increase. Therefore, an evolution that

preserves entropy is one that preserver orthonormality and therefore it is unitary. That is, condition

The evolution preserves information entropy (DR-INFO)

is equivalent to DR-SCEQ.

Quasi-static processes

Note that in quantum mechanics both information entropy and thermodynamic entropy coincide, in the sense that we do not have two different definition in quantum statistical mechanics as we have in classical statistical mechanics (i.e. logarithm of count of states and Shannon entropy). However, we saw that projectors are more fundamental as they are implied by the mere definition of a Hilbert space, and projectors can be understood as equilibration processes. In thermodynamics, reversible processes are quasi-static processes. We can show that unitary evolution is, in this sense, a quasi-static process: unitary evolution can be understood as an infinite sequence of projections that perturb the system minimally.

To give intuition, suppose that a beam of light passes through two linear polarizers, the first oriented vertically and the second horizontally. No light will pass through. Recall, in fact, that the intensity decreases by a factor of $\cos^2\varphi$ where φ is the difference in angle between the polarizers. However, if you put another polarizer in between at a 45 degree angle, then some light will have a chance to be pass through. You can put another two polarizers so that the angle between any consecutive pairs is 22.5 degrees. More light will go through. We can imagine to repeat this process, until we have a large sequence of polarizers at a small angle. In that case, $\cos^2\varphi \approx 1 - \frac{\varphi^2}{2}$. Note that, to a first order, all light will go through. Therefore, in the limit, the net effect of the polarizers is to rotate the polarization of light from vertical to horizontal. This idea generalizes.

We saw, in fact, that for a unitary evolution $\langle\psi(t+dt)|\psi(t)\rangle = 1$, that is the projection of the state at a future time step on the previous time step is one. This can be understood as making a projective measurement on an observable that is slightly different to one for which $\psi(t)$ is an eigenstate. That is, we can understand unitary evolution as an infinitesimal sequence of projections at each time step. Note that the direction of the projection depends on the initial state. This is consistent with the evolution being deterministic: if we assume that the final state is the outcome of a projective measurement, the process is deterministic if and only if the choice of projective measurement depends on the initial state.

Another way of understanding this is that determinism and reversibility can be used for both measuring and preparing states. That is, if we prepare a system in a given state, we can use unitary evolution to prepare a system in the future state. Conversely, if we can measure a system in a given state, we can use unitary evolution to infer the state of the system at a prior time. Therefore, we can understand determinism and reversibility as a series of preparations or measurements. Since measurements in quantum mechanics are projections, it makes sense that we can understand unitary evolution as a sequence of projections. Therefore

Time evolution is a quasi-static process (DR-PSEQ)

and

The evolution is an infinite sequence of reversible measurements (DR-MSEQ)

are equivalent to DR-EV.

1.4 Projection and measurements

WTS: Every state is an eigenstate of a unitary evolution, of a projection, and of an Observable.

Given a pure state $|\psi\rangle$ we can always create the operator $P_\psi = |\psi\rangle\langle\psi|$. This operator has $|\psi\rangle$ as an eigenvector with eigenvalue one, and any state $|\phi\rangle$ that is orthogonal to $|\psi\rangle$ will also be an eigenvector with eigenvalue zero. In fact:

$$\begin{aligned} P_\psi|\psi\rangle &= |\psi\rangle\langle\psi|\psi\rangle = |\psi\rangle \\ P_\psi|\phi\rangle &= |\psi\rangle\langle\psi|\phi\rangle = |\psi\rangle 0 = 0 \end{aligned} \tag{1.49}$$

Note that $P_\psi^\dagger = P_\psi$ which means that this is a Hermitian operator and since $P_\psi^\dagger P_\psi = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle 1\langle\psi| = |\psi\rangle\langle\psi|$

$X = |x_i\rangle x_i \langle x_i|$

Projections are processes with equilibria (all fine states are equilibria) * (?) Projection is not enough: need compatility with a unitary evolution * Show that projections cannot decrease entropy * Eigenstates of projections are equlibria =¿ all quantum states are equilibria of projection - Mathematically, this is what Hilbert spaces add on top of Banach spaces * Analogy to thermodynamics (context is like different type of ensembles) * Unitary evolution is quasi-static evolution (like in thermodynamics) - Make the parallel to S-matrix calculation where we put initial state at minus infinity, and final state at plus infinity for a process that actually last "femtoseconds"

1.5 Observables

Observables are self-adjoint operators (textbook definition)

Definition 1.50. *Observables are associated with self-adjoint operators $\mathcal{O} : \mathcal{D}_\mathcal{O} \to \mathcal{H}$, where the domain $\mathcal{D}_\mathcal{O}$ is a dense subset[4] of $\mathcal{H}$.*

This means that the adjoint $\mathcal{O}^\dagger$ of $\mathcal{O}$, implicitly defined by

$$\langle \mathcal{O}^\dagger \psi | \phi \rangle = \langle \psi | \mathcal{O}\phi \rangle \tag{1.51}$$

with domain

$$\mathcal{D}_{\mathcal{O}^\dagger} = \{\psi \in \mathcal{H} \mid \exists \xi \in \mathcal{H} : \forall \phi \in \mathcal{D}_\mathcal{O} : \langle \xi | \phi \rangle = \langle \psi | \mathcal{O}\phi \rangle\} \tag{1.52}$$

is identical to $\mathcal{O}$ itself, including their domains.

Observables are generators of unitary transformations

Theorem 1.53 (Stone's Theorem)**.** *For any strongly continuous one-parameter unitary group $U_t : \mathcal{H} \to \mathcal{H}$, there exists a unique self-adjoint operator $A : \mathcal{D}_A \to \mathcal{H}$, with dense domain*

$$\mathcal{D}_A = \left\{ \psi \in \mathcal{H} \,\middle|\, \lim_{\epsilon \to 0} \frac{1}{i\epsilon}(U_\epsilon(\psi) - \psi) \text{ exists} \right\} \tag{1.54}$$

that generates the group via

$$U_t = e^{itA} \tag{1.55}$$

Conversely, any self-adjoint operator generates a strongly continuous one-parameter unitary group.

[4] On a Hilbert space continuous is equivalent to bounded. Unbounded observables are therefore either discontinuous or have a restricted domain.

Remark. The domain $\mathcal{D}_A$ coincides with all of $\mathcal{H}$, if and only if the map $t \to U_t$ is continuous w.r.t. the operator norm.

TODO Do we want to use this? Something like this holds in classical mechanics, too. Maybe we can find a common reason then.

WORK IN PROGRESS: Observables have real spectrum

From Wikipedia:

- Self-adjoint operators have real spectrum, but there are also unbounded operators with real spectrum that are **NOT** self-adjoint.
- Symmetric operators with a real spectrum are self-adjoint (Wikipedia references [3]).

I don't need a spectral theorem to talk about the spectrum. The theorems are about spectral decompoisitons (diagonalization).

There exist multiple different versions of the spectral theorem in the literature

- In terms of projection valued measures (PVMs) (see Theorem 10.4 in [1])
- In terms of multiplication operators (see Theorem 10.10 in [1])
- A nuclear spectral theorem in the context of rigged Hilbert spaces (Sec. 4.5 in [2])

TODO Figure out which one we want to use.

TODO Find formulation of nuclear spectral theorem that doesn't require too many prerequisites.

XXX There is no general spectral theorem for unbounded operators, but maybe for normal operators. Characterizing self-adjoint operators as those with real spectrum might be hard (or not useful). Alternatively, they might be exaclty those constructed from a projection-valued measure on the real line.

Theorem 1.56 (Spectral Theorem)**.** *Suppose A is a self-adjoint operator on $\mathcal{H}$. Then there is a unique projection-valued measure μ^A on $\sigma(A)$ with values in $\mathcal{B}(\mathcal{H})$ such that*

$$\int_{\sigma(A)} \lambda \mathrm{d}\mu^A(\lambda) = A \,. \tag{1.57}$$

Theorem 1.58 (Spectral Theorem [Multiplication Operators])**.** *Suppose A is a self-adjoint operator on $\mathcal{H}$. Then there is a σ-finite measure space (X, μ), a measurable, real-valued function h on X, and a unitary map $U : \mathcal{H} \to L^2(X, \mu)$ such that*

$$U(\mathcal{D}_A) = \{\psi \in L^2(X, \mu) \,|\, h\psi \in L^2(X, \mu)\} \tag{1.59}$$

and such that

$$(UAU^{-1}(\psi))(x) = h(x)\psi(x) \tag{1.60}$$

for all $\psi \in U(\mathcal{D}_A)$.

WORK IN PROGRESS: Observables are linear functionals of observables

In [1] we can find the converse claim, i.e., states are linear functionals of observables (and more).

Theorem 1.61. *Let $\Phi : \mathcal{B}(\mathcal{H}) \to \mathbb{C}$ be a **family of expectation values**, i.e.,*

- $\Phi(\alpha A + \beta B) = \alpha\Phi(A) + \beta\Phi(B)$
- $\Phi(\mathrm{Id}_{\mathcal{H}}) = 1$
- $\Phi(A) \in \mathbb{R}$ *for A self-adjoint and* $\Phi(A) \geq 0$ *if A is non-negative as well*
- Φ *is strongly continuous in the sense that for any sequence* $\{A_n\}_{n=1}^{\infty} \subset \mathcal{B}(\mathcal{H})$ *with* $\|A_n\psi - A\psi\| \to 0 \quad \forall \psi \in \mathcal{H}$ *it follows that* $\Phi(A_n) \to \Phi(A)$.

Then there exists a unique density matrix ρ, s.t.

$$\Phi(A) = \mathrm{tr}(\rho A) \,. \tag{1.62}$$

Conversely, any density matrix defines a family of expectation values via eq. (1.62).

Issues:

- It's obviously in the wrong direction. We don't want to characterize observables before states.
- Not all observables are represented by bounded operators.
- Not all bounded operators represent observables.

TODO Find finite-dimensional proof of what we actually want. Should be easier.

TODO list

* (?) Convex maps of mixed states are Hermitian operators * Is it useful to note that any observable is compatible to some unitary? That is, any observable is left unchanged by a unitary?

1.6 next

Open quantum systems * Review open quantum system (Lindblad master equation) * (?) recover CPTP maps - Linear maps : map mixed states to mixed states while preserving mixtures - Trace preserving : map trace one operators to trace one operators - Positive : mixed states have non-negative eigenvalues - Completely Positive: (?) need a characterization of completely positive in terms of only the system, without the ancilla * (?) Kraus operator, jump operatos: how are they related? If they are? * (?) What are the possible motions on a Bloch sphere? That is, what are the possible vector fields described by the Lindblad equation

Classical limit * Classical mechanics is the high entropy limit of quantum mechanics - Find classical transformations that increase entropy, show that they are all "unitaries" plus stretch of phase space - Find equivalent of phase space stretching in quantum mechanics - See that it is a CPTP map only defined in the anti-normal ordering - Show that it rescales the commutator by the factor for phase space stretching - Show that this is equivalent to the limit of $\hbar \to 0$

Negative probability in quantum mechanics * QM on phase space (Wigner functions - Hussimi - GlauberSudarshan) * (real) convex combination vs affine combination vs linear

combination * =¿ QM on phase space is using affine combinations, since convex combinations are not sufficient

Quantum states as equibria * Show that for a unitary evolution, eigenstates are equilibria * Show that any quantum state is an eigenstate of some unitary * =¿ all quantum states are equilibria of unitary

(?) Recover spin 1/2 (two state systems) * Space of ensembles that is fully characterized by an average direction. - Gaussian states are fully identified by average and standard deviation - Suppose we have "guassian states" of directions with same standard deviation - -¿ the space is a ball - (?) how much can we recover? * Space of directional pure states - I have a state space for directions in space - Recycle the argument that we have to be able to put a frame invariant distribution over it - -¿ two sphere is the only symplectic sphere and therefor is the only space - (?) why are ensembles the Bloch?

1.7 Problems with infinite dimensional spaces

Problems with Hilbert spaces

- All infinite-dimesional (separable) Hilbert spaces are isomorphic to each other (The state space of one particle is indistinguishable from that of a billion particles)
- $[X, P] \neq \imath\hbar$
- Hilbert spaces necessarily contain states with undefined (e.g., divergent) moments/expectation values

Connection between observables and states

- Expectations of polynomials of x and p are not enough to recover Schwartz functions
- Expectations of Weil operators are enough to recover Schwartz functions
- Probability distributions of all linear combinations of x and p are enough to recover Schwartz functions
- Expectations of polynomials of x and p are enough to recover functions in $S^{1/2}_{1/2}$
- What is the space of all functions with well defined expectations for all polynomials of x and p? Is it Schwartz or bigger?

In the previous sections we restricted ourselves to finite dimensional spaces. In these spaces all measurements can be understood as having finitely many outcomes, and all those outcomes can be understood as quantum states. In this section we will extend the discussion to the infinite case and see the extension is problematic. Infinite dimensional Hilbert spaces, in fact, seem to hide the source of the infinity and require the existence of states that cannot be thought as being physically meaningful. The exact mathematical representation of these cases, then, is still an open problem.

There are two potential sources of infinity in physics: the infinitely large and the infinitely small. The infinitely large comes from unbounded quantities. For example, the number of particles in a gas or the distance of a particle from the origin of our reference system can be, in principle, arbitrarily large. In this case, infinity is just the range of possible values, and not a value itself. It would not make sense, for example, to say that a gas has infinitely many particles or that a particle is infinitely distant from the origin unless we are talking about the

limit of a process that takes an infinite amount of time. Note that the infinitely large does not change the nature of the quantity: the number of particles is a discrete quantity and the position is a continuous quantity regardless of whether we are allowing an infinite range or not.

The infinitely small comes from the ability to refine measurements indefinitely. For example, we assume we can measure the position of a particle with arbitrary precision. While the infinite precision measurement is never realizable, the infinite precision value can be understood as the information needed to specify the finite precision outcome at all level of precision. That is, if we knew the position with infinite precision we would know all the possible finite precision intervals in which the particle can be. The infinitely small, then, changes the nature of the quantity, going from discrete to continuous. Over a finite range, a discrete quantity will have finitely many possible values while a continuous quantity will have infinitely many.

We can characterize these differences in the following way. A measurement will tell us whether a particular value is within a set of possible values. If the quantity is continuous, all measurements will always restrict the range of possible values to an infinite set. That is, any measurement of position will have a finite uncertainty, which will include infinitely many possible positions. If the quantity is discrete, at some point, we will have a measurement that identifies each possible case. That is, we can count exactly the number of particles, which restricts the measurement to only one possible case. Intuitively, the range is infinite in either cases if it cannot be always covered with finitely many measurements. That is, if we are given measurements with finite ranges, we are not going to be able to cover the infinite range with finitely many measurements.

Mathematically, this maps to properties of open sets of the topology. The topology, in fact, keeps track of the notion of closeness, and measurement resolution is about that closeness. The topology of a real line, then, is different from the topology of sets of points: the first one is topologically connected, while the second one is not. A space with a finite range is topologically compact, while one that has an infinite range is not. Establishing a perfect mapping between these concepts is not the goal of Reverse Physics, but rather of Physical Mathematics.

In classical mechanics, the mathematics characterizes and keeps track of these differences. The problem is that in quantum mechanics, the mathematical framework does not keep track of these differences. This is the problem we are going to explore in this section, a problem that is ultimately not solved. Given that the mathematical framework does not capture all the elements and only the elements that are physically meaningful, the Reverse Physics program cannot be fully completed for quantum mechanics.

Equivalence of Hilbert spaces

One feature of Hilbert spaces is that the cardinality of the base fully characterizes the space. In fact, let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces with the same cardinality and let $\{e_i\}_{i\in I} \in \mathcal{H}_1$ and $\{g_i\}_{i\in I} \in \mathcal{H}_2$ be two orthonormal basis of the respective spaces. Then we can define a map $m : \mathcal{H}_1 \to \mathcal{H}_2$ such that $m(\sum a^i e_i) = \sum a^i g_i$. The map is unitary since

$$\langle a^i e_i | a^i e_i \rangle = \langle a^i g_i | a^i g_i \rangle = \sum |a^i|^2 \tag{1.63}$$

which means the two spaces are unitarily equivalent and therefore they are the same Hilbert space.

All spaces we are interested in quantum mechanics are going to have a countable basis. If we imagine to extend to infinite the range of a discrete observable, we can see that we will have a countable set of possible outcomes and therefore a countable basis. If we take the space of wave functions, in either a finite or infinite range, we will get the space of square integrable functions, which also has a countable basis. To see that, note that the Hamiltonian for the harmonic oscillator has a discrete spectra and therefore has countably many eigenstates which will form a basis. All wave-functions, then, can be written as a superposition of eigenstates of the harmonic oscillator and therefore the space has a countable basis.

In quantum mechanics, then, if the space has an infinite basis, it will be a countable one, which leads to the following observation: all infinite dimensional spaces in quantum mechanics are equivalent. If we have an infinite dimensional space we are not going to be able to know whether we have a discrete quantity over an infinite range, a continuous quantity over a finite range or a continuous quantity over an infinite range. Mathematically, there will be no difference in the states, unlike in classical mechanics.

For example, the state space of a single DOF (i.e. $L^2(\mathbb{R})$) is going to be equivalent to the state space of n DOFs (i.e. $L^2(\mathbb{R}^n)$)). As we said before, the Hermite functions, the eigenstates of a Harmonic oscillator, provide a countable basis for a single DOF. We can imagine a Harmonic oscillator over 2 DOFs. In that case, the products of the Hermite functions across degrees of freedom provide a countable basis for the space. Through diagonalization, we can map one set of basis to the other

$$h_0(x) \mapsto h_0(x)h_0(y) \tag{1.64}$$
$$h_1(x) \mapsto h_0(x)h_1(y) \tag{1.65}$$
$$h_2(x) \mapsto h_1(x)h_1(y) \tag{1.66}$$
$$h_3(x) \mapsto h_0(x)h_2(y) \tag{1.67}$$
$$\dots \tag{1.68}$$

This idea can be generalized to map any set of n DOFs to any set of m DOFs. Potentially, it means that we can construct a process that can encode the information of a billion particles into a single particle.

TODO: add picture of diagonalization

Note that the above transformation is not possible in classical mechanics. One degree of freedom, topologically, is $\mathbb{R}^2$ while n DOFs are $\mathbb{R}^{2n}$, which are not topologically equivalent. The issue is that the topology of the Hilbert spaces in quantum mechanics is the one induced by the inner product and not the one induced by the observables, like in classical mechanics. Therefore the mathematical representation in quantum mechanics knows "too little" about the physical system it is supposed to describe.

Domain of operators

In the infinite dimensional case, the HellingerToeplitz theorem tells us that any Hermitian[5] operator O that is define on the whole space is bounded. Therefore any unbounded operator O will not be defined on the whole Hilbert space. Since variables like position and energy are unbounded, there will be some state ψ for which the energy is infinite, or the average position

[5] More precisely, symmetric. which means $\langle O\psi|\phi\rangle = \langle\psi|O\phi\rangle$.

is undefined. This is clearly a problem, and it is instructive to understand where the problem comes from.

TODO: cannibalize the paper

TODO: add plots for the wave-functions

For example, consider the following wave function

$$\psi(x) = \sqrt{\frac{1}{\sqrt{\pi}} e^{-x^2}} \tag{1.69}$$

$$\rho_\psi(x) = \psi^\dagger(x)\psi(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}. \tag{1.70}$$

This is a Gaussian wave packet with expectation of zero and variance of $\frac{1}{2}$. Now consider the following wave function

$$\phi(x) = \sqrt{\frac{1}{\pi(x^2+1)}} \tag{1.71}$$

$$\rho_\phi(x) = \phi^\dagger(x)\phi(x) = \frac{1}{\pi(x^2+1)}. \tag{1.72}$$

Note $phi(x)$ goes to zero as $\frac{1}{x^2}$. The expectation, then, will converge and will converge to zero since the distribution is symmetric. However, the variance will diverge since $\lim_{x\to\infty} x^2 \frac{1}{\pi(x^2+1)} = \frac{1}{\pi}$.

Both of these will correspond to two vectors in the Hilbert space, meaning we are going to be able to find a unitary transformation that changes one to the other. In this case, we can do that by a change of variable. That is, we are going to look for a transformation$y = y(x)$ that transforms one wave function into the other. What we require is that the integral of one function over one region equals the integral of the second function on the second region. That is

$$\begin{aligned}
\int_0^{y(x)} \phi^\dagger(\hat{y})\phi(\hat{y})d\hat{y} &= \int_0^x \psi^\dagger(\hat{x})\psi(\hat{x})d\hat{x} \\
\int_0^{y(x)} \frac{1}{\pi(\hat{y}^2+1)} d\hat{y} &= \int_0^x \frac{1}{\sqrt{\pi}} e^{-x^2} d\hat{x} \\
\frac{\tan^{-1}(y(x))}{\pi} &= \frac{\operatorname{erf}(x)}{2} \\
y(x) &= \tan\left(\frac{\pi}{2}\operatorname{erf}(x)\right).
\end{aligned} \tag{1.73}$$

This change of variable, then, maps a state with finite expectation of position square to a state with infinite expectation. Changes of variable are unitary transformation on the Hilbert space, so in general a unitary transformation can map finite expectations to infinite expectations.

To be clear, the position in each reference frame corresponds to different observables. That is, X and Y are going to map to two different Hermitian operators. TODO finish the paragraph.

Moreover, in a Hilbert space we can always map one vector to another vector through a continuous unitary transformation, through continuous evolution. Mathematically, we can always rotate one vector on top of another. Physically, this means that we can construct a time

evolution operator that oscillates between the two states. For example, suppose the evolution is such that the position chances in the following way

$$x(t) = \cos(\omega t)x_0 + \sin(\omega t)\tan\left(\frac{\pi}{2}\operatorname{erf}(x_0)\right). \tag{1.74}$$

This is a continuous transformation in t. For $t = 0$ we get the identity as $x(0) = x_0$. For $t = \frac{\pi}{2}$ we get $x(0) = \tan\left(\frac{\pi}{2}\operatorname{erf}(x_0)\right)$. Therefore $\psi(x)$ is the initial state, the evolution will oscillate between the two states, transforming finite expectation to infinite expectation and vice-versa in finite time, over and over.

Continuous spectra

In the finite dimensional case, we are used to associated possible values of an observable to states (i.e. the eigenstates) associated to that value. In the infinite dimensional case, we have operators that have a continuous spectra, like position or energy. There is a temptation to extend the previous scheme to the continuous one, adding eigenstates of continuous values. This is problematic, not just mathematically, but physically as well. Simply put, since we cannot prepare or measure a continuous quantity with infinite precision, these states are neither physically realizable nor measurable.

As we saw in the classical mechanics section, states as points in phase space (i.e. point particles) do not make sense in classical mechanics either. Volumes and areas define the geometry of phase space as these define the count of configurations per DOF and the count of states. Hamiltonian mechanics is exactly the conservation of those areas and volumes. A single point, having no volume, can have no such notions. Classical point particles, then, should be understood as an infinitesimal region of phase space. A region so small that we do not care about its size, but still a region.

In quantum mechanics, talking about particles at a point makes even less sense. If we shrink the spread over position to zero, we are forced to stretch the spread over momentum to infinity. A uniform distribution over an infinite range makes even less sense than a distribution all concentrated at one point. Mathematically, a probability distribution over a single point is still a measure, a uniform probability distribution over an infinite range is not a measure. One way to see this, is that it does not satisfy countable additivity. We can imagine dividing the real line into countable intervals of equal size. Each should correspond to the same probability, and the sum of all the contributions should be one. If each interval corresponds to finite probability, the sum is infinite; if each interval corresponds to zero probability, the sum corresponds to zero probability as well. The solution would be to sacrifice countable additivity, so that each interval has zero probability while the total is one. Therefore, this would not satisfy the axioms of measure theory and probability theory.

Suppose we do want to go ahead and include distributions wholly concentrated at one point (i.e. delta function) in our state space. These are not square integrable function as they diverge, they are infinite, at one point. What is the inner product between such distribution and a standard state? It will be

$$\int \delta(x)\psi(x)dx = \psi(0). \tag{1.75}$$

This is called the sifting or sampling property of the delta function.

TODO:

, meaning that they do not have "proper" eigenstates. For example, for the position operator, the eigenvectors would correspond to the delta functions, which are not in the Hilbert space, but rather in the space of distributions. Therefore, there are no states with perfectly prepared position and the unitary transformation generated by position has no fixed points (i.e. equilibria, eigenstates).

Schwartz spaces

For multiple independent DOF of position and momentum, the requirement to have all polynomials of position and momentum with a finite expectation value recovers the Schwartz space, which is a dense subspace of the Hilbert space (i.e. any element of the Hilbert space can be understood as the limit of a sequence of Schwarz functions).

TODO: talk about the difference between Schwartz, Hilbert and Schwartz dual (distributions)

TODO: Does it make physical sense to have a mixed state over Schwartz where the probability coefficients converge as a polynomial.

Probability on a continuum

For any observable with any spectra, the probability measure can be recovered in the following way. Take a Borel set. Construct the projector onto that Borel set. For example, take X as the position operator. Take a Borel set A and take the indicator function 1_A and calculated $\langle\psi|1_A|\psi\rangle = \int_{\mathbb{R}} \psi^\dagger 1_A \psi dx = \int_A \psi^\dagger 1 \psi dx = \int_A \psi^\dagger \psi dx$.

Conjecture: whether finite expectations map to finite expectation if and only if the velocity is bound.

Self-adjoint vs Hermitian

In some spaces, self-adjoint is not equivalent to Hermitian. Consider the half real-line $[0, +\infty]$. Consider the momentum operator $\imath\hbar\partial_x$. The exponential $e^{-\lambda x}$ where $\lambda > 0$ is an eigenstate of momentum

$$\imath\hbar\partial_x e^{-\lambda x} = -\lambda\imath\hbar e^{-\lambda x} \tag{1.76}$$

with eigenvalue is $-\lambda\imath\hbar$ which is imaginary.

In the finite dimensional case, all self-adjoint are Hermitian.

Definition of Hermitian adjoint. Given O, the Hermitian adjoint $O^\dagger$ is such that:

$$\langle\psi|O\phi\rangle = \langle O^\dagger\psi|\phi\rangle \tag{1.77}$$

where $\psi, \phi \in D(O)$ (i.e. the operator is defined on the vectors).

Wavefunctions and equivalent classes

Schauder basis

TODO consider moving this into a bare minimum for quantum mechanics

Usually, one defines a basis B of a vector space V in linear algebra as a linearly independent subset of V that spans V. More explicitly

linear independence means that every *finite* subset $\{e_1, \ldots, e_n\} \subset B$ is linearly independent, i.e., $a_1 e_1 + \ldots + a_n e_n = 0 \iff \forall 1 < i < n : a_i = 0$.

spanning means that for every vector $v \in V$ there exists a finite linear combination of some elements in B, s.t., $v = a_1 e_1 + \ldots + a_n e_n$.

This definition is unpractical[6] for infinite-dimensional Hilbert spaces and the typical notion of basis differs from this. To differentiate we call the previous defintion that of a Hamel basis, while the following is called a Schauder basis.

A **Schauder basis** B of a separable Hilbert space $\mathcal{H}$ is a *sequence* e_n of vectors, s.t., for every vector $v \in \mathcal{H}$ there exists a *unique* sequence a_n such that

$$v = \sum_{n=1}^{\infty} a_n e_n \ . \tag{1.78}$$

Convergence is defined w.r.t the topology implicitly defined through the inner product on the Hilbert space.

Note that we had to explicitly define an *order* for the basis elements, since (1.78) may not converge unconditionally.

Physical states and well-defined moments

As we have seen before expectation values of observables are not always well-defined. Here we find another explicit example where a moment is not necessarily infinite, but depends on the order of basis.

Consider the state

$$\rho = \sum_{n=1}^{\infty} p_n |\psi_n\rangle\langle\psi_n| = \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} |\psi_n\rangle\langle\psi_n| \tag{1.79}$$

and the observable

$$\mathcal{O} = \sum_{n=1}^{\infty} \mathcal{O}_n |\psi_n\rangle\langle\psi_n| = \sum_{n=1}^{\infty} (-1)^n n |\psi_n\rangle\langle\psi_n| \tag{1.80}$$

with some domain $\mathcal{D}_{\mathcal{O}}$, which makes $\mathcal{O}$ into a self-adjoint operator. Both of them are defined w.r.t. some ONB $\{\psi_n\}_{n=1}^{\infty} \subset \mathcal{H}$.

While calculating the expectation value of $\mathcal{O}$ for ρ in our ONB, we encounter the harmonic series[7]

$$\langle\mathcal{O}\rangle_\rho := \mathrm{tr}\,(\rho\mathcal{O}) = \sum_{n=1}^{\infty} p_n \mathcal{O}_n = \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^n n}{n^2} = \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} = \frac{6}{\pi^2} \ln(2) \tag{1.81}$$

This series is *conditionally* convergent and according to Riemann's rearrangement theorem it can therefore take *any* value in $[-\infty, +\infty]$ or not converge at all, just by rearranging the order of its terms, which is equivalent to a reordering of the basis elements.

[6] I couldn't find any explicit example for a definition of a basis in the above sense for a separable Hilbert space. In fact, since it's existence is only guaranteed by the axiom of choice, I suspect it to be impossible to explicitly construct an example.

[7] This is of course no coincidence, but by construction.

Example for pure states

One can construct a similar example with a pure state. Consider

$$\psi = \sum_{n=1}^{\infty} \psi_n |\psi_n\rangle = \frac{\sqrt{6}}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} |\psi_n\rangle \tag{1.82}$$

and then calculate the expectation value as

$$\langle \mathcal{O} \rangle_\psi := \langle \psi | \mathcal{O} | \psi \rangle = \sum_{n=1}^{\infty} |\psi_n|^2 \mathcal{O}_n = \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^n n}{n^2} = \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} = \frac{6}{\pi^2} \ln(2) \, . \tag{1.83}$$

The result is the same and will depend on the order of the basis elements in the same way.

While it's obvious that such a state can never be found in the lab, we need a criterion to distinguish the physical from the unphysical states.

For a pure state we can consider the domain of the operator as an indicator of what states have well defined expectation values. Let's check if ψ could possibly be in the domain of $\mathcal{O}$. For that, calculate

$$\begin{aligned} \mathcal{O}|\psi\rangle &= \sum_{m=1}^{\infty} (-1)^m m |\psi_m\rangle\langle\psi_m| \frac{\sqrt{6}}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} |\psi_n\rangle \\ &= \frac{\sqrt{6}}{\pi} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (-1)^m \frac{m}{n} |\psi_m\rangle\langle\psi_m|\psi_n\rangle \\ &= \frac{\sqrt{6}}{\pi} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (-1)^m \frac{m}{n} \delta_{mn} |\psi_m\rangle \\ &= \frac{\sqrt{6}}{\pi} \sum_{n=1}^{\infty} (-1)^n |\psi_n\rangle \end{aligned} \tag{1.84}$$

and notice that the result is not normalized

$$\||\mathcal{O}\psi\||^2 = \frac{6}{\pi^2} \sum_{n=1}^{\infty} |(-1)^n|^2 = \frac{6}{\pi^2} \sum_{n=1}^{\infty} 1 = \infty \tag{1.85}$$

and we conclude that ψ can not be in any sensible domain of $\mathcal{O}$.

Operator domain and the second moment

The previous observation that the domain indicates what states have finite expectation values can be generalized. Consider any unbounded observables $\mathcal{O} : \mathcal{D}_\mathcal{O} \to L^2(\mathbb{R}^n)$ and let's assume a very conservative domain of

$$\mathcal{D}_\mathcal{O} = \mathcal{D}_{\max} := \left\{ \psi \in L^2(\mathbb{R}^n) \mid \mathcal{O}\psi \in L^2(\mathbb{R}^n) \right\} \, , \tag{1.86}$$

i.e., the resulting function must at least be in $L^2(\mathbb{R}^n)$.

Then, by definition it must be true that

$$\||\mathcal{O}\psi\||^2 < \infty \tag{1.87}$$

which is equivalent to

$$\||\mathcal{O}\psi\||^2 = \langle \psi | \mathcal{O}^\dagger \mathcal{O} | \psi \rangle = \langle \psi | \mathcal{O}^2 | \psi \rangle = \langle \mathcal{O}^2 \rangle_\psi < \infty \tag{1.88}$$

which means that at least the second moment has to be well-defined.

Insight 1.89. *States in the domain of unbounded operator must have at least a well-defined second moment.*

1.8 WARNINGS

Every state (even if I restrict to Schwartz space) is an eigenstate of some projection, observable and unitary, because $|\psi\rangle\langle\psi|$ is always defined.

Not every projector, observable or unitary can be understood as having eigenstates. For example, projection on a Borel set of position.

Questions: * Can we see the position operator as the limit of a discretized position operator where the discretization is smaller and smaller? If we did the same with momentum, would the commutator between them become $\imath\hbar$.

(?) Generalize this to arbitrary dimensions

(?) Random things to look at to see whether they are helpful * Khler manifold, interplay of symplectic structure with metric tensor * Is anything of the old arguments salvageable? * See if there is anything we can get from GPT or other reconstructions

1.9 Open problem: Extremality of quantum ensembles

Category: Quantum mechanics - Tags: Convex spaces, Choquet theory, Projective spaces

Classical ensembles are exactly the ones in which every probability distribution over the pure states correspond do a distinct ensemble. The question is whether quantum ensembles have, in some sense, the opposite property: each quantum ensemble represents the largest equivalence class of indistinguishable distributions over pure states.

As a motivating example, consider a two dimensional sphere. This can represent the space of classical space, and therefore any probability measure over the sphere represents a distinct ensemble. But it can also represent the Bloch sphere, in which case the space of ensemble is the Bloch ball, and every probability distribution over pure states corresponds to the barycenter of that distribution. Note that, since the sphere is a two dimensional object, the three dimensional ball represents the smallest convex space that has the sphere as its distinct extreme points. Physically, a classical ensemble gives us perfect reconstruction of the preparation, while the quantum ensemble gives us the least information about the preparation. The question is whether this insight is particular to the two dimensional case, or it can be generalized. The first step would be to generalize it to the finite dimensional case.

The formal setup would be the following. Let X be a symplectic manifold which characterizes the space of pure states. Let $\mathcal{P}(X)$ be the space of probability measures, which is naturally equipped with the total variation topology and a convex (continuous) operation representing statistical mixing. Let $\sim$ be an equivalence class that is compatible with both the topology and the convex structure. Moreover, $\sim$ should keep extreme point (i.e. pure states) extreme and distinct. Equivalently, the equivalence class for each pure states is its singleton. The minimal equivalence class is the identity, which recovers the full probability simplex. The question is whether there exists some maximality condition that recovers the quantum case.

1.10 Open problem: Spectral theorem on Fréchet spaces

Category: Quantum mechanics - Tags: Spectral theorem, Functional analysis

As we try to understand what is the correct physically motivated mathematical framework for quantum mechanics, we need to generalize the spectral theorem to a subset of the Hilbert space that can be justified mathematically.

From other work, it seems that a more physically motivated setting for quantum mechanics is a complete separable Hausdorff locally convex topological vector space, where the topology is the one that makes a core set of observables (i.e. self-adjoint operators) continuous. The idea is that all observables have to be continuous (i.e. small state changes lead to small observable changes), the topology is generated by the observables (i.e. we distinguish state through them) and the space is complete in that topology (i.e. any valid combination of observable is a valid physical state).

Additionally, observables have to be understood as generator of physical transformation (i.e. one parameter unitary group), so the commutation relationships between them must be understood as relationship between the groups (i.e. $[X,Y]=0 \iff e^{\frac{Xt}{\imath\hbar}}e^{\frac{Ys}{\imath\hbar}} = e^{\frac{Ys}{\imath\hbar}}e^{\frac{Xt}{\imath\hbar}}$). Similarly, an eigenstate of X has to be understood as an eigenstate of $e^{\frac{Xt}{\imath\hbar}}$.

In this context, we want a spectral theorem like the following. Let V be a separable Fr'echet space, and X_i be a maximal set of generators that (strongly) commute. Then there is a representation of V in which those operators are multiplicative. That is, there exists a measurable space (Y,Σ,μ) and a continuous unitary operator $U : V \to W \subseteq L^2(Y,\mu)$ such that $U^{-1}T_iU = X_i$ and T_i are multiplicative operators $(Tf)(y) = \lambda(y)f(y)$.

1.11 Open problem: Ensembles for Schwartz spaces

Category: Quantum mechanics - Tags: Spectral theorem, density operators

While Schwartz spaces are the right space for pure states identified by position and momentum, we need to understand what are the corresponding ensembles.

On a Hilbert space, density operators are defined as trace-one positive semi-definite self-adjoint operators. On a Schwartz space, this definition does not work as we have no general notion of trace-class operators. There are two approaches that, hopefully, are equivalent.

As a subset of density operator. Let $\mathcal{S}$ be a Schwartz space and $\mathcal{H}$ its Hilbert completion. Let $\mathcal{D}(\mathcal{H})$ be the space of density operators. Let $\mathcal{D}(\mathcal{S})$ be the set of density operators such that $\rho(\mathcal{S}) \subseteq \mathcal{S}$ (i.e. each density operator maps Schwartz functions to Schwartz functions) and $\operatorname{tr}[X^nP^m\rho] < \infty$ (i.e. the expectation of all monomials of position and momentum is finite). We call these Schwartz density operators. Are the eigenstates of Schwartz density operators Schwartz functions?

As the completion of the convex set of projectors. Let $\mathcal{S}$ be a Schwartz space. Let $D(\mathcal{S})$ be the convex set generated by all one-dimensional projectors. Note that $D(\mathcal{S})$ contains only self-adjoint operators. Given a self-adjoint operator X and $\rho = \sum_i p_i P_{\psi_i} \in D(\mathcal{S})$, define the expectation $E[X|\rho] = \sum_i p_i E[X|P_{\psi_i}] = \sum_i p_i \frac{\langle\psi_i|X\psi_i\rangle}{\langle\psi_i|\psi_i\rangle}$. Note that $E[I|\rho] = \sum_i p_i \frac{\langle\psi_i|\psi_i\rangle}{\langle\psi_i|\psi_i\rangle} = \sum_i p_i = 1$. We now want to complete $D(\mathcal{S})$ by taking all limits of sequences of $D(\mathcal{S})$ where the expectation of all everywhere-defined self-adjoint operators remains finite. Let $\mathcal{D}(\mathcal{S})$ be such completion. Does $\mathcal{D}(\mathcal{S})$ consist of only self-adjoint operators that map Schwartz to Schwartz? Can they always be expressed as the infinite convex combination of countably many projectors?

Do the two definition coincide?

Bibliography

[1] Brian C. Hall. *Quantum Theory for Mathematicians*, Springer, New York, 2013, http://dx.doi.org/10.1007/978-1-4614-7116-5.

[2] I. M. Gelfand and N. Y. Vilenkin. *Generalized functions. Applications of harmonic analysis*, Academic Press, New York 1964.

[3] N. Bebiano and J. da Providncia *Non-self-adjoint operators with real spectra and extensions of quantum mechanics*, Journal of Mathematical Physics **60**(1), 2019, https://doi.org/10.1063/1.5048577.