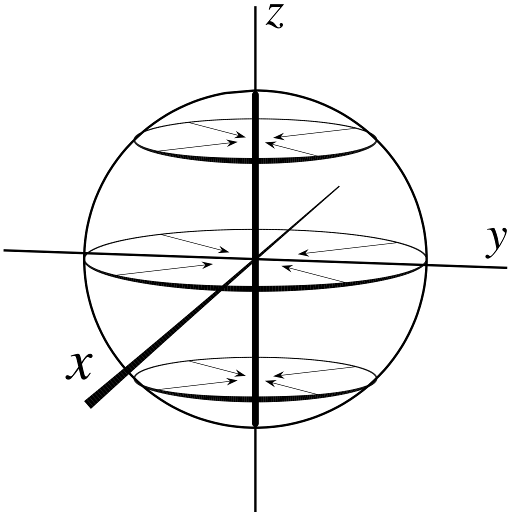
Introduction to quantum information and computing

Bas Janssens

2023/2024



Contents

Preface

These are lecture notes for the master course *introduction to quantum information and computing* (WI4645) at the TU Delft. They are used together with the book [NC00] by Nielsen and Chuang. With the possible exception of §3.3 (which is based on [JM06]), most of the content of these notes can be found in many other sources. Rather than attempting originality, we have based our exposition on the above mentioned book [NC00] by Nielsen and Chuang, and on the excellent set of lecture notes [M04].

Many problems are an integral part of the text. On a first reading I would advise to also read the problems, and perhaps reflect briefly on how one could approach them. If a problem is phrased as a statement, then the implicit challenge is to prove this statement. Here is an approximate table of contents for the course.

Lecture 1	Bell inequalities and the Orsay experiment Notes §1
Lecture 2	States, events, observables and qubits Notes §2.1, §2.2, §2.3.1, §2.3.3
Lecture 3	Time evolution, Observables and time evolution for qubits Notes §2.4, §2.5.1, §2.5.2, §2.5.3
Lecture 4	Functional analysis of the Harmonic oscillator Notes §2.7
Lecture 5	Super dense coding and Tensor products NC §2.3, Notes §2.6
Lecture 6	Circuits and the quantum Fourier transform NC §4.1, §4.2, §4.3, §5.1
Lecture 7	Phase estimation, quantum order finding, and the factorization algorithm. NC §5.2, §5.3
Lecture 8	Symmetry and conserved quantities Notes §2.8 (under construction)
Lecture 9	Open systems Notes §3.1
Lecture 10	Transformations Notes §3.2
Lecture 11	Completely positive maps, collapse of the wave function Notes §3.3, 3.4
Lecture 12	GNS representations and Stinespring's Theorem Notes §3.5.1, 3.5.2
Lecture 13	Choi, Kadison-Schwarz, no-cloning theorem and quantum teleportation. Notes §3.5.3, 3.5.4, 3.5.5, and NC §1.3.7
Lecture 14	Measures of distance Notes §3.6
Lecture 15	Quick intro to semidefinite programming and quantum error correction Notes §3.7, and NC §10.1, 10.2

1 Introduction: a game and an experiment

We describe a game that cannot be won using classical strategies, but which *can* be won according to the rules of quantum mechanics.

The game was proposed by John Clauser, Michael Horne, Abner Shimony and Richard Holt in 1969, and converted into an experiment by John Clauser and Stuart Friedman in 1972. The version of the experiment we describe was performed by Alain Aspect, Jean Dalibart and Gérard Roger in Orsay in 1982. Together with Anton Zeilinger, Clauser and Aspect received the 2022 Nobel Prize of Physics for their work on this circle of ideas.

1.1 The CHSH game

The rules of the game are as follows. Alice and Bob are at opposite ends of a long table, with a referee in the middle. Before the game started, Alice and Bob met to agree on a strategy. But now they are separated from each other and from the referee by screens that prevent any form of communication.

The referee has four boxes, labelled Q_1, Q_2, R_1 and R_2. Each round, the following happens:

1) The referee puts either a black or a white marble in each of the four boxes. When the boxes are filled, he sends Q_1 and Q_2 to Alice, and R_1 and R_2 to Bob. The procedure for choosing the marbles can be either deterministic or randomized, but the same procedure is used in every round.

2) Alice and Bob are allowed to open *one* of the two boxes they receive. On a sheet of paper, they record $+1$ if they found a white marble and -1 if they found a black one. They also record which box they chose to open.

Then the next round begins. If the game ends after N rounds, the two lists hat Alice and Bob compiled may look as follows:

Round	Alice	Bob
1	$Q_1 = -1$	$R_2 = +1$
2	$Q_1 = +1$	$R_1 = +1$
3	$Q_2 = +1$	$R_1 = -1$
$\vdots$	$\vdots$	$\vdots$
N	$Q_1 = +1$	$R_2 = -1.$

The screens are lifted, and Alice and Bob compare their list of outcomes. They calculate the average value $\overline{Q_1R_1}$ of $Q_1R_1 \in \{\pm 1\}$ over all the instances where Alice chose to open Q_1 and Bob chose to open R_1. Similarly, they calculate $\overline{Q_1R_2}$, $\overline{Q_2R_1}$ and $\overline{Q_2R_2}$, and then determine the number

$$\overline{Q_1R_1} + \overline{Q_1R_2} + \overline{Q_2R_1} - \overline{Q_2R_2}. \tag{1}$$

The objective of Alice and Bob is to make this number strictly larger than 2.

Unfortunately for Alice and Bob, the CHSH-inequality (named after Clauser, Horne, Shimony and Holt) implies that they do not have a winning strategy, regardless of the way in which the referee fills the boxes.

Theorem 1.1 (CHSH-inequality)**.** *Let Q_1, Q_2, R_1 and R_2 be random variables on a probability space $(\Omega, \Sigma, \mathbb{P})$ that take only the values ± 1. Then*

$$|\mathbb{E}(Q_1R_1) + \mathbb{E}(Q_1R_2) + \mathbb{E}(Q_2R_1) - \mathbb{E}(Q_2R_2)| \leq 2.$$

Proof. Recall that

$$\mathbb{E}(Q_1R_1)+\mathbb{E}(Q_1R_2)+\mathbb{E}(Q_2R_1)-\mathbb{E}(Q_2R_2) = \mathbb{E}(Q_1R_1+Q_1R_2+Q_2R_1-Q_2R_2),$$

regardless whether the variables are independent or not. The random variable $Q_1R_1 + Q_1R_2 + Q_2R_1 - Q_2R_2 = Q_1(R_1 + R_2) + Q_2(R_1 - R_2)$ takes only the values ± 2. Indeed, if $R_1 = R_2$, then $R_1 - R_2 = 0$ and $Q_1(R_1 + R_2) = \pm 2$. Similarly, if $R_1 \neq R_2$, then $R_1 + R_2 = 0$ and $Q_2(R_1 - R_2) = \pm 2$. Since the values are ± 2, the expectation lies between 2 and -2. □

Although the average $\overline{Q_1R_1}$ is only calculated from the rounds in which Alice chose to open Q_1 and Bob chose to open R_1, these rounds form a *random sample* from the set of all boxes that the referee has prepared, because Alice and Bob do not exchange information with the referee.

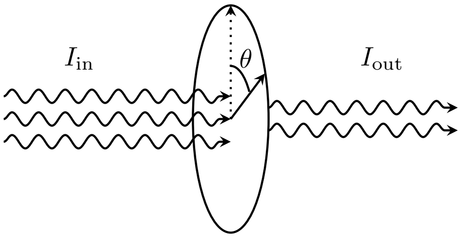
Since the referee fills the boxes according to the same (probabilistic) strategy in each round, the weak law of large numbers guarantees that the average $\overline{Q_1R_1}$ over the random sample of opened boxes converges to the expectation $\mathbb{E}(Q_1R_1)$, and, similarly, that $\overline{Q_1R_1} + \overline{Q_1R_2} + \overline{Q_2R_1} - \overline{Q_2R_2}$ converges to $\mathbb{E}(Q_1R_1) + \mathbb{E}(Q_1R_2) + \mathbb{E}(Q_2R_1) - \mathbb{E}(Q_2R_2)$.

Since this is at most 2 in absolute value by the CHSH-inequality, Alice and Bob will lose the game with arbitrarily high probability if N is large enough, regardless of the way in which the referee chooses to fill the boxes.

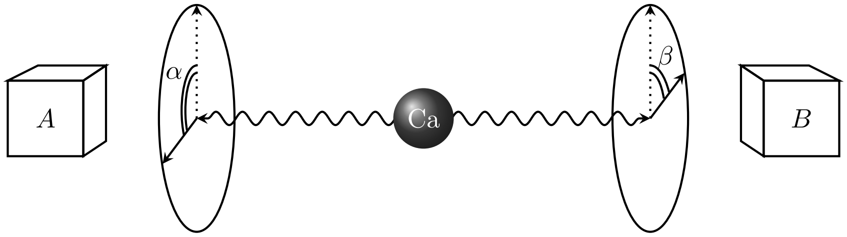
Problem 1.1. Suppose that Alice and Bob play the CHSH game. Bob cheats by looking at Alice's outcome before he chooses his box. Does there *exist* a strategy for the referee to fill the boxes in such a way that Alice and Bob can win the game? Can Alice and Bob win the game for *any* strategy that the referee may choose?

1.2 The Orsay experiment

A beam of light consists of photons, and each photon has a *polarization direction* perpendicular to the direction of the beam. If a polarization filter is placed in the beam of light, then the photons whose polarization is aligned with that of the filter can pass unobstructed, whereas those photons whose polarization is perpendicular to that of the beam are blocked. If the polarization of the light beam makes an angle θ with the polarization of the filter, then the light beam that comes out of the filter has an intensity $I_{\text{out}} = \cos^2(\theta) I_{\text{in}}$.



The following experiment was performed by Alain Aspect in Orsay in 1982. In the middle of a long table there is a calcium atom, which has been prepared in an excited state by means of a laser beam. When the calcium atom decays to its ground state, it emits two photons in opposite directions. These photons have opposite polarization. At one end of the table, we place a polarization filter, with a polarization direction that makes an angle α with the z-axis. Detector A measures whether or not the emitted photon passes through the filter. At the other end of the table, we place a filter with a polarization direction that makes an angle β with the z-axis. Detector B measures whether or not the photon passes through that filter.



In fact, the two filters can be spun around their axis *while the photon is in flight*, resulting in a truly random choice of α and β. Running this experiment many times, it was found that the probability that both photons pass is $\frac{1}{2}\sin^2(\alpha-\beta)$, and the probability that both photons are blocked is also $\frac{1}{2}\sin^2(\alpha-\beta)$.

We define $Q(\alpha)$ to be ± 1 according to whether or not a photon is detected at A, and $R(\beta) = \pm 1$ according to whether a photon is detected at B. Then the results of the Orsay experiment can then be summarized as follows:

$$\mathbb{P}[Q(\alpha) = R(\beta)] = \sin^2(\alpha-\beta). \tag{2}$$

At first sight formula (2) looks rather harmless. It is consistent with the classical input-output law $I_{\text{out}} = \cos^2(\theta) I_{\text{in}}$, with the cosine replaced by a sine because the two photons have opposite (rather than identical) polarization. An important difference is that formula (2) does not pertain to the *intensity* of a beam of light, but to the *probability* that single photons are detected. This turns out to have rather profound ramifications.

1.3 A winning strategy for Alice and Bob

Unexpectedly, the Orsay experiment yields a version of the game from §1.1 where Alice and Bob *do* have a winning strategy!

The calcium atom plays the role of the referee, and sends Alice and Bob one photon each. Alice aligns her polarization filter either in the direction α_1 or α_2. If she chooses the direction α_i, she writes down $Q(\alpha_i) = +1$ if she detects that the photon has passed through the filter, and $Q(\alpha_i) = -1$ if it was blocked. In the same vein, Bob aligns his polarization filter either in the direction β_1 or β_2, and writes down $R(\beta_i) = \pm 1$ according to whether or not the photon has passed through his filter. Since $Q(\alpha)$ and $R(\beta)$ have the same sign with probability $\sin^2(\alpha - \beta)$, and opposite signs with probability $\cos^2(\alpha - \beta)$, we have

$$\mathbb{E}\big(Q(\alpha)R(\beta)\big) = \sin^2(\alpha - \beta) - \cos^2(\alpha - \beta). \tag{3}$$

So the quantity

$$\Delta := \mathbb{E}(Q(\alpha_1)R(\beta_1)) + \mathbb{E}(Q(\alpha_1)R(\beta_2)) + \mathbb{E}(Q(\alpha_2)R(\beta_1)) - \mathbb{E}(Q(\alpha_2)R(\beta_2)) \tag{4}$$

from Theorem 1.1 takes the value

$$\Delta = 2 + 2\Big(\cos^2(\alpha_2 - \beta_2) - \cos^2(\alpha_1 - \beta_1) - \cos^2(\alpha_1 - \beta_2) - \cos^2(\alpha_2 - \beta_1)\Big). \tag{5}$$

If Alice and Bob choose to align their filters as

Alice	Bob
$\alpha_1 = \frac{1}{3}\pi$	$\beta_1 = \frac{2}{3}\pi$
$\alpha_2 = \pi$	$\beta_2 = 0,$

then $\Delta = 2 + 2(1 - \frac{1}{4} - \frac{1}{4} - \frac{1}{4}) = 2\frac{1}{2}$, violating the CHSH-inequality. Apparently, if the referee from the CHSH-game is replaced by a calcium atom, then Alice and Bob do have winning strategy.

1.4 Interpretation of the experiment

Since the Orsay experiment violates the CHSH-inequality, we must conclude that it can *not* be described by classical probability theory, at least not along the lines of the game described in §1.1. At least one of the following two assumptions in the CHSH game must be violated:

1) The assumption that the polarizations $Q(\alpha_1)$, $Q(\alpha_2)$ of Alice's photon and the polarizations $R(\beta_1)$ and $R(\beta_2)$ of Bob's photon have definite values ± 1, which exist regardless whether they are measured or not. This assumption is called *realism*

2) The assumption that the measurements made by Alice and Bob do not influence each other, and that their choice of polarization direction is not influenced by the calcium atom. This assumption is called *locality.*

Although it may seem attractive to drop locality, this is not the majority point of view in the physics community. To appreciate the reluctance to abandon locality, it is good to realize that in order for Alice's measurement to influence Bob's outcomes, information would have to travel *faster than light.* In view of special relativity, this is equivalent to information *travelling back in time*, which opens up a Pandora's Box of paradoxes in its own right. Perhaps more decisively, serious attempts to create a theory of Nature in which locality is violated have led to constructions that, although logically consistent, are widely perceived as unnecessarily complicated and not very enlightening.

The majority view, then, is to drop the assumption of realism. One then abandons the picture where each of the two photons carries a little list of 'correct answers', each of which can be read off by a different type of measurement. Instead, one views a measurement outcome as something that arises in the interaction between photon, polarization filter and detector. If Alice chooses to align the detector in the direction α_1, then for this particular photon, the information about the direction α_2 is forever lost. In fact, in view of the CHSH inequality, one may ask if this information was ever really there in the first place.

1.5 Quantum mechanics

At the time of the experiment (1982), the philosophical discussion about locality versus realism was not at all new. The behaviour of single photons is governed by *quantum mechanics.* Starting with the pioneering work of Planck (1900), Einstein (1905) and Bohr (1913), quantum mechanics developed into a comprehensive theory of nature with Heisenberg's matrix mechanics (1925) and Schrödinger's wave equation (1926). It found its definitive mathematical formulation in the work of Dirac (1930) and von Neumann (1932), and has remained a cornerstone of modern physics, confirmed by every single experiment performed to date.

The formula (2) and the resulting violation of the CHSH inequality are precisely what quantum mechanics predicts, and it is this prediction that inspired the experiment. It was immediately clear to its inventors in the early 20^{th} century that quantum *theory* violates either locality and/or realism. The reason that Clauser and Aspect received the 2022 Nobel Prize of Physics is that their experiments show that *nature* violates locality and/or realism, independent of its description by quantum theory.

1.6 Outlook

The somewhat counterintuitive nature of quantum mechanics is therefore not an accidental property of a theory that describes Nature, but a fundamental property of Nature itself.

In the remainder of these notes we will leave aside the philosophical question of what this *means*, and focus on the more practical question of *what one can do with this.* Apparently, we can use quantum mechanics to win a game that cannot be won classically. Admittedly the CHSH game may not be very interesting, but can we use quantum mechanics to win some games that we *do* care about?

To answer this question, we first have to build a solid understanding of quantum theory, with an emphasis on its counterintuitive facets.

2 Quantum mechanics of isolated systems

We first describe the mathematical framework for *isolated* quantum systems, wich evolve without any interaction with their environment. Because quantum information and quantum computing is ultimately about manipulating quantum systems, we will move to *open* systems as soon as possible.

The mathematical framework of quantum mechanics is surprisingly simple because of the essentially linear nature of the theory – for isolated as well as for open quantum systems. To underline this point, we first describe quantum systems with finitely many degrees of freedom using only linear algebra. After this, we will comment on the functional analytic refinements that are needed to describe systems with infinitely many degrees of freedom.

2.1 The postulates of quantum mechanics

The basic postulates of quantum mechanics tell us which type of mathematical structure should be used to model:

1) The *state* of a quantum system[1].

2) The *observables* of a quantum system.

3) The *time evolution* of a quantum system.

The state captures all the information about the future behaviour of the system. The observables are the properties of the system that one can measure, and the time evolution tells us how the state of the system evolves in time.

2.2 First postulate: states

The state space of an isolated system is described by a *Hilbert space*.

Definition 2.1. A finite dimensional Hilbert space is a finite dimensional complex vector space $\mathcal{H}$ equipped with an inner product $\langle \cdot, \cdot \rangle \colon \mathcal{H} \times \mathcal{H} \to \mathbb{C}$.

We adopt the physics convention that the inner product is linear in the *right* argument. So for all $\psi, \chi, \chi_1, \chi_2 \in \mathcal{H}$ and $\alpha, \beta \in \mathbb{C}$, we have

$$\begin{aligned} \langle \chi, \psi \rangle &= \overline{\langle \psi, \chi \rangle}, \\ \langle \psi, \alpha\chi_1 + \beta\chi_2 \rangle &= \alpha\langle \psi, \chi_1 \rangle + \beta\langle \psi, \chi_2 \rangle, \text{ and} \\ \langle \psi, \psi \rangle &\geq 0. \end{aligned}$$

[1] De nederlandse vertaling van *state* is *toestand*.

Further, $\langle \psi, \psi \rangle = 0$ if and only if $\psi = 0$. This allows us to define the *norm*

$$\|\psi\| := \sqrt{\langle \psi, \psi \rangle} \tag{6}$$

of a vector ψ. A *unit vector* is an element $\psi \in \mathcal{H}$ with unit norm, $\|\psi\| = 1$.

Example 2.1. The basic example of a finite dimensional Hilbert space is $\mathcal{H} = \mathbb{C}^n$, where the inner product between $\psi = (\psi_1, \ldots, \psi_n)$ and $\chi = (\chi_1, \ldots, \chi_n)$ is given by

$$\langle \psi, \chi \rangle = \overline{\psi}_1 \chi_1 + \ldots + \overline{\psi}_n \chi_n.$$

Problem 2.1. Check that this is indeed a Hilbert space.

Problem 2.2. The inner product on a Hilbert space $\mathcal{H}$ is conjugate linear in the first argument, $\langle \alpha\psi_1 + \beta\psi_2, \chi \rangle = \overline{\alpha}\langle \psi_1, \chi \rangle + \overline{\beta}\langle \psi_2, \chi \rangle$ for all $\psi_1, \psi_2, \chi \in \mathcal{H}$ and $\alpha, \beta \in \mathbb{C}$.

We also adopt the physics notation $A^\dagger$ for the adjoint of an operator A.

Definition 2.2. The *adjoint* of a linear map $A\colon \mathcal{H} \to \mathcal{K}$ from a Hilbert space $\mathcal{H}$ to a Hilbert space $\mathcal{K}$ is the unique linear map $A^\dagger\colon \mathcal{K} \to \mathcal{H}$ that satisfies $\langle \psi, A\chi \rangle = \langle A^\dagger \psi, \chi \rangle$ for all $\psi \in \mathcal{K}$ and $\chi \in \mathcal{H}$.

Problem 2.3. If $A\colon \mathcal{H} \to \mathcal{K}$ has coefficients a_{ij} with respect to an orthonormal basis $e_1, \ldots, e_n$ of $\mathcal{H}$ and $f_1, \ldots, f_m$ of $\mathcal{K}$, then $A^\dagger$ has coefficients $\overline{a}_{ji}$.

Every vector $\psi \in \mathcal{H}$ gives rise to a linear map $\mathcal{H} \to \mathbb{C}$ defined by $\chi \mapsto \langle \psi, \chi \rangle$. This linear map $\mathcal{H} \to \mathbb{C}$ is denoted by $\langle \psi |$ and called a *bra.* If we think of $\chi \in \mathcal{H}$ simply as a vector, we denote it by $|\chi\rangle$ and call it a *ket.* The logic of this *Dirac notation* is that $\langle \psi | \chi \rangle$, the bra $\langle \psi |$ applied to the ket $|\chi\rangle$, is the same as $\langle \psi, \chi \rangle$, the inner product of ψ and χ. We can now make sense of expressions like

$$|\chi\rangle\langle\psi|$$

by simply concatenating the symbols. For example, $|\chi\rangle\langle\psi|$ is the linear map $\mathcal{H} \to \mathcal{H}$ that sends ξ to $\langle \psi, \xi \rangle \chi$.

Problem 2.4. Let e_i be an orthonormal basis of a finite dimensional Hilbert space $\mathcal{H}$. Then

$$\sum_{i=1}^{n} |e_i\rangle\langle e_i| = \mathrm{Id}_{\mathcal{H}}. \tag{7}$$

Problem 2.5. Show that $(|\psi\rangle\langle\chi|)^\dagger = |\chi\rangle\langle\psi|$.

Physically, the *state* of a quantum system describes everything there is to know about this system. We can now state the first postulate of quantum mechanics.

Postulate 1

The state space of a quantum system is modelled by a Hilbert space $\mathcal{H}$. Every state of this system is described by a unit vector $\psi \in \mathcal{H}$.

The dimension of the Hilbert space corresponds to the number of degrees of freedom of the quantum system. For the moment, we will focus on systems with finitely many degrees of freedom because 1) they are at the centre of the theory of quantum information and computation and 2) they are easier to handle. Nonetheless, systems with infinitely many degrees of freedom are ubiquitous in physics, so we will come back to infinite dimensional Hilbert spaces later on.

2.3 Second postulate: events and observables

The second postulate of quantum mechanics states, briefly, that events are modelled by orthogonal projections $P \colon \mathcal{H} \to \mathcal{H}$. We describe this postulate in more detail in §2.3.1. In §2.3.2, we use this postulate to refine our understanding of the physical states of a closed system, and show that they are described by a *projective* Hilbert space. In §2.3.3, we show that *observables* – the analogues of random variables in quantum theory – should be described by Hermitian operators.

2.3.1 Events

Events are modelled by orthogonal projections $P \colon \mathcal{H} \to \mathcal{H}$.

Definition 2.3. An orthogonal projection is a linear map $P \colon \mathcal{H} \to \mathcal{H}$ that satisfies $P^2 = P$ and $P^\dagger = P$.

Orthogonal projections P on a finite dimensional Hilbert space $\mathcal{H}$ correspond bijectively with linear subspaces $V \subseteq \mathcal{H}$. Indeed, a projection P gives rise to an orthogonal decomposition $\mathcal{H} = V \oplus V^\perp$ into the image $V = \mathrm{Im}(P)$ and the kernel $V^\perp = \mathrm{Ker}(P)$. Conversely, given a linear subspace $V \subseteq \mathcal{H}$, the unique linear map $P_V \colon \mathcal{H} \to \mathcal{H}$ that satisfies $P_V \psi = \psi$ for $\psi \in V$ and $P\psi = 0$ for $\psi \perp V$ is an orthogonal projection. It is called the *projection onto* V.

Problem 2.6. Let $P \colon \mathcal{H} \to \mathcal{H}$ be linear. If $P^\dagger = P$, then $\mathrm{Ker}(P) \perp \mathrm{Im}(P)$. If $P^2 = P$, then $\mathrm{Ker}(P) = \mathrm{Im}(\mathbf{1} - P)$. Conclude that if P is an orthogonal projection, then $\mathcal{H} = V \oplus V^\perp$ with $V = \mathrm{Im}(P)$ and $V^\perp = \mathrm{Ker}(P)$.

We postulate that for a system in state $\psi \in \mathcal{H}$, the probability that the event P occurs is $\langle \psi, P\psi \rangle$. This is always a real number between 0 and 1.

Problem 2.7. If P is a projection, then so is $\mathbf{1} - P$. Since $0 \leq \langle \psi, P\psi \rangle$ and $0 \leq \langle \psi, (\mathbf{1} - P)\psi \rangle$, we have $0 \leq \langle \psi, P\psi \rangle \leq 1$ for every orthogonal projection.

Since an orthogonal projection P is uniquely determined by its image V, we can define the operations $P_V \vee P_W := P_{V+W}$ and $P_V \wedge P_W := P_{V \cap W}$. We say that P and Q *commute* if $PQ = QP$. More generally, the failure of two operators to commute is measured by their *commutator*

$$[P, Q] := PQ - QP. \tag{8}$$

For commuting projections, the operations $P \wedge Q$ and $P \vee Q$ can be expressed as follows.

Problem 2.8. Let P and Q be projections onto V and W, respectively.

a) The linear map PQ is a projection if and only if P and Q commute.

b) If so, then $P \wedge Q = PQ$ is the projection onto $V \cap W$.

c) The linear map $P+Q-PQ$ is a projection if and only if P and Q commute.

d) If so, then $P \vee Q = P + Q - PQ$ is the projection onto $V + W$.

We can now state the second postulate of quantum mechanics.

Postulate 2

Events are modelled by orthogonal projections. If the event P is measured for a system in state ψ, then the probability that P occurs is $\langle \psi, P\psi \rangle$. Two projections P and Q can be simultaneously measured if and only if they commute. If this is the case, then $P \wedge Q$ is interpreted as the event "P and Q occur", and $P \vee Q$ as the event "P or Q occurs".

This postulate reveals that quantum mechanics is at its core a *probabilistic* theory. It predicts the probability that an event P occurs for a system in state ψ, but it generally does not (and, as we will see, *cannot*) predict the outcome of a single experiment. Predictions of this probabilistic type can be tested by performing the same measurement in a large ensemble of systems, all of which have been prepared in the same state ψ.

Problem 2.9. Let $P = |\psi\rangle\langle\psi|$ and $Q = |\chi\rangle\langle\chi|$ for two non-orthogonal vectors ψ and χ in $\mathcal{H} = \mathbb{C}^2$.

a) Show that P and Q do not commute.

b) What goes wrong if we still try to interpret $P \vee Q$ as the event "P or Q occurs", and $P \wedge Q$ as the event "P and Q occurs"?

2.3.2 Projective space

As a first consequence of this postulate, we see that it is impossible to distinguish a closed system in state $\psi \in \mathcal{H}$ from one in state $e^{i\phi}\psi$. Indeed, since the probability of occurrence for *any* possible event P is the same in these two states,

$$\langle \psi, P\psi \rangle = \langle e^{i\phi}\psi, Pe^{i\phi}\psi \rangle,$$

there are no experiments that could conceivably tell these states apart. So although every possible state of the system corresponds to a unit vector in $\mathcal{H}$, two such vectors describe the *same* state if and only if they differ by a complex number $e^{i\phi}$ of modulus 1.

On the set of unit vectors $\psi \in \mathcal{H}$, we therefore define the equivalence relation

$$\psi \sim \chi \quad \text{if} \quad \chi = e^{i\phi}\psi \text{ for some phase } \phi \in \mathbb{R}.$$

The equivalence classes

$$[\psi] = \{e^{i\phi}\psi\,;\, \phi \in \mathbb{R}\}$$

corresponding to physical states are called *rays*, and the set

$$\mathcal{P}(\mathcal{H}) = \{[\psi]\,;\, \psi \in \mathcal{H}\,;\, \|\psi\| = 1\}.$$

of rays is called the *projective space.* We therefore obtain the following refinement of the first postulate.

Physical states are rays in $\mathcal{P}(\mathcal{H})$

> *Physical states of a closed quantum system correspond bijectively to rays $[\psi]$ in the projective Hilbert space $\mathcal{P}(\mathcal{H})$.*

Remark 2.1. In the mathematics literature, the projective space $\mathcal{P}(\mathbb{C}^n)$ is often denoted $\mathbb{C}\mathrm{P}^{n-1}$. This is because in complex (or algebraic) geometry, it is considered as a complex manifold (or smooth algebraic variety) of dimension $n-1$.

2.3.3 Observables

Starting from the above model for events, we now develop the quantummechanical analogues of probability measures and random variables: projection valued measures (PVMs) and observables.

Definition 2.4 (Finite PVMs)**.** A *projection valued measure* (PVM) on a finite set Ω is a collection $\{P_\omega\,;\, \omega \in \Omega\}$ of projections that satisfy $P_\omega P_{\omega'} = 0$ for $\omega \neq \omega'$ and $\sum_{\omega\in\Omega} P_\omega = \mathbf{1}_{\mathcal{H}}$

The projections P_ω are simultaneously measurable (since $P_\omega P_{\omega'}$ and $P_{\omega'} P_\omega$ are both zero), and every measurement yields a single outcome ω. Indeed, the probability that two different outcomes $\omega \neq \omega'$ both occur is $\langle \psi, P_\omega P_{\omega'}\psi \rangle = 0$, and the probability that at least one outcome occurs is $\langle \psi, \sum_{\omega\in\Omega} P_\omega \psi \rangle = 1$.

Problem 2.10 (PVMs and direct sum decompositions)**.** If $\{P_\omega\,;\,\omega \in \Omega\}$ is a PVM with $V_\omega = \mathrm{Im}(P_\omega)$, then $\mathcal{H} = \bigoplus_{\omega\in\Omega} V_\omega$ is an orthogonal direct sum decomposition. Conversely, every such decomposition gives rise to a PVM.

Given a projection valued measure $\{P_\omega\,;\,\omega \in \Omega\}$ and a state $\psi \in \mathcal{H}$, we obtain a *probability density* on Ω,

$$p_\psi(\omega) := \langle \psi, P_\omega \psi \rangle. \tag{9}$$

Note that $0 \leq p_\omega \leq 1$ for every $\omega \in \Omega$, and $\sum_{\omega\in\Omega} p_\omega = 1$.

A *random variable* on Ω is a function $a\colon \Omega \to \mathbb{R}$. With respect to the (classical) probability density p_ψ, it has (classical) expectation

$$\mathbb{E}_\psi(a) = \sum_{\omega\in\Omega} a(\omega)p_\psi(\omega). \tag{10}$$

Note that for the Hermitian operator $A\colon \mathcal{H} \to \mathcal{H}$ defined by

$$A := \sum_{\omega} a(\omega)P_\omega, \tag{11}$$

we have

$$\mathbb{E}_\psi(a) = \langle \psi, A\psi \rangle. \tag{12}$$

The operator A is called the *observable* associated to the PVM $\{P_\omega\,;\,\omega \in \Omega\}$ and the random variable a.

In fact, *any* Hermitian operator $A\colon \mathcal{H} \to \mathcal{H}$ is of the form (11) for some PVM. This follows from the spectral theorem. Recall that $a \in \mathbb{C}$ is an eigenvalue of the linear map $A\colon \mathcal{H} \to \mathcal{H}$ if the eigenspace

$$V_a := \{\psi \in \mathcal{H}\,;\, A\psi = a\psi\}$$

is nonzero. The orthogonal projection P_a onto V_a is called a *spectral projection*, and the set of eigenvalues is called the *spectrum* of A, denoted $\mathrm{spec}(A)$. If A is Hermitian, then every eigenvalue is real, and $V_a \perp V_b$ if $a \neq b$.

Problem 2.11. Both of these statements follow from $\langle \psi, A\chi \rangle = \langle A\psi, \chi \rangle$ applied to unit vectors $\psi \in V_a$ and $\chi \in V_b$, first with $\psi = \chi$ and then with $a \neq b$.

Theorem 2.1 (Spectral theorem)**.** *If $A\colon \mathcal{H} \to \mathcal{H}$ is a Hermitian operator, then*

$$\mathcal{H} = \bigoplus_{a\in\mathrm{spec}(A)} V_a$$

is an orthogonal direct sum of eigenspaces.

This allows us to express the Hermitian operator as $A = \sum_{\mathrm{spec}(A)} aP_a$. Indeed, the left and the right hand side agree on V_a for every $a \in \mathrm{spec}(A)$, and therefore on the direct sum $\mathcal{H} = \bigoplus_{a\in\mathrm{spec}(A)} V_a$ by linearity. Since the spectral projections $\{P_a\,;\,a \in \mathrm{spec}(A)\}$ form a PVM by Problem 2.10, we obtain the following equivalent formulation of the spectral theorem.

Theorem 2.2 (Spectral theorem)**.** *Let $A\colon \mathcal{H} \to \mathcal{H}$ be a Hermitian operator. Then $\{P_a\,;\, a \in \operatorname{spec}(A)\}$ is a PVM, and*

$$A = \sum_{a \in \operatorname{spec}(A)} a P_a. \tag{13}$$

Choosing an orthonormal basis for every eigenspace V_a and combining them into an orthonormal basis of $\mathcal{H}$, we obtain a basis $\{\psi_i\,;\, i = 1, \ldots, n\}$ of eigenvectors, $A\psi_i = a_i\psi_i$. This yields an alternative spectral decomposition

$$A = \sum_{i=1}^{n} a_i \,|\psi_i\rangle\langle\psi_i|$$

into projections $P_i = |\psi_i\rangle\langle\psi_i|$ of rank 1, where the a_i are not necessarily distinct.

Proof of Theorem 2.1. We proceed by induction on $n = \dim(\mathcal{H})$. The case $n = 0$ is clear, so suppose that $n \neq 0$ and that the statement holds for all Hilbert spaces of dimension less than n. Because the characteristic polynomial $p_A(a) = \det(a\mathbf{1} - A)$ has at least one root $a \in \mathbb{C}$, the operator A has at least one eigenspace V_a. If $\psi \in V_a$ and $\chi \in V_a^{\perp}$, then $\langle\psi, A\chi\rangle = \langle A\psi, \chi\rangle = \overline{a}\langle\psi, \chi\rangle = 0$, so A maps V_a to V_a and $V_a^{\perp}$ to $V_a^{\perp}$. Since $\dim(V_a^{\perp}) < n$, the induction hypothesis yields an eigenspace decomposition of $V_a^{\perp}$ into eigenspaces for $A|_{V_a^{\perp}} : V_a^{\perp} \to V_a^{\perp}$. Since $\mathcal{H} = V_a \oplus V_a^{\perp}$, this results in an eigenspace decomposition for $\mathcal{H}$. □

So far, one might get the impression that quantum mechanics is secretly just classical probability theory on $\Omega = \operatorname{spec}(A)$. As long as one considers only commuting observables, there is indeed much to be said for this point of view. But what makes quantum mechanics fundamentally different is that two observables A and B can *not* be represented on the same classical probability space if they do not commute.

Proposition 2.3. *Two operators A and B commute if and only if all their spectral projections P_a and P_b commute.*

Proof. If the spectral projections commute, then clearly A and B commute, since they are linear combinations of spectral projections. So suppose that A and B commute. If V_a is an eigenspace of A and $\psi \in V_a$, then $B\psi \in V_a$ as well, since $A(B\psi_a) = B(A\psi_a) = a(B\psi_a)$. So $BV_a \subseteq V_a$, and $BP_a = P_aBP_a$. Since the right hand side is Hermitian, the left hand side is Hermitian as well, $BP_a = (BP_a)^{\dagger}$. So $BP_a = P_aB$, and B commutes with all spectral projections P_a of A. Applying this to the commuting operators P_a and B (instead of B and A), we conclude that P_a commutes with the spectral projections P_b of B. □

So if $[A, B] = 0$, then the product P_aP_b of a spectral projection for A and one for B is again a projection operator.

Problem 2.12. Verify that $\{P_aP_b\,;\, (a, b) \in \operatorname{spec}(A) \times \operatorname{spec}(B)\}$ is a PVM, and conclude that two Hermitian operators A and B can be expressed with respect to the *same* PVM if and only if they commute.

Combined with the spectral theorem, this leads us to the following interpretation of Hermitian operators:

Observables

> *Observables are modelled by Hermitian operators* $A\colon \mathcal{H} \to \mathcal{H}$*. If the system is in state* $\psi \in \mathcal{H}$*, then a measurement of* A *will yield outcome* $a \in \operatorname{spec}(A)$ *with probability* $\mathbb{P}_\psi(a) = \langle \psi, P_a \psi \rangle$*. Two observables* A *and* B *can be simultaneously measured if and only if they commute.*

Remark 2.2. Note that using *finite dimensional* Hilbert spaces, we can only model observables with a *discrete* spectrum of possible measurement outcomes.

If the expectation of $[A, B]$ in the state ψ is nonzero, then this places a lower bound on the variance of A and B. This is known as the *Heisenberg uncertainty relation*. In the form below, it is due to Robertson.

Problem 2.13. Show that if $A = \sum_{\omega \in \Omega} a(\omega) P_\omega$, then $A^2 = \sum_{\omega \in \Omega} a^2(\omega) P_\omega$. Conclude that the *variance* $\mathbf{Var}_\psi(a) := \mathbb{E}_\psi(a^2) - \mathbb{E}_\psi(a)^2$ of the random variable a with respect to the probability density $\mathbb{P}_\psi(\omega) = \langle \psi, P_\omega \psi \rangle$ can be expressed as

$$\mathbf{Var}_\psi(a) = \langle \psi, A^2 \psi \rangle - \langle \psi, A\psi \rangle^2.$$

Problem 2.14 (Heisenberg uncertainty relation)**.** Let A and B be Hermitian operators on $\mathcal{H}$, and let $\psi \in \mathcal{H}$ be a unit vector. Derive the uncertainty relation

$$\mathbf{Var}_\psi(a)\mathbf{Var}_\psi(b) \geq |\tfrac{1}{2i}\langle \psi, [A, B]\psi \rangle|^2. \tag{14}$$

a) Set $\langle A \rangle := \langle \psi, A\psi \rangle$, and show that $\mathbf{Var}_\psi(a) = \|(A - \langle A \rangle \mathbf{1})\psi\|^2$.

b) Conclude that $\mathbf{Var}_\psi(a)\mathbf{Var}_\psi(b) \geq |\langle (A - \langle A \rangle \mathbf{1})\psi, (B - \langle B \rangle \mathbf{1})\psi \rangle|^2$.

c) The right hand side satisfies $\operatorname{Im}\langle (A - \langle A \rangle \mathbf{1})\psi, (B - \langle B \rangle \mathbf{1})\psi \rangle = \frac{1}{2i}\langle \psi, [A, B]\psi \rangle$.

d) Derive the Heisenberg uncertainty relation (14).

2.4 Third postulate: time evolution

Transformations of a closed quantum system are modelled by unitary operators $U\colon \mathcal{H} \to \mathcal{H}$.

Definition 2.5. A unitary operator is an invertible linear map $U\colon \mathcal{H} \to \mathcal{H}$ that respects the inner product, $\langle U\psi, U\chi \rangle = \langle \psi, \chi \rangle$ for all $\psi, \chi \in \mathcal{H}$. Equivalently, U is unitary if and only if $U^\dagger U = U U^\dagger = \mathbf{1}$.

Time evolution on a closed system is modelled by a continuous 1-parameter group of unitary operators.

Definition 2.6. A *continuous 1-parameter group of unitary operators* is a family $U_t\colon \mathcal{H} \to \mathcal{H}$ of unitary operators, indexed by $t \in \mathbb{R}$, that satisfies:

i) $U_0 = \mathbf{1}$.

ii) $U_{t+s} = U_s U_t$ for all $t \in \mathbb{R}$.

iii) For every $\psi \in \mathcal{H}$, the map $\psi_t := U_t\psi$ is continuous in t.

The first property says that time evolution from time 0 to time 0 is trivial, and the third property means that the state of the system depends continuously on time. The second property is a *Markov condition*: if time evolution takes a state $\psi_0 = \psi$ at time 0 to a state $\psi_t = U_t\psi$ at time t, then the state ψ_t at time t contains all the relevant information for what happens afterwards. So to determine the state $\psi_{t+s} = U_{t+s}\psi$ at time $t + s$, we can run the time evolution up to time t, pretend that the system is born again at time 0 in state ψ_t, and run the time evolution up to time s with initial condition ψ_t. Note that the second property also implies that the time evolution is invertible, $U_{-t} = U_t^{-1}$.

Postulate 3

The time evolution of a closed system is modelled by a continuous 1-parameter group $\{U_t\,;\, t \in \mathbb{R}\}$ of unitary operators.

There are two equivalent ways to implement the unitary transformations. In the *Schrödinger picture*, the states $|\psi\rangle$ evolve and the observables A are fixed,

$$\begin{aligned} |\psi\rangle &\to |\psi\rangle_t = U_t\,|\psi\rangle \\ A &\to A. \end{aligned}$$

For a system initially in state $|\psi\rangle$, the observable initially given by A then has expectation $\langle U_t\psi, AU_t\psi\rangle$ at time t.

In the *Heisenberg picture*, the states $|\psi\rangle$ are fixed and the observables A evolve in time,

$$\begin{aligned} |\psi\rangle &\to |\psi\rangle \\ A &\to U_t^{-1}AU_t. \end{aligned}$$

For a system initially in state $|\psi\rangle$, the observable initially given by A then has expectation $\langle \psi, U_t^{-1}AU_t\psi\rangle$ at time t.

The Schrödinger picture and the Heisenberg picture are equivalent: since U_t is unitary, we have

$$\langle \psi, U_t^{-1}PU_t\psi\rangle = \langle U_t\psi, PU_t\psi\rangle$$

for every projection P, so the probability that P occurs for a system initially in state ψ is the same in both pictures.

Remark 2.3 (Projective unitary group)*.* Since a physical state of a closed system is only determined up to a phase, $\psi \sim e^{i\phi}\psi$ for all $\phi \in [0, 2\pi]$, we can identify unitary transformations if they agree up to a phase, $U \sim e^{i\phi}U$. The group of physical transformations of a closed system is therefore not the group $\mathrm{U}(\mathcal{H})$ of unitary operators on $\mathcal{H}$, but the *projective unitary group* $\mathrm{PU}(\mathcal{H}) = \mathrm{U}(\mathcal{H})/\mathbb{T}$. This is the quotient of $\mathrm{U}(\mathcal{H})$ by the normal subgroup $\mathbb{T} = \{e^{i\phi}\mathbf{1}\,;\, \phi \in [0, 2\pi]\}$ of unitary operators that act by a phase factor.

In order to further investigate continuous 1-parameter groups of unitary operators, we will need a *functional calculus* for normal operators.

Definition 2.7. An operator $A\colon \mathcal{H} \to \mathcal{H}$ is *normal* if $[A, A^\dagger] = 0$.

Proposition 2.4. *The following are equivalent:*

i) *A is normal.*

ii) *$A = X + iY$ for Hermitian operators X and Y with $[X, Y] = 0$.*

iii) *A has a spectral decomposition with complex eigenvalues:*

$$A = \sum_{a \in \mathrm{spec}(A)} aP_a. \tag{15}$$

Proof. For i) ⇔ ii), take $X = \frac{1}{2}(A + A^\dagger)$ and $Y = \frac{1}{2i}(A - A^\dagger)$. For ii) ⇔ iii), note that if $A = X + iY$ is normal, then by Problem 2.12 the commuting Hermitian operators X and Y can be expressed as

$$X = \sum_{\omega \in \Omega} x(\omega) P_\omega \quad \text{and} \quad Y = \sum_{\omega \in \Omega} y(\omega) P_\omega$$

for the *same* PVM $\{P_\omega\,;\, \omega \in \Omega\}$. So $A = \sum_{\omega \in \Omega}(x(\omega) + iy(\omega))P_\omega$ has the desired spectral decomposition. Conversely, every operator with such a spectral decomposition is normal, as $A^\dagger = \sum_{a \in \mathrm{spec}(A)} \overline{a} P_a$ commutes with A. □

We will need the following *function calculus* for normal operators $A\colon \mathcal{H} \to \mathcal{H}$.

Definition 2.8. For $f\colon \mathrm{spec}(A) \to \mathbb{C}$, we define $f(A) := \sum_{a \in \mathrm{spec}(A)} f(a) P_a$.

Problem 2.15. The function calculus respects addition and multiplication, $(f + g)(A) = f(A) + g(A)$ and $(f \cdot g)(A) = f(A)g(A)$. For a polynomial $p(z) = a_0 + a_1 z + \ldots + a_n z^n$, we have $p(A) = c_0 \mathbf{1} + c_1 A + \ldots + c_n A^n$.

The following problem provides a source of continuous 1-parameter groups of unitary operators.

Problem 2.16. If A is Hermitian, then $U_t := \exp(-itA)$ is a continuous 1-parameter group of unitary operators.

In fact, *every* continuous 1-parameter group of unitary operators is of this form.

Theorem 2.5 (Stone's theorem, finite dimensional version)**.** *For every continuous 1-parameter group U_t of unitary operators, there exists a Hermitian operator A such that $U_t = \exp(-itA)$.*

A variant of this theorem remains true for infinite dimensional Hilbert spaces. In this context, the result is due to Marshall Stone in 1932.

Proof. Since $U_tU_s = U_{t+s} = U_sU_t$, we have $[U_t, U_s] = 0$ for all $s, t \in \mathbb{R}$. It follows that all U_t can be simultaneously diagonalized,

$$U_t = \sum_{\omega \in \Omega} z_\omega(t) P_\omega \tag{16}$$

for a PVM that does not depend on t. For each $\omega \in \Omega$, the map $t \mapsto z_\omega(t)$ is a *continuous* map from $\mathbb{R}$ to the unit circle $\mathrm{U}(1) = \{z \in \mathbb{C}\,;\, |z| = 1\}$. Note that since $U_{kt} = U_t^k$, we have $z_\omega(kt) = z_\omega^k(t)$ for all $k \in \mathbb{Z}$. Indeed, U_t applied to a nonzero vector $\psi_\omega \in \mathrm{Im}(P_\omega)$ yields $U_t\psi_\omega = z_\omega(t)\psi_\omega$, so $U_t^k\psi_\omega = z_\omega^k(t)\psi_\omega$, whereas $U_{kt}\psi_\omega = z_\omega(kt)$. In particular, we have $z_\omega(t) = z_\omega^2(t/2)$ for all $t \in \mathbb{R}$.

Let $\mathrm{U}(1)_+ := \{e^{i\phi}\,;\, \phi \in (-\pi/2, \pi/2)\}$ be the right hand side of the unit circle. Since $z_\omega \colon \mathbb{R} \to \mathrm{U}(1)$ is continuous, we can choose $\varepsilon > 0$ such that $z_\omega(t) \in \mathrm{U}(1)_+$ for all $t \in [-\varepsilon, \varepsilon]$. If $t \in [-\varepsilon, \varepsilon]$, then $z_\omega(t) = e^{i\phi}$ for some $\phi \in (-\pi/2, \pi/2)$. Now comes the main trick. Since $z_\omega(t) = z_\omega^2(t/2)$, we have $z_\omega(t/2) = \pm e^{i\phi/2}$. But since $t/2 \in [-\varepsilon, \varepsilon]$, we have $z_\omega(t/2) \in \mathrm{U}(1)_+$. Since $+e^{i\phi} \in \mathrm{U}(1)_+$ and $-e^{i\phi} \notin \mathrm{U}(1)_+$, we conclude that $z_\omega(t/2) = +e^{i\phi/2}$.

Repeating this procedure, we find $z_\omega(\frac{1}{2^n}t) = \exp(i\frac{1}{2^n}\phi)$ for all $n \in \mathbb{N}$. Combined with the fact that $z_\omega(k\tau) = z_\omega^k(\tau)$ for all $k \in \mathbb{Z}$, we conclude that

$$z_\omega(\tfrac{k}{2^n}t) = \exp(i\tfrac{k}{2^n}\phi)$$

for all $k \in \mathbb{Z}$ and $n \in \mathbb{N}$. Since z_ω is continuous and $\{\frac{k}{2^n}\,;\, k \in \mathbb{Z}, n \in \mathbb{N}\} \subseteq \mathbb{R}$ is dense, we conclude that $z_\omega(st) = \exp(is\phi)$ for all $s \in \mathbb{R}$. So there exists a real number $a_\omega \in \mathbb{R}$, independent of τ, such that $z_\omega(\tau) = \exp(-i\tau a_\omega)$ for all $\tau \in \mathbb{R}$.

If we define

$$A := \sum_{\omega \in \Omega} a_\omega P_\omega,$$

then (16) with $z_\omega(t) = e^{-ita_\omega}$ shows that $U_t = \exp(-itA)$ for all $t \in \mathbb{R}$. □

We conclude that the time evolution of a *closed* system is given by

$$U_t = \exp(-i\tfrac{t}{\hbar}Ht) \tag{17}$$

for a Hermitian operator H. The observable H is interpreted as the *energy* of the system, and the operator $H \colon \mathcal{H} \to \mathcal{H}$ is called the *Hamilton operator.*

Remark 2.4 (Planck's constant). The *reduced Planck constant* $\hbar$ is equal to $\frac{h}{2\pi}$, where h is Planck's constant

$$h = 6.62607015 \cdot 10^{-34}\,\mathrm{J\,s}. \tag{18}$$

The SI units in which Planck's constant is expressed reflect the fundamental relation between time and energy imposed by (17). If H is to be interpreted as energy (Joules, $\mathrm{J} = \mathrm{kg\,m^2\,s^{-2}}$) and t as time (seconds, s), then $\hbar$ must be expressed in J s in order for $\exp(i\frac{t}{\hbar}H)$ to be dimensionless. The numerical value of h therefore depends on the units of time and energy, and hence on the unit of mass. Until 2018, the kilogram was defined – rather unpractically – as the mass of a small polished cylinder of platinum and iridium locked away in a vault in Paris. This situation came to an end in November 2018, when the the kilogram was redefined so as to make equation (18) exact on the nose. It is of course also possible to choose units in which $\hbar = 1$, and we will often do so.

If the time evolution of a system is governed by the Hamiltonian H, then a system that starts in state $|\psi\rangle$ at time 0 will be in state $|\psi\rangle_t = \exp(-i\frac{t}{\hbar}H)\,|\psi\rangle$ at time t.

Theorem 2.6 (Schrödinger's equation)**.** *The state $|\psi\rangle_t$ satisfies*

$$i\hbar\tfrac{d}{dt}\,|\psi\rangle_t = H\,|\psi\rangle_t \tag{19}$$

with initial condition $|\psi\rangle_0 = |\psi\rangle$.

Proof. Using the spectral decomposition, it is not hard to show that

$$\tfrac{d}{dt}\exp(-i\tfrac{t}{\hbar}H)\,|\psi\rangle = -\tfrac{i}{\hbar}H\exp(-i\tfrac{t}{\hbar}H)\,|\psi\rangle\,.$$

Indeed, the equality manifestly holds on the eigenspace V_E of H, as $\frac{d}{dt}e^{-i\frac{t}{\hbar}E}\,|\psi_E\rangle = -\frac{i}{\hbar}Ee^{-i\frac{t}{\hbar}E}\,|\psi_E\rangle$ for all $|\psi_E\rangle \in V_E$. Since the eigenspaces span $\mathcal{H}$ and since both sides of the equation are linear in ψ, we have equality on all of $\mathcal{H}$. Multiplying both sides with $i\hbar$, we obtain the Schrödinger equation. □

Remark 2.5. If we change the Hamilton operator H by an additive constant, $H \to H + \Delta E\,\mathbf{1}$, then the time evolution changes by a time-dependent phase, $U_t \to e^{-i\frac{\Delta E}{\hbar}}U_t$. Since this is physically equivalent to the old time evolution by Remark 2.3, Hamiltonians which differ by an additive constant can be considered equivalent.

Problem 2.17. Let U_t be a continuous 1-parameter group of unitary operators. In this problem, we sketch an alternative proof of Stone's theorem under the additional condition that $t \mapsto U_t\psi$ is *differentiable* for all $\psi \in \mathcal{H}$.

a) Define $A\psi := i\frac{d}{dt}|_{t=0}U_t\psi$. Then A is Hermitian.
Hint: since $\langle U_t\psi, U_t\psi\rangle = 1$, we have $\frac{d}{dt}\langle U_t\psi, U_t\psi\rangle = 0$.

b) Show that $\frac{d}{dt}|_{t_0}U_t\psi = AU_{t_0}\psi$.

c) By the part a of this problem, $V_t := \exp(-itA)$ is a 1-parameter group of unitary operators. Show that $\langle V_t\psi, U_t\psi\rangle = 1$ for all $t \in \mathbb{R}$.
Hint: it is true for $t = 0$, so it suffices to prove $\frac{d}{dt}\langle V_t\psi, U_t\psi\rangle = 0$.

d) If $\psi, \chi \in \mathcal{H}$ are unit vectors with $\langle \psi, \chi\rangle = 1$, then $\psi = \chi$.

e) Conclude that $U_t\psi = V_t\psi = \exp(-itA)\psi$ for all $t \in \mathbb{R}$.

2.5 The qubit

A *qubit* is a quantum system with two degrees of freedom. It is modelled by the 2-dimensional Hilbert space $\mathcal{H} = \mathbb{C}^2$, equipped with the *computational basis*

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Unlike a classical bit, which is either in state 0 or 1, a qubit can be in state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ for any pair $\alpha, \beta \in \mathbb{C}$ with $|\alpha|^2 + |\beta|^2 = 1$.

We think of a qubit not as a concrete quantum system, but as an abstract description of the minimal features that a system should have in order to store one 'quantum bit' of information. Realistic implementations of qubits are usually much more complicated systems, with a 2-dimensional subspace of states that behave as a qubit.

2.5.1 Observables

An observable on a qubit is a Hermitian operator $A \colon \mathbb{C}^2 \to \mathbb{C}^2$. Expressed in terms of the computational basis, an operator $A \colon \mathbb{C}^2 \to \mathbb{C}^2$ yields a Hermitian 2×2 matrix

$$A = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix}$$

with entries $a_{ij} = \langle i|\, A\, |j\rangle$. If no confusion arises, we will be sloppy and identify the operator A with the corresponding matrix.

Problem 2.18. The adjoint $A^\dagger$ is represented by the matrix

$$A^\dagger = \begin{pmatrix} \overline{a}_{00} & \overline{a}_{10} \\ \overline{a}_{01} & \overline{a}_{11} \end{pmatrix}.$$

So A is Hermitian, $A^\dagger = A$, if and only if

$$\begin{aligned} a_{00} &= \overline{a}_{00} \in \mathbb{R}, \\ a_{11} &= \overline{a}_{11} \in \mathbb{R}, \text{ and} \\ a_{10} &= \overline{a}_{01} \in \mathbb{C}. \end{aligned}$$

It follows that Hermitian 2×2-matrices A are precisely those of the form

$$A = \begin{pmatrix} t + z & x - iy \\ x + iy & t - z \end{pmatrix} \tag{20}$$

for some $t, x, y, z \in \mathbb{R}$, and the real vector space of Hermitian matrices (observables) is 4-dimensional. A convenient basis is given by

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The observables X, Y and Z are called the *Pauli matrices*. They are sometimes also called $\sigma_x = X$, $\sigma_y = Y$ and $\sigma_z = Z$, respectively. These names come from the physical realization of a qubit as the internal degrees of freedom of a spin-$\frac{1}{2}$ particle (e.g. an electron). In this setting, the observables

$$S_x := \tfrac{\hbar}{2}\sigma_x, \quad S_y := \tfrac{\hbar}{2}\sigma_y, \quad S_z := \tfrac{\hbar}{2}\sigma_z$$

are interpreted as the *spin* (internal angular momentum) of the particle in the x, y and z direction.

Since the observable Z has eigenvalues $|0\rangle$ and $|1\rangle$ with eigenvalues ± 1, it has spectral decomposition $Z = P_0 - P_1$, with spectrum $\operatorname{spec}(Z) = \{\pm 1\}$ and spectral projections $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$. If we measure a system in state $\psi = \alpha\,|0\rangle + \beta\,|1\rangle$, we therefore obtain the outcome "$Z = +1$" with probability

$$\langle\psi|\, P_0 \,|\psi\rangle = \langle\psi|0\rangle\langle 0|\psi\rangle = |\langle\psi|0\rangle|^2 = |\alpha|^2,$$

and the outcome "$Z = -1$" with probability

$$\langle\psi|\, P_1 \,|\psi\rangle = \langle\psi|1\rangle\langle 1|\psi\rangle = |\langle\psi|1\rangle|^2 = |\beta|^2.$$

So for a spin-$\frac{1}{2}$ particle in state $\psi = \alpha\,|0\rangle + \beta\,|1\rangle$, measurement of the spin in the z-direction yields outcome $+\frac{\hbar}{2}$ with probability $|\alpha|^2$, and outcome $-\frac{\hbar}{2}$ with probability $|\beta|^2$. Note that although the expectation

$$\mathbb{E}_\psi(S_z) = \langle\psi, S_z\psi\rangle = \tfrac{\hbar}{2}(|\alpha|^2 - |\beta|^2)$$

can take any value in $[-\frac{\hbar}{2}, +\frac{\hbar}{2}]$ depending on the state ψ, the only possible measurement outcomes in a single experiment are $\pm\frac{\hbar}{2}$. We say that spin in the z-direction is *quantized*.

Problem 2.19. Determine the spectrum $\operatorname{spec}(X)$ of the first Pauli matrix X, and find its spectral decomposition. Determine the probability that a measurement outcome $x \in \operatorname{spec}(X)$ occurs if we measure X for a qubit in state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$.

Problem 2.20. Do the same for the second Pauli matrix Y. Can X and Y be simultaneously measured?

The *commutator* is defined by $[A, B] = AB - BA$. The *anticommutator* by $\{A, B\} := AB + BA$.

Problem 2.21.

a) If A and B are Hermitian, then so is $i[A, B]$.

b) Verify that $[\sigma_x, \sigma_y] = 2i\sigma_z$, $[\sigma_y, \sigma_z] = 2i\sigma_x$, and $[\sigma_z, \sigma_x] = 2i\sigma_y$.

Problem 2.22.

a) If A and B are Hermitian, then so is $\{A, B\}$.

b) Verify that $\{\sigma_i, \sigma_j\} = 0$ for $i, j \in \{x, y, z\}$, $i \neq j$.

c) Verify that $\sigma_i^2 = \mathbf{1}$.

Problem 2.23. Show that the Hermitian operators $A \colon \mathbb{C}^n \to \mathbb{C}^n$ constitute a real vector space. What is its dimension?

2.5.2 States and the Bloch sphere

Recall that two unit vectors vectors ψ and ψ' in $\mathcal{H} = \mathbb{C}^2$ correspond to the same physical state if $\psi' = e^{i\phi}\psi$ for some $\phi \in [0, 2\pi]$. Physical states of the qubit therefore correspond to rays $[\psi] = \{e^{i\phi}\psi\,;\,\phi \in [0, 2\pi]\}$ in the projective Hilbert space $\mathcal{P}(\mathbb{C}^2) = \mathbb{C}\mathrm{P}^1$.

Two vectors ψ and ψ' in the same ray define the same projection operator:

$$|\psi'\rangle\langle\psi'| = \left|e^{i\phi}\psi\right\rangle\left\langle e^{i\phi}\psi\right| = e^{i\phi}\overline{e^{i\phi}}\,|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi|\,.$$

Conversely, the ray $[\psi]$ can be recovered from the projection operator $|\psi\rangle\langle\psi|$ as the set of unit vectors in the image of the projector.

Since any projection operator P is Hermitian, it can be written as

$$P = \frac{1}{2}\begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix} \tag{21}$$

for some $t, x, y, z \in \mathbb{R}$. If $P = |\psi\rangle\langle\psi|$ is a projector of rank 1, then it has trace $\mathbf{tr}(P) = 1$ and determinant $\det(|\psi\rangle\langle\psi|) = 0$. Plugging this into (21), we find $t = 1$ (for the trace) and $\frac{1}{4}(1 - x^2 - y^2 - z^2) = 0$ (for the determinant), so

$$|\psi\rangle\langle\psi| = \tfrac{1}{2}(\mathbf{1} - x\sigma_x - y\sigma_y - z\sigma_z)$$

for a vector $\vec{r} = (x, y, z)$ of norm 1. Apparently, then, the set of physical states of the qubit looks like a 2-sphere.

Definition 2.9 (Bloch sphere)**.** The *Bloch sphere* is the set of rank 1 projection operators

$$|\psi\rangle\langle\psi| = \tfrac{1}{2}(\mathbf{1} - x\sigma_x - y\sigma_y - z\sigma_z),$$

parameterized by *Bloch vectors* $\vec{r} = (x, y, z)$ of norm one.

Remark 2.6. The matrix $x\sigma_x + y\sigma_y + z\sigma_z$ is sometimes denoted $\vec{r}\cdot\vec{\sigma}$. With this notation, we have $|\psi\rangle\langle\psi| = \frac{1}{2}(\mathbf{1} - \vec{r}\cdot\vec{\sigma})$.

Problem 2.24. Show that $\mathbf{tr}(|\psi\rangle\langle\psi|) = 1$ and $\det(|\psi\rangle\langle\psi|) = 0$ for any unit vector ψ in $\mathbb{C}^2$.

Remark 2.7. The set of unit vectors in $\mathbb{C}^2$ is a 3-sphere. Indeed, $\psi = \alpha\,|0\rangle + \beta\,|1\rangle$ satisfies $\|\psi\| = 1$ if and only if $|\alpha|^2 + |\beta|^2 = 1$. Expanding $\alpha = x + iy$ and $\beta = u + iv$ into real and imaginary parts, we recover the equation $x^2 + y^2 + u^2 + v^2 = 1$ for the 3-dimensional sphere $\mathbb{S}^3 \subseteq \mathbb{R}^4$. The map from the 3-sphere $\mathbb{S}^3$ to the 2-sphere $\mathbb{S}^2$ that sends a unit vector $\psi \in \mathbb{S}^3$ to its Bloch vector $\vec{r} \in \mathbb{S}^2$ is called the *Hopf fibration.*

2.5.3 Time evolution

The hamiltonian $H\colon \mathbb{C}^2 \to \mathbb{C}^2$ of a qubit depends on the particular physical realization of the system. If we choose the computational basis $|0\rangle$, $|1\rangle$ to be a basis of eigenvectors for H, then H has matrix

$$H = \begin{pmatrix} E_0 & 0 \\ 0 & E_1 \end{pmatrix}$$

with respect to this basis. Here E_0 and E_1 are the energy levels of the eigenstates $|0\rangle$ and $|1\rangle$ respectively, $H|0\rangle = E_0|0\rangle$ and $H|1\rangle = E_1|1\rangle$. Since H is already in diagonal form, $U_t = \exp(-i\frac{t}{\hbar}H)$ is equal to

$$U_t = \begin{pmatrix} \exp(-i\frac{t}{\hbar}E_0) & 0 \\ 0 & \exp(-i\frac{t}{\hbar}E_1) \end{pmatrix} .$$

The Schrödinger equation $i\hbar\frac{d}{dt}|\psi\rangle = H|\psi\rangle$ with initial condition $|\psi\rangle_0 = \alpha|0\rangle + \beta|1\rangle$ then has solution $|\psi\rangle_t = U_t|\psi\rangle$ given by

$$|\psi\rangle_t = \alpha e^{-\frac{i}{\hbar}E_0}|0\rangle + \beta e^{-\frac{i}{\hbar}E_1}|1\rangle .$$

To find the stationary states for the time evolution, note that the initial condition $|\psi\rangle_0 = |0\rangle$ give rise to the state $|\psi\rangle_t = e^{-\frac{i}{\hbar}E_0}|0\rangle$ at time t which differs from the initial state by a global phase factor. Similarly, the initial condition $|\psi\rangle_0 = |1\rangle$ gives rise to $|\psi\rangle_t = e^{-\frac{i}{\hbar}E_1}|1\rangle$. Considered as rays $[\psi_t]$ in the projective space $\mathbb{CP}^1$, the eigenvectors $|0\rangle$ and $|1\rangle$ are therefore *stationary states* of the time evolution.

Remark 2.8. This is a general feature of the Schrödinger equation. Considered as rays in $\mathcal{P}(\mathcal{H})$, the stationary states are precisely the classes $[\psi_E]$ of eigenvectors ψ_E of the Hamilton operator, satisfying $H\psi = E\psi$. The latter equation is sometimes called the *time independent Schrödinger equation*.

For an electron coupled to a magnetic field with strength $B \in \mathbb{R}$, the Hamiltonian is proportional to $H = -\frac{\hbar}{2}B\sigma_x$ if the magnetic field points in the x-direction, and to $H = -\frac{\hbar}{2}B\sigma_y$ (or $H = -\frac{\hbar}{2}B\sigma_z$) if the magnetic field points in the y-direction (or z-direction).

Problem 2.25. a) Solve the Schrödinger equation for $H = -\frac{\hbar}{2}B\sigma_z$ with initial condition $\psi_0 = \alpha|0\rangle + \beta|1\rangle$. Do the same for $H = -\frac{\hbar}{2}B\sigma_x$ and $H = -\frac{\hbar}{2}B\sigma_y$.

Hint: remember Problems 2.19 and 2.20.

b) Sketch the orbits of $|\psi\rangle_t = U_t|\psi\rangle$ on the Bloch sphere for $H = -\frac{\hbar}{2}B\sigma_z$.

2.5.4 Tensor products of qubits

A single qubit is described by $\mathbb{C}^2$, the 2-dimensional Hilbert space with orthonormal basis

$$|0\rangle , |1\rangle . \tag{22}$$

A system of *two* qubits is described by the *tensor product* $\mathbb{C}^2 \otimes \mathbb{C}^2$, the 4-dimensional Hilbert space with orthonormal basis

$$\begin{aligned} |0\rangle \otimes |0\rangle &=: |00\rangle \\ |0\rangle \otimes |1\rangle &=: |01\rangle \\ |1\rangle \otimes |0\rangle &=: |10\rangle \\ |1\rangle \otimes |1\rangle &=: |11\rangle . \end{aligned}$$

Similarly, a system with n qubits is described by the n-fold tensor product $\underbrace{\mathbb{C}^2 \otimes \ldots \otimes \mathbb{C}^2}_{n \text{ times}}$, the 2^n-dimensional Hilbert space with orthonormal basis

$$\begin{array}{rcl} |0\rangle \otimes \ldots \otimes |0\rangle & =: & |0 \ldots 0\rangle \\ |0\rangle \otimes \ldots \otimes |1\rangle & =: & |0 \ldots 1\rangle \\ \vdots & \vdots & \vdots \\ |1\rangle \otimes \ldots \otimes |0\rangle & =: & |1 \ldots 0\rangle \,. \end{array}$$

The inner product on the tensor product is fixed by the requirement that the above basis is orthonormal.

For two vectors $|\psi\rangle, |\chi\rangle \in \mathbb{C}^2$, we define the tensor product $|\psi\rangle \otimes |\chi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ by bilinear expansion. For

$$\begin{array}{rcl} |\psi\rangle & = & \alpha_0 |0\rangle + \alpha_1 |1\rangle \\ |\chi\rangle & = & \beta_0 |0\rangle + \beta_1 |1\rangle \,, \end{array}$$

we set

$$|\psi\rangle \otimes |\chi\rangle := \alpha_0\beta_0 |00\rangle + \alpha_0\beta_1 |01\rangle + \alpha_1\beta_0 |10\rangle + \beta_1\alpha_1 |11\rangle \,. \tag{23}$$

We thus obtain a bilinear map $\otimes \colon \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$.

Problem 2.26. Show that this bilinear map is *universal*: for every bilinear map $B \colon \mathbb{C}^2 \times \mathbb{C}^2 \to U$ into a complex vector space U, there exists a unique *linear* map $\beta \colon \mathbb{C}^2 \otimes \mathbb{C}^2 \to U$ such that $B(\psi, \chi) = \beta(\psi \otimes \chi)$ for all $\psi, \chi \in \mathbb{C}^2$.

For linear maps $A \colon \mathbb{C}^2 \to \mathbb{C}^2$ and $B \colon \mathbb{C}^2 \to \mathbb{C}^2$, we define the tensor product

$$A \otimes B \colon \mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$$

by specifying it on the orthonormal basis $|i\rangle \otimes |j\rangle$ for $i, j \in \{0, 1\}$,

$$A \otimes B(|i\rangle \otimes |j\rangle) := (A\,|i\rangle) \otimes (B\,|j\rangle). \tag{24}$$

Problem 2.27. Verify that $A \otimes B(|\psi\rangle \otimes |\chi\rangle) := (A\,|\psi\rangle) \otimes (B\,|\chi\rangle)$ for all $\psi, \chi \in \mathbb{C}^2$, and conclude that the above definition is basis independent.

If A and B are Hermitian, we interpret $A \otimes \mathbf{1}$ as an observable that pertains only to the first qubit, and $\mathbf{1} \otimes B$ as one that pertains only to the second qubit.

Problem 2.28. Show that $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$.

Problem 2.29. For all linear maps $A \colon \mathbb{C}^2 \to \mathbb{C}^2$ and $B \colon \mathbb{C}^2 \to \mathbb{C}^2$, we have $[A \otimes \mathbf{1}, \mathbf{1} \otimes B] = 0$, also if A and B do not commute. So observables on *different* systems can *always* be simultaneously measured.

2.5.5 The CHSH game revisited

We return to the CHSH game from §1, and give the mathematical description of the Orsay experiment.

The internal degrees of freedom of a single photon are modelled by the Hilbert space $\mathcal{H} = \mathbb{C}^2$, and a photon with polarization angle α is described by the unit vector

$$|\psi_\alpha\rangle = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}.$$

The projection corresponding to the event that the photon has polarization in the α-direction is $P(\alpha) = |\psi_\alpha\rangle \langle\psi_\alpha|$, with matrix

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

The probability that a photon with vertical polarization ($\alpha = 0$) is measured to have polarization in the α-direction is therefore $\langle\psi_0, P(\alpha)\psi_0\rangle = \cos^2(\alpha)$. If a beam of vertically polarized photons passes through a filter that is polarized in the α-direction, then a fraction $\cos^2(\alpha)$ of the photons will pass through, resulting in the transmission equation $I_{\text{out}} = \cos^2(\alpha) I_{\text{in}}$ from §1.2.

We turn to the Orsay experiment, where the internal degrees of freedom for a pair of photons are described by the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B = \mathbb{C}^2 \otimes \mathbb{C}^2$. The pair of photons emitted by the calcium atom are in the entangled state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \tag{25}$$

on $\mathcal{H}_B = \mathbb{C}^2$. The probability that *both* Alice and Bob see a photon come through their filter is therefore

$$\begin{aligned}
\langle\Psi, P(\alpha) \otimes P(\beta)\Psi\rangle &= \tfrac{1}{2}\big(\cos^2(\alpha)\sin^2(\beta) + \sin^2(\alpha)\cos^2(\beta)\big) \\
&\quad - \cos(\alpha)\sin(\alpha)\cos(\beta)\sin(\beta) \\
&= \tfrac{1}{2}\big(\cos(\alpha)\sin(\beta) - \sin(\beta)\cos(\alpha)\big)^2 \\
&= \tfrac{1}{2}\sin^2(\alpha - \beta).
\end{aligned}$$

Similarly, the probability that both Alice and Bob see their photon blocked is

$$\langle\Psi, P^\perp(\alpha) \otimes P^\perp(\beta)\Psi\rangle = \tfrac{1}{2}\sin^2(\alpha - \beta),$$

where $P^\perp = \mathbf{1} - P$ denotes the complementary projection. The probability that Alice and Bob see the same thing (either both photons are blocked or both of them come through) is therefore

$$\sin^2(\alpha - \beta),$$

predicting the result (2) of the Orsay experiment.

Problem 2.30. Have a look at the 'scientific description' for the Nobel Prize of Physics 2022, [NP22]. You should be able to follow much of this with what you've learned so far.

2.6 Tensor products

The following abstract definition of tensor products captures not so much what tensor products *are*, but what they *do*.

Definition 2.10 (Tensor products)**.** A tensor product of two vector spaces V and W is a vector space $V \otimes W$, together with a bilinear map $\otimes \colon V \times W \to V \otimes W$ that is *universal*: for every *bilinear* map $B \colon V \times W \to U$ into a vector space U, there exists a unique *linear* map $\beta \colon V \otimes W \to U$ such that $B = \beta \circ \otimes$.

$$\begin{tikzcd}
V \times W \arrow[r, "\otimes"] \arrow[d, "\forall B"] & V \otimes W \arrow[dl, dashed, "\exists ! \beta"] \\
U &
\end{tikzcd}$$

The following result shows that tensor products are unique up to linear isomorphism.

Theorem 2.7 (Uniqueness of tensor products)**.** *Let $V \otimes W$ and $V \widetilde{\otimes} W$ be tensor products of V and W. Then there exists a unique linear isomorphism $\iota \colon V \otimes W \xrightarrow{\sim} V \widetilde{\otimes} W$ such that $\iota \circ \otimes = \widetilde{\otimes}$.*

Proof. Consider the commutative diagram

$$\begin{tikzcd}
& V \times W \arrow[dl, "\otimes"'] \arrow[dr, "\widetilde{\otimes}"] & \\
V \otimes W \arrow[rr, dashed, shift left, "\iota"] & & V \widetilde{\otimes} W. \arrow[ll, dashed, shift left, "\kappa"]
\end{tikzcd}$$

The universal property for $\otimes$ yields a linear map $\iota \colon V \otimes W \to V \widetilde{\otimes} W$ such that $\iota \circ \otimes = \widetilde{\otimes}$. Similarly, the universal property for $\widetilde{\otimes}$ yields a linear map $\kappa \colon V \widetilde{\otimes} W \to V \otimes W$ with $\kappa \circ \widetilde{\otimes} = \otimes$.

To show that $\kappa \circ \iota$ is the identity, consider the commutative diagram

$$\begin{tikzcd}
& V \times W \arrow[dl, "\otimes"'] \arrow[dr, "\otimes"] & \\
V \otimes W \arrow[rr, dashed, shift left, "\kappa \circ \iota"] \arrow[rr, dashed, shift right, "\mathrm{Id}"'] & & V \otimes W.
\end{tikzcd}$$

Note that both the identity $\mathrm{Id} \colon V \otimes W \to V \otimes W$ and the composition $\kappa \circ \iota \colon V \otimes W \to V \otimes W$ respect the tensor product, $\otimes = \kappa \circ \widetilde{\otimes} = \kappa \circ \iota \circ \otimes$. But by the universal property of $\otimes$, the linear map with this property is *unique*, so $\kappa \circ \iota = \mathrm{Id}$. Similarly, the universal property of $\widetilde{\otimes}$ guarantees that $\iota \circ \kappa = \mathrm{Id}$, so κ is the inverse of ι. □

Because the tensor product is essentially unique, one often speaks of *the* tensor product of V and W. We will show momentarily that tensor products always exist, but before we do so, two remarks on Definition 2.10 are in order. First of all, there are in fact *many* ways to construct the tensor product, each with its advantages and disadvantages. By Theorem 2.7 we are free to choose whichever construction we find most convenient for the purpose at hand. Secondly, for some purposes Definition 2.10 itself is quite convenient. Here is an example, to be compared with (24) in §2.5.4.

Proposition 2.8. *If $A\colon V_0 \to V_1$ and $B\colon W_0 \to W_1$ are linear maps, then there exists a unique linear map $A \otimes B\colon V_0 \otimes W_0 \to V_1 \otimes W_1$ such that*

$$A \otimes B(v \otimes w) = (Av) \otimes (Bw)$$

for all $v \in V_0$, $w \in W_0$.

Proof. Use the universal property for $V_0 \otimes W_0$ for the bilinear map $B\colon V_0 \times W_0 \to V_1 \otimes W_1$ defined by $B(v, w) := (Av) \otimes (Bw)$,

$$\begin{tikzcd}
V_0 \times W_0 \arrow[r, "\otimes"] \arrow[d, "B"] & V_0 \otimes W_0 \arrow[dl, dashed, "\exists! A\otimes B"] \\
V_1 \otimes W_1 &
\end{tikzcd}$$

□

To show that tensor products of *finite dimensional* vector spaces exist, we can follow the line of reasoning from §2.5.4.

Problem 2.31. In §2.5.4, we have constructed $\mathbb{C}^2 \otimes \mathbb{C}^2$, together with the bilinear map $\otimes\colon \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$. Extend this construction to the tensor product $\mathbb{C}^n \otimes \mathbb{C}^m$, and show that $\otimes\colon \mathbb{C}^n \times \mathbb{C}^m \to \mathbb{C}^n \otimes \mathbb{C}^m$ has the universal property.

2.6.1 Quotient construction of tensor products

The following construction yields a tensor product for *any* two vector spaces V and W, finite dimensional or not. Let $\mathcal{F}(V \times W)$ be the free vector space[2] with basis $\delta_{(v,w)}$ labelled by elements of $V \times W$. An element $f \in \mathcal{F}(V \times W)$ is given by a finite (but arbitrarily large) formal linear combination

$$f = \sum_{i=1}^{N} \alpha_i \delta_{(v_i,w_i)}$$

of different basis vectors $\delta_{(v_i,w_i)}$. Let $\mathcal{R} \subseteq \mathcal{F}(V \times W)$ be the linear subspace of *relations*, spanned by the vectors

$$\delta_{(v,\alpha w+\beta w')} \quad - \quad \big(\alpha\delta_{(v,w)} + \beta\delta_{(v,w')}\big) \tag{26}$$

$$\delta_{(\alpha v+\beta v',w)} \quad - \quad \big(\alpha\delta_{(v,w)} + \beta\delta_{(v',w')}\big), \tag{27}$$

with $v, v' \in V$, $w, w' \in W$ and $\alpha, \beta \in \mathbb{C}$. The tensor product of V and W is constructed as the quotient vector space

$$V \otimes W = \mathcal{F}(V \times W)/\mathcal{R},$$

and $v \otimes w := [\delta_{(v,w)}]$ is the equivalence class of the basis vector $\delta_{(v,w)}$ under the equivalence relation $f \sim f' \Leftrightarrow f - f' \in \mathcal{R}$.

[2] One can think of elements of $\mathcal{F}(V \times W)$ as functions $V \times W \to \mathbb{C}$ with finite support, and of $\delta_{(v,w)}\colon V \times W \to \mathbb{C}$ as the function that takes the value 1 on (v, w) and 0 on $V \times W \backslash \{(v, w)\}$.

Problem 2.32 (Quotient vector spaces)**.** Let F be a vector space, and let $R \subseteq F$ be a linear subspace. Show that $f \sim f' \Leftrightarrow f - f' \in R$ is an equivalence relation. Let $F/R := \{[f]\,;\, f \in F\}$ be the set of equivalence classes. Show that addition and scalar multiplication on F/R are well defined by $[f]+[f'] := [f+f']$ and $\lambda[f] := [\lambda f]$, and that F/R is a vector space.

Proposition 2.9. *The vector space $V \otimes W = \mathcal{F}(V \times W)/\mathcal{R}$ with the bilinear map $v \otimes w = [\delta_{(v,w)}]$ is a tensor product.*

Proof. To show that $\otimes\colon V \times W \to V \otimes W$ is bilinear, we need to show that $v \otimes (\alpha w + \beta w') = \alpha(v \otimes w) + \beta(v \otimes w')$. This follows from

$$[\delta_{(v,\alpha w+\beta w')}] - \alpha[\delta_{(v,w)}] - \beta[\delta_{(v,w')}] = [\delta_{(v,\alpha w+\beta w')} - (\alpha\delta_{(v,w)} + \beta\delta_{(v,w')})] = [0],$$

where the last step follows because (26) is by definition an element of $\mathcal{R}$. Similarly, $(\alpha v + \beta v') \otimes w = \alpha(v \otimes w) + \beta(v' \otimes w)$ because (27) is in $\mathcal{R}$.

Next, note that *any* map $B\colon V \times W \to U$ into a vector space U, linear or not, gives rise to a unique *linear* map $\widetilde{\beta}\colon \mathcal{F}(V \times W) \to U$ with

$$\widetilde{\beta}(\delta_{(v,w)}) = B(v,w)$$

for all $(v,w) \in V \times W$. Indeed, a linear map on $\mathcal{F}(V \times W)$ is uniquely determined by what it does on the basis. If B is bilinear, then $\widetilde{\beta}$ vanishes on $\mathcal{R}$, so $\beta([f]) := \widetilde{\beta}(f)$ is a well-defined linear map $\beta\colon V \otimes W \to U$ with $\beta(v \otimes w) = B(v,w)$. The map β is uniquely determined by B because $\widetilde{\beta}$ is. □

Problem 2.33. Prove Proposition 2.8 using the quotient construction of the tensor product.

2.6.2 Tensor product construction using linear maps

We give a more explicit construction of the tensor product in the case that $\mathcal{H}$ is a finite dimensional Hilbert space.

For every Hilbert space $\mathcal{H}$, the *conjugate* Hilbert space $\overline{\mathcal{H}}$ is equal to $\mathcal{H}$ as a set, with the same addition but with scalar multiplication twisted by complex conjugation. If we write $\overline{\psi} \in \overline{\mathcal{H}}$ for $\psi \in \mathcal{H}$ considered as an element of $\mathcal{H}$, then

$$\overline{\psi} + \overline{\chi} = \overline{\psi + \chi}$$
$$\overline{\lambda} \cdot \overline{\psi} = \overline{\lambda\psi}$$

for all $\psi, \chi \in \mathcal{H}$. The inner product on $\overline{\mathcal{H}}$ is $\langle\overline{\psi}, \overline{\chi}\rangle := \overline{\langle\psi, \chi\rangle}$.

Proposition 2.10. *The vector space $\mathcal{H}_A \otimes \mathcal{H}_B = \mathcal{L}(\overline{\mathcal{H}}_B, \mathcal{H}_A)$ of linear maps from $\mathcal{H}_B$ to $\mathcal{H}_A$ is a tensor product, with $\psi \otimes \chi = |\psi\rangle\,\langle\overline{\chi}|$.*

Proof. To show that $\otimes\colon \mathcal{H}_A \times \mathcal{H}_B \to \mathcal{L}(\overline{\mathcal{H}}_B, \mathcal{H}_A)$ is bilinear, we use that the inner product is *antilinear* in the left argument:

$$\begin{aligned} \psi \otimes (\alpha\chi + \beta\chi') &= |\psi\rangle\,\langle\overline{\alpha}\,\overline{\chi} + \overline{\beta}\,\overline{\chi}'| \\ &= \alpha\,|\psi\rangle\,\langle\overline{\chi}| + \beta\,|\psi\rangle\,\langle\overline{\chi}'| \\ &= \alpha(\psi \otimes \chi) + \beta(\psi \otimes \chi'). \end{aligned}$$

This is the reason we need the complex conjugate in the definition of the tensor product. The other equation $(\alpha\psi + \beta\psi') \otimes \chi = \alpha(\psi \otimes \chi) + \beta(\psi' \otimes \chi)$ is similar, except that one doesn't have to conjugate twice.

If e_i is a basis of $\mathcal{H}_A$ and f_j is a basis of $\mathcal{H}_B$, then $|e_i\rangle \langle \overline{f}_j|$ is a basis of $\mathcal{L}(\overline{\mathcal{H}}_B, \mathcal{H}_A)$. It follows that for every bilinear form $B\colon \mathcal{H}_A \times \mathcal{H}_B \to U$, there exists a unique linear map $\beta\colon \mathcal{L}(\overline{\mathcal{H}}_B, \mathcal{H}_A) \to U$ with $\beta(|e_i\rangle \langle \overline{f}_j|) = B(e_i, f_j)$ for all e_i and f_j. Since the bilinear map $\beta \circ \otimes\colon \mathcal{H}_A \times \mathcal{H}_B \to U$ agrees with B on basis vectors of $\mathcal{H}_A$ and $\mathcal{H}_B$, it agrees with B on all of $\mathcal{H}_A \times \mathcal{H}_B$. □

On $\mathcal{H}_A \otimes \mathcal{H}_B = \mathcal{L}(\overline{\mathcal{H}}_B, \mathcal{H}_A)$, we have a natural inner product

$$\langle X, Y\rangle := \mathbf{tr}(X^\dagger Y). \tag{28}$$

(Recall from Definition 2.2 that $X^\dagger\colon \mathcal{H}_A \to \overline{\mathcal{H}}_B$ is the unique linear map such that $\langle X^\dagger\psi, \overline{\chi}\rangle_{\overline{\mathcal{H}}_B} = \langle \psi, X\overline{\chi}\rangle_{\mathcal{H}_A}$ for all $\psi \in \mathcal{H}_A$ and $\overline{\chi} \in \mathcal{H}_B$.)

Problem 2.34. This is the unique inner product on $\mathcal{H}_A \otimes \mathcal{H}_B$ such that

$$\langle \psi_1 \otimes \chi_1, \psi_2 \otimes \chi_2\rangle = \langle \psi_1, \psi_2\rangle_A \langle \chi_1, \chi_2\rangle_B \tag{29}$$

for all $\psi_1, \psi_2 \in \mathcal{H}_A$ and $\chi_1, \chi_2 \in \mathcal{H}_B$.

Problem 2.35. The four Bell states $\frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$ and $\frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$ constitute an orthonormal basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$.

Problem 2.36. Let $U^{(1)}_t$ and $U^{(2)}_t$ be continuous one-parameter groups of unitary operators on $\mathcal{H}_1 = \mathbb{C}^{n_1}$ and $\mathcal{H}_2 = \mathbb{C}^{n_2}$, with generators given by $H^{(1)}$ and $H^{(2)}$, respectively.

a) Show that $U^{(1)}_t \otimes U^{(2)}_t$ is a continuous one-parameter group of unitary operators on $\mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2}$.

b) Express its generator in terms of $H^{(1)}$ and $H^{(2)}$.

2.6.3 Entanglement and the Schmidt decomposition

An element of $\mathcal{H}_A \otimes \mathcal{H}_B$ is called *elementary* if it is of the form $\psi \otimes \chi$, and *entangled* otherwise. Entanglement is a primary resource in quantum information and quantum computing. Although entanglement is generic (the elementary vectors form a subset of measure zero in $\mathcal{H}_A \otimes \mathcal{H}_B$), it is technically very challenging to preserve entanglement between different systems.

Among other things, the *Schmidt decomposition* allows us to quantify entanglement. Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be Hilbert spaces of dimension n and m, respectively. Without loss of generality, we may assume that $n \geq m$.

Theorem 2.11 (Schmidt decomposition)**.** *For every $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, there exist orthonormal bases $|u_i\rangle$ of $\mathcal{H}_A$ and $|v_j\rangle$ of $\mathcal{H}_B$ such that*

$$|\psi\rangle = \sum_{i=1}^{m} \lambda_i |u_i\rangle \otimes |v_i\rangle \tag{30}$$

for nonnegative numbers $\lambda_i \geq 0$. These Schmidt coefficients *are unique up to reordering.*

Proof. This is essentially the singular value decomposition for $\Psi \in \mathcal{L}(\overline{\mathcal{H}}_B, \mathcal{H}_A)$. Suppose that $\Psi = U\Sigma V^\dagger$ for unitary maps $U\colon \mathcal{H}_A \to \mathcal{H}_A$ and $V\colon \overline{\mathcal{H}}_B \to \overline{\mathcal{H}}_B$, and for a linear map $\Sigma\colon \overline{\mathcal{H}}_B \to \mathcal{H}_A$ that has matrix

$$\Sigma = \begin{pmatrix} \lambda_1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & \lambda_m \\ 0 & \dots & 0 \\ & \vdots & \\ 0 & \dots & 0 \end{pmatrix}$$

with respect to an orthonormal basis $|e_i\rangle$ of $\mathcal{H}_A$ and $\left|\overline{f}_j\right\rangle$ of $\mathcal{H}_B$. The singular values λ_i are uniquely determined by Ψ up to reordering, since the eigenvalues of $\Psi^\dagger\Psi$ are $|\lambda_i|^2$. In terms of the orthonormal basis $|u_i\rangle = U\,|e_i\rangle$ of $\mathcal{H}_A$ and $|\overline{v}_i\rangle = V\left|\overline{f}_i\right\rangle$ of $\overline{\mathcal{H}}_B$, the singular value decomposition $\Psi = U\Sigma V^\dagger$ reads

$$\Psi = \sum_{i=1}^{m} \lambda_i\,|u_i\rangle\,\langle\overline{v}_i|\,.$$

Identifying $\mathcal{L}(\overline{\mathcal{H}}_B, \mathcal{H}_A)$ with $\mathcal{H}_A \otimes \mathcal{H}_B$, this yields the desired result. □

Note that since $\langle\psi,\psi\rangle = \mathbf{tr}(\Psi^\dagger\Psi) = 1$, the Schmidt coefficients of a unit vector $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ satisfy

$$\sum_{i=1}^{m} \lambda_i^2 = 1.$$

Since an elementary unit vector $|\psi\rangle \otimes |\chi\rangle$ corresponds to the rank 1 operator $|\psi\rangle\,\langle\overline{\chi}|$ in $\mathcal{L}(\overline{\mathcal{H}}_B, \mathcal{H}_A)$, it has Schmidt coefficients $1, 0, \dots, 0$.

Problem 2.37. Determine the Schmidt coefficients of the Bell state $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, and conclude that it is entangled.

Problem 2.38. Show that the unit vector $\psi = \alpha\,|00\rangle + \beta\,|11\rangle + \gamma\,|10\rangle + \delta\,|01\rangle$ in $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ is entangled if and only if $\alpha\beta - \gamma\delta \neq 0$.

The Schmidt coefficients can be used to *quantify* the amount of entanglement.

Problem 2.39. The *entanglement entropy* of $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is defined by $S(\psi) = -\sum_i \lambda_i^2 \log(\lambda_i^2)$. Show that $S(\psi) \geq 0$, and $S(\psi) \neq 0$ if and only ψ is entangled.

Problem 2.40. The entanglement entropy is invariant under unitary transformations of $\mathcal{H}_A$ and $\mathcal{H}_B$ separately, but not under unitary transformations of the whole space $\mathcal{H}_A \otimes \mathcal{H}_B$.

Problem 2.41. Show that $0 \le S(\psi) \le \log(d)$ for $d = \min(n, m)$.

Hint: maximize $-\sum_{i=1}^d \lambda_i^2 \log(\lambda_i^2)$ *with constraints* $\lambda_i^2 \ge 0$ *and* $\sum_{i=1}^d \lambda_i^2 = 1$.

We say that a state ψ is *maximally entangled* if $S(\psi) = \log(d)$.

Problem 2.42. Let e_i be an orthonormal basis of a Hilbert space $\mathcal{H}$ of dimension d. Then the state $|\psi_{\max}\rangle := \frac{1}{\sqrt{d}} \sum_{i=1}^d e_i \otimes e_i$ in $\mathcal{H} \otimes \mathcal{H}$ is maximally entangled.

Problem 2.43. For any maximally entangled state $|\psi\rangle$ in $\mathcal{H} \otimes \mathcal{H}$, there exist unitary operators U and V on $\mathcal{H}$ such that $|\psi\rangle = U \otimes V \, |\psi_{\max}\rangle$.

2.7 Systems with infinitely many degrees of freedom

In quantum information theory and quantum computing, we often deal with systems that have finitely many degrees of freedom. Since these arise in practise as simplifications/partial descriptions of quantum systems that *do* have infinitely many degrees of freedom, we briefly sketch the adaptations to our setting that are necessary to include this case as well.

2.7.1 States

Since the number of degrees of freedom of a quantum system corresponds to the dimension of the inner product space $\mathcal{H}$, we will need to handle the case where $\mathcal{H}$ is infinite dimensional. In this setting, we require an additional completeness assumption. Recall that a metric space is *complete* if every Cauchy sequence converges. Since the metric on $\mathcal{H}$ is given by $d(\psi, \chi) = \|\psi - \chi\|$, a sequence ψ_n in $\mathcal{H}$ is *Cauchy* if for every $\varepsilon > 0$, there exists a number $N \in \mathbb{N}$ such that $\|\psi_n - \psi_m\| \le \varepsilon$ for all $n, m > N$. We say that $\lim_{n\to\infty} \psi_n = \psi$ if $\lim_{n\to\infty} \|\psi - \psi_n\| = 0$. As usual, we abbreviate $\lim_{N\to\infty} \sum_{n=0}^N \chi_n$ by $\sum_{n=0}^\infty \chi_n$.

Definition 2.11. A Hilbert space is a complex vector space $\mathcal{H}$ with an inner product $\langle \cdot, \cdot \rangle \colon \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ such that $\mathcal{H}$ is *complete* as a metric space.

Since finite dimensional vector spaces are automatically complete, this extra assumption is redundant there. This marks a sharp distinction in the mathematical description of quantum mechanics: systems with *finitely many* degrees of freedom can be handled using linear algebra and discrete probability theory. For systems with *infinitely many* degrees of freedom, one needs more refined tools from functional analysis and measure theory.

Problem 2.44. In the space $\mathcal{H} = C([0,1], \mathbb{C})$ of continuous functions with inner product $\langle \psi, \chi \rangle = \int_0^1 \overline{\psi(x)} \chi(x) dx$, the sequence

$$\psi_n(x) := \begin{cases} nx & \text{for } x \in [0, 1/n) \\ 1 & \text{for } x \in [1/n, (n-1)/n] \\ n(1-x) & \text{for } x \in ((n-1)/n, 1] \end{cases}$$

is Cauchy, but not convergent. So $\mathcal{H} = C([0,1], \mathbb{C})$ is not a Hilbert space.

For a linear map $A\colon \mathcal{H} \to \mathcal{H}$, we define the *operator norm* by

$$\|A\| := \sup\{\|A\psi\|\,;\,\psi \in \mathcal{H}, \|\psi\| = 1\}.$$

A linear map $A\colon \mathcal{H} \to \mathcal{H}$ is called *bounded* if $\|A\| < \infty$.

Problem 2.45. A linear map is bounded if and only if it is continuous.

Definition 2.12. An *orthonormal basis* of $\mathcal{H}$ is a sequence $\psi_n \in \mathcal{H}$ such that $\langle \psi_n, \psi_m \rangle = \delta_{n,m}$ for all $n, m \in \mathbb{N}$, and such that for all $\psi \in \mathcal{H}$, there exist $c_n \in \mathbb{C}$ such that $\sum_{n=0}^{\infty} c_n \psi_n = \psi$.

Problem 2.46. If $\psi = \sum_{n=0}^{\infty} c_n \psi_n$ for an orthonormal basis ψ_n of $\mathcal{H}$, then $c_n = \langle \psi_n, \psi \rangle$.

2.7.2 Events

The description of *events* is to a large extent the same as for finite dimensional systems. An event is modelled by a projection $P\colon \mathcal{H} \to \mathcal{H}$ with $P^2 = P^\dagger = P$, or, equivalently, by the *closed* linear subspace $V \subseteq \mathcal{H}$ onto which it projects.

As for probability measures, PVMs are best formulated in terms of a measurable space $(\Omega, \mathcal{F})$, where $\mathcal{F}$ is a σ-algebra on the set Ω.

Definition 2.13 (PVMs)**.** A *projection valued measure* (PVM) on $(\Omega, \mathcal{F})$ is a mapping $E \mapsto P(E)$ from $\mathcal{F}$ into the projections on $\mathcal{H}$ such that

i) $P(\emptyset) = 0$ and $P(\Omega) = \mathbf{1}_{\mathcal{H}}$.

ii) $P(E \cap F) = P(E)P(F)$ for all measurable subsets $E, F \subseteq \Omega$.

iii) If the subsets $E_n \subseteq \Omega$ are measurable and pairwise disjoint, then

$$P(\sqcup_{n=0}^{\infty} E_n)\psi = \sum_{n=0}^{\infty} P(E_n)\psi \quad \text{for all } \psi \in \mathcal{H}.$$

Problem 2.47. Let $E \mapsto P(E)$ be a PVM on a sigma algebra $(\Omega, \mathcal{F})$, and let $\psi \in \mathcal{H}$ be a unit vector. Then the assignment $E \mapsto \langle \psi, P(E)\psi \rangle$ is a probability measure. It is often denoted by $E \mapsto \mathbb{P}_\psi(E)$

Problem 2.48. Let Ω be a finite set, and let $\mathcal{F}$ be its power set (the collection of all possible subsets). Explain the relation between Definition 2.4 and Definition 2.13.

Example 2.2 (EM field)**.** A single mode (frequency) in the electromagnetic field is described by the Hilbert space $\mathcal{H} = \ell^2(\mathbb{N})$ of square integrable sequences $(a_n)_{n\in\mathbb{N}}$ of complex numbers with inner product $\langle a, b \rangle = \sum_{n=0}^{\infty} \overline{a}_n b_n$. On the measurable space $(\mathbb{N}, \mathcal{F})$ where $\mathcal{F}$ is the power set of $\mathbb{N}$, we define the PVM $E \mapsto P(E)$ by

$$(P(E)a)_n = \begin{cases} a_n & \text{if } n \in E \\ 0 & \text{if } n \notin E. \end{cases} \tag{31}$$

The projection $P(\{n\})$ is interpreted as the event "there are n photons in this mode".

Example 2.3 (Particle in $\mathbb{R}^d$)**.** A particle moving in 1 dimension is described by the Hilbert space $\mathcal{H} = L^2(\mathbb{R})$ of square integrable functions $\psi\colon \mathbb{R} \to \mathbb{C}$ (modulo null sets), with inner product $\langle \psi, \chi\rangle = \int_{\mathbb{R}} \overline{\psi}(x)\chi(x)dx$. On the Borel σ-algebra $(\mathbb{R}, \mathcal{F})$ we define the PVM $E \mapsto P(E)$ by

$$(P(E)\psi)(x) = \begin{cases} \psi(x) & \text{if } x \in E \\ 0 & \text{if } x \notin E. \end{cases} \tag{32}$$

The projection $P(E)$ is interpreted as the event "the position x of the particle is an element of $E \subseteq \mathbb{R}$". For a single particle moving in $\mathbb{R}^d$, the relevant Hilbert space is $L^2(\mathbb{R}^d)$.

2.7.3 Observables

Mimicking (11), we wish to construct for any random variable $a\colon \Omega \to \mathbb{R}$ a Hermitian operator

$$A = \int_{\Omega} a(\omega)P(d\omega) \tag{33}$$

on $\mathcal{H}$ with the property that $\langle \psi, A\psi\rangle$ is the expectation of the random variable $a\colon \Omega \to \mathbb{R}$ with respect to the probability measure $\mathbb{P}_{\psi}(E) = \langle \psi, P(E)\psi\rangle$ induced by any unit vector $\psi \in \mathcal{H}$. If the random variable is unbounded, we encounter a rather serious technical difficulty.

To illustrate this, consider the setting of Example 2.2. For a single mode in the EM field with (angular) frequency ω, the *energy* $E(n) = (n + \frac{1}{2})\hbar\omega$ is an unbounded random variable on $\Omega = \mathbb{N}$. The corresponding Hermitian operator

$$H = \hbar\omega \sum_{n=0}^{\infty} (n + \tfrac{1}{2})P(\{n\}) \tag{34}$$

maps the sequence a_n to the sequence $\hbar\omega(n + \frac{1}{2})a_n$. Unfortunately, $\sum_{n=0}^{\infty} a_n^2 < \infty$ does *not* imply $\sum_{n=0}^{\infty}((n+\frac{1}{2})a_n)^2 < \infty$, so the expression (34) does *not* define a map on all of $\mathcal{H} = \ell^2(\mathbb{N})$. To resolve the issue, we consider H as a linear map $H\colon \mathrm{Dom}(H) \to \ell^2(\mathbb{N})$, defined on the dense linear subspace $\mathrm{Dom}(H) = \{(a_n)_{n\in\mathbb{N}}\,;\, \sum_{n=0}^{\infty}((n + \frac{1}{2})a_n)^2 < \infty\}$ of $\ell^2(\mathbb{N})$.

Similarly, in the setting of Example 2.3, the *position* $x \mapsto x$ is an unbounded random variable on $\Omega = \mathbb{R}$. It seems reasonable to interpret $X = \int_{-\infty}^{\infty} xP(dx)$ in equation (33) as the linear map into $L^2(\mathbb{R})$ given by

$$(X\psi)(x) = x\psi(x).$$

Again, $\int_{-\infty}^{\infty} |\psi(x)|^2 < \infty$ does *not* imply $\int_{-\infty}^{\infty} |x\psi(x)|^2 < \infty$, so we consider X as a linear map $X\colon \mathrm{Dom}(X) \to L^2(\mathbb{R})$, defined on the dense linear subspace $\mathrm{Dom}(X) = \{\psi \in L^2(\mathbb{R})\,;\, \int_{-\infty}^{\infty} |x\psi(x)|^2 < \infty\}$.

Definition 2.14. A *linear operator* on $\mathcal{H}$ is a linear subspace $\mathrm{Dom}(A) \subseteq \mathcal{H}$, together with a linear map $A\colon \mathrm{Dom}(A) \to \mathcal{H}$. It is *densely defined* if $\mathrm{Dom}(A)$ is dense in $\mathcal{H}$.

So the energy H and position X are linear operators on $\ell^2(\mathbb{N})$ and $L^2(\mathbb{R})$, respectively. The *spectrum* $\mathrm{spec}(A)$ is the set of all $\lambda \in \mathbb{C}$ for which $\lambda\mathbf{1} - A$ does not have a bounded inverse.

Definition 2.15. A densely defined linear operator A is *symmetric* if

$$\langle \psi, A\chi \rangle = \langle A\psi, \chi \rangle \tag{35}$$

for all $\psi, \chi \in \mathrm{Dom}(A)$.

Problem 2.49. The linear operators H and X are symmetric.

It turns out that symmetric operators are not *quite* the appropriate generalization of Hermitian operators on finite dimensional Hilbert spaces.

Definition 2.16. A symmetric operator A is *self-adjoint* if $\psi \in \mathrm{Dom}(A)$ if and only if the linear map $\mathrm{Dom}(A) \to \mathbb{C}$ defined by $\chi \mapsto \langle \psi, A\chi \rangle$ is continuous.

Remark 2.9. This condition says that $\mathrm{Dom}(A)$ is in some sense the 'largest possible' domain for A. Indeed, suppose that $\overline{A}$ is a *symmetric extension* of A, i.e. a symmetric operator $\overline{A}$ that agrees with A on $\mathrm{Dom}(A)$, but with possibly larger domain $\mathrm{Dom}(\overline{A}) \supseteq \mathrm{Dom}(A)$. Then since $\langle \psi, A\chi \rangle = \langle \overline{A}\psi, \chi \rangle$ for all $\chi \in \mathrm{Dom}(A)$ and $\psi \in \mathrm{Dom}(\overline{A})$, the linear functional $\chi \mapsto \langle \psi, A\chi \rangle$ is continuous for all $\psi \in \mathrm{Dom}(\overline{A})$. The fact that A is self-adjoint then implies $\mathrm{Dom}(\overline{A}) = \mathrm{Dom}(A)$.

Problem 2.50. If A is symmetric and $\psi \in \mathrm{Dom}(A)$, then $\chi \mapsto \langle \psi, A\chi \rangle$ is automatically continuous.

For self-adjoint operators (but not for merely symmetric ones), there exist suitable generalizations of the Spectral Theorem 2.2 and Stones' Theorem 2.5.

Theorem 2.12 (Observables from random variables)**.** *Let P be a PVM on $(\Omega, \mathcal{F})$, and let $a \colon \Omega \to \mathbb{R}$ be measurable. Then there exists a unique self-adjoint operator A such that:*

i) $\mathrm{Dom}(A) = \{\psi \in \mathcal{H}\,;\, \int_\Omega |a(\omega)|^2 \mathbb{P}_\psi(d\omega) < \infty\}$,

ii) $\langle \psi, A\psi \rangle = \int_\Omega a(\omega)\mathbb{P}_\psi(d\omega)$ *for every unit vector* $\psi \in \mathrm{Dom}(A)$,

iii) $\langle A\psi, A\psi \rangle = \int_\Omega |a(\omega)|^2 \mathbb{P}_\psi(d\omega)$ *for every unit vector* $\psi \in \mathrm{Dom}(A)$.

Proof. See for example Theorem 4.7 in Section X of the book [Co07]. □

We denote this operator by $A =: \int_\Omega a(\omega)P(d\omega)$, giving rigorous meaning to (33).

Problem 2.51. Verify that H and X satisfy i), ii) and iii).

Theorem 2.13 (Spectral theorem)**.** *For every self-adjoint operator A, there exists a PVM on $\mathbb{R}$ such that*

$$A = \int_{\mathbb{R}} aP(da).$$

This PVM is concentrated on $\mathrm{spec}(A)$*, so* $P(E) = 0$ *if* $E \cap \mathrm{spec}(A) = \emptyset$.

Proof. See for example Theorem 4.11 in Section X of the book [Co07]. □

So a self-adjoint operator A is determined by its spectrum $\mathrm{spec}(A) \subseteq \mathbb{R}$ and an PVM on $\mathrm{spec}(A)$ in much the same way that a Hermitian matrix A is determined by its eigenvalues and the corresponding eigenspaces.

Problem 2.52. Verify that $\mathrm{spec}(H) = \mathbb{N}$ and $\mathrm{spec}(X) = \mathbb{R}$.

2.7.4 Time evolution

For a self-adjoint operator A, the spectral theorem allows us to extract a PVM on $\mathrm{spec}(A) \subseteq \mathbb{R}$. Using Theorem 2.12 we can define a continuous 1-parameter group of unitary operators $U_t = \exp(-itA)$ by

$$\exp(-iAt) := \int_{\mathbb{R}} e^{-iat} P(da).$$

By Stone's Theorem, every continuous 1-parameter group of unitary operators is of this form.

Theorem 2.14 (Stone, 1932)**.** *For every continuous 1-parameter group of unitary operators on $\mathcal{H}$, there exists a self-adjoint operator A such that $U_t = \exp(-itA)$.*

The proof is unfortunately beyond the scope of these notes, we refer to Theorem 5.6 in Section X of [Co07]. As in the finite-dimensional case, the self-adjoint operator H that generates time translations, $U_t = \exp(-i\frac{t}{\hbar}H)$, is interpreted as the *energy* of the system.

2.7.5 A particle moving in one dimension

Recall that a particle moving in 1 dimension is described by the Hilbert space $\mathcal{H} = L^2(\mathbb{R})$. Translation over a distance s is then described by the 1-parameter group of unitary operators

$$(T_s\psi)(x) = \psi(x - s). \tag{36}$$

By Stone's theorem, it is of the form $T_s = \exp(-isP)$, with a self-adjoint generator P that is interpreted as the *momentum* of the particle. To find P, we simply differentiate (36) at $s = 0$,

$$(P\psi)(x) = i\tfrac{d}{ds}\big|_{s=0} \exp(-isP)\psi(x) = i\tfrac{d}{ds}\big|_{s=0}\psi(x - s) = -i\tfrac{d}{dx}\psi(x).$$

We conclude that P is the self-adjoint operator $P\psi = -i\frac{d}{dx}\psi$ with domain $\mathrm{Dom}(P) = \{\psi \in L^2(\mathbb{R})\,;\, \int_{-\infty}^{\infty} |\frac{d}{dx}\psi|^2 < \infty\}$. In fact one usually scales this by $\hbar$, resulting in the expression

$$P = -i\hbar\tfrac{d}{dx}. \tag{37}$$

For a classical particle moving in one dimension under the influence of a potential $V(x)$, the energy is given by the expression $h(x,p) = \frac{1}{2m}p^2 + V(x)$. For a quantum particle in one dimension, it is postulated that the Hamilton operator H that generates the 1-parameter group of time translations is given by

$$H = \frac{1}{2m}P^2 + V, \tag{38}$$

the symmetric operator on $L^2(\mathbb{R})$ defined by

$$(H\psi)(x) = -\tfrac{\hbar^2}{2m}\tfrac{d^2}{dx^2}\psi(x) + V(x)\psi(x). \tag{39}$$

If this operator is self-adjoint, then it generates a continuous 1-parameter group U_t, interpreted as the group of time translations. For a system initially in state ψ_0, the solution $\psi_t = U_t\psi_0$ satisfies the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi_t(x) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) + V(x)\psi(x), \tag{40}$$

considered as a PDE with initial condition $\psi_0(x)$ at time zero.

Problem 2.53. A particle with mass m that moves freely on a circle with radius R is described by the Hilbert space $\mathcal{H} = L^2([0, 2\pi])$ with Hamilton operator $H = \frac{1}{2m}P^2$. The momentum operator P is given by

$$P\psi = i\frac{\hbar}{R}\frac{d}{d\theta}\psi$$

on the space $\mathcal{S} \subset \mathcal{H}$ of smooth functions $\psi\colon [0, 2\pi] \to \mathbb{C}$ whose left derivatives at 0 agrees with their right derivatives at 2π (so they are smooth on the circle).

a) Show that for every $n \in \mathbb{Z}$, the function $\psi_n(\theta) = \exp(in\theta)$ is an eigenvector for H. Determine the corresponding eigenvalue E_n.

b) Give a solution $\psi(t, \theta)$ for the Schrödinger equation with initial condition $\psi(0, \theta) = \psi_n(\theta)$.

2.7.6 The harmonic oscillator

A *harmonic oscillator* is a particle that moves on the real line under the influence of a potential $V(x) = \frac{1}{2}kx^2$. This is an important system for at least three reasons:

1) It can be solved analytically (we will outline how to do this below).

2) Even more exceptionally, many-particle systems consisting of harmonic oscillators can also be solved analytically.

3) Any potential $V(x)$ with a local minimum at x_0 can be approximated in a second order Taylor expansion by $V(x) \simeq V(x_0) + \frac{1}{2}kx^2$ with $k = V''(x_0)$. So the harmonic oscillator approximates this system well for low energies.

As a dense common domain for the linear operators on $L^2(\mathbb{R})$ that we will encounter, we take the *Schwartz space* of rapidly decaying smooth functions

$$\mathcal{S}(\mathbb{R}) = \left\{ \psi \colon \mathbb{R} \to \mathbb{C}\,;\, \sup_{x\in\mathbb{R}} |x^n \partial_x^m \psi(x)| < \infty \text{ for all } n, m \in \mathbb{N} \right\}.$$

The position and momentum operators X and P are defined on $\mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R})$ by

$$(X\psi)(x) = x\psi(x), \quad (P\psi)(x) = -i\partial_x \psi(x).$$

Since these operators map $\mathcal{S}(\mathbb{R})$ into itself, the Schwarz functions constitute an *invariant domain.* Since $(XP\psi - PX\psi)(x) = i(\partial_x x\psi - x\partial_x\psi)(x) = i\psi(x)$, we have

$$[X, P] = i$$

on $\mathcal{S}(\mathbb{R})$. (In the literature one usually finds $[X, P] = i\hbar$, but we choose units in which $\hbar = 1$.)

The harmonic oscillator with $k = 1$ has potential $V(x) = \frac{1}{2}x^2$, so for $m = 1$ the Hamilton operator for this model is

$$H = \frac{1}{2}(P^2 + X^2).$$

We show that it has eigenvalues $E_n = (n + \frac{1}{2})$, where n runs over the integers.

For this, it is convenient to introduce the creation and annihilation operators

$$A^\dagger = \frac{1}{\sqrt{2}}(X - iP), \quad A = \frac{1}{\sqrt{2}}(X + iP).$$

Problem 2.54. Show that $[A, A^\dagger] = 1$. Show that $H = A^\dagger A + \frac{1}{2}$, and conclude that $[H, A^\dagger] = A^\dagger$ and $[H, A] = -A$.
Hint: everything follows from the fact that $[X, P] = i$.

It follows that if $\psi \in \mathcal{S}(\mathbb{R})$ satisfies $H\psi = E\psi$, then

$$\begin{aligned} HA\psi &= (E-1)A\psi \\ HA^\dagger\psi &= (E+1)A^\dagger\psi. \end{aligned}$$

Indeed,

$$\begin{aligned} HA^\dagger\psi &= A^\dagger H\psi + [H, A^\dagger]\psi \\ &= EA^\dagger\psi + A^\dagger\psi = (E+1)A^\dagger\psi. \end{aligned}$$

Problem 2.55. The reasoning for A is similar.

So if $\psi \in \mathcal{S}(\mathbb{R})$ is an eigenvector of H with eigenvalue E, then $A^\dagger\psi$ is either zero or an eigenvector with eigenvalue $E + 1$. If we find a single eigenvector $\Omega \in \mathcal{S}(\mathbb{R})$, we can therefore use this to find new eigenvectors.

Problem 2.56. If Ω satisfies $A\Omega = 0$, then $H\Omega = \frac{1}{2}\Omega$.

Problem 2.57. The solutions to $A\Omega = 0$ are multiples of $\Omega(x) = \exp(-\frac{1}{2}x^2)$.

Since Ω is a Schwartz function, we have found a single eigenvalue $\Omega \in \mathcal{S}(\mathbb{R})$ for H with eigenvalue $E_0 = \frac{1}{2}$. We could try to find eigenvectors with *lower* energy by applying A, but $A\Omega = 0$ so this will not work. We can, however, apply $A^\dagger$ repeatedly to find eigenvectors $\psi_n = (A^\dagger)^n\Omega$ with eigenvalue $E_n = n + \frac{1}{2}$.

Problem 2.58. Show that $\psi_0 = e^{-\frac{1}{2}x^2}$, $\psi_1 = (\sqrt{2}x)e^{-\frac{1}{2}x^2}$, and that $\psi_2 = ((\sqrt{2}x)^2 - 1)e^{-\frac{1}{2}x^2}$ are eigenvectors with eigenvalues $E_0 = \frac{1}{2}$, $E_1 = \frac{3}{2}$, $E_2 = \frac{5}{2}$.

Problem 2.59. Show that $\psi_n(x) = p_n(\sqrt{2}x)e^{-\frac{1}{2}x^2}$, where p_n is a *monic* polynomial of degree n that satisfies the recurrence relation

$$p_{n+1}(y) = yp_n(y) - p_n'(y)$$

with $p_0 = 1$.

In particular, ψ_n is not identically zero, and $n+\frac{1}{2}$ is an eigenvalue of H. The polynomials p_n are called *Hermite polynomials.* One can show that $\psi_n/\|\psi_n\|$ is an orthonormal basis of $L^2(\mathbb{R})$, and that $\operatorname{spec}(H) = \{n + \frac{1}{2}\,; n \in \mathbb{N}\}$. Since measuring the energy of a harmonic oscillator *always* yields a value in $\operatorname{spec}(H)$, the energy levels of the harmonic oscillator are said to be *quantized.*

Problem 2.60. If the energy is measured for a harmonic oscillator is in the state $\psi \in L^2(\mathbb{R})$, then the probability that the outcome $E_n = (n + \frac{1}{2})$ occurs is

$$\frac{\left|\int_{-\infty}^{\infty} p_n(\sqrt{2}x)e^{-\frac{1}{2}x^2}\psi(x)dx\right|^2}{\int_{-\infty}^{\infty} p_n^2(\sqrt{2}x)e^{-x^2}dx}$$

Problem 2.61. In the above argument, we could try to switch the roles of A and $A^\dagger$. If we write $H = AA^\dagger - \frac{1}{2}$, then a solution $\tilde{\Omega}$ of $A^\dagger\tilde{\Omega} = 0$ would yield an eigenvalue $\tilde{E}_0 = -\frac{1}{2}$, and we could apply A repeatedly to lower the energy, yielding eigenvalues $-(n + \frac{1}{2})$. Why doesn't this work?

2.7.7 Reproducing Kernel Hilbert Spaces

Let S be a set and K a *kernel* on S, i.e. a function $K\colon S \times S \to \mathbb{C}$. The kernel K is *positive definite* if

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \overline{z}_i z_j K(x_i, x_j) \geq 0$$

for all $z_1, \dots, z_n \in \mathbb{C}$ and $x_1, \dots, x_n \in S$.

Theorem 2.15 (Kolmogorov-Aronszajn dilations)**.** *For every positive definite kernel K on S, there exists a Hilbert space $\mathcal{H}_K$ and a map $e\colon S \to \mathcal{H}_K$ such that*

$$\langle e(x), e(y)\rangle = K(x, y)$$

for all $x, y \in S$, and such that the linear span of $\{e(x), x \in S\}$ is dense in $\mathcal{H}_K$.

Proof. On the vector space $\mathcal{H}_0$ of finitely supported functions $z\colon S \to \mathbb{C}$, the sesquilinear form

$$\langle z, w\rangle := \sum_{x_i \in \mathrm{supp}(z)} \sum_{y_j \in \mathrm{supp}(w)} \overline{z}(x_i) w(y_j) K(x_i, y_j)$$

is positive semidefinite, $\langle z, w\rangle \geq 0$ for all $z, w \in \mathcal{H}_0$. Let

$$\mathcal{N} = \{z \in \mathcal{H}_0\,;\, \langle z, z\rangle = 0\}$$

be the null space. For $x \in \mathcal{N}$ and $y \in \mathcal{H}_0$, we have $|\langle x, y\rangle| \leq \|x\|\|y\| = 0$ by the Cauchy-Schwartz inequality, so the sesquilinear form $\langle [x], [y]\rangle := \langle x, y\rangle$ is well defined on the quotient space $\mathcal{H}_0/\mathcal{N}$. It is positive definite because $\langle [x], [x]\rangle = 0$ implies $\langle x, x\rangle = 0$, so $x \in \mathcal{N}$ and its class $[x]$ in $\mathcal{H}_0/\mathcal{N}$ is zero.

One can show that the completion $\mathcal{H}_K$ of the metric space $\mathcal{H}_0/\mathcal{N}$ (in the sense of [N22, Thm. D.6]) is a Hilbert space, and $e(x) := [\delta_x]$ gives the required map from S to $\mathcal{H}_K$. □

The Hilbert space $\mathcal{H}_K$ is called a *Reproducing Kernel Hilbert Space* (RKHS). Reproducing kernel Hilbert spaces are unique up to unitary isomorphism. If $(\widetilde{\mathcal{H}}_K, \widetilde{e})$ is another RKHS, then the linear map

$$U_0\colon \mathcal{H}_0 \to \widetilde{\mathcal{H}}_K, \qquad U_0(z) := \sum_{x_i \in \mathrm{supp}(z)} z(x_i)\widetilde{e}(x_i)$$

satisfies $\langle U_0(z), U_0(z)\rangle = \langle z, z\rangle$ for all $z \in \mathcal{H}_0$. In particular, $U_0(z) = 0$ for $z \in \mathcal{N}$, so U_0 defines an isometry $U_0\colon \mathcal{H}_0/\mathcal{N} \to \widetilde{\mathcal{H}}_K$. Since the linear span of $e(S)$ and $\widetilde{e}(S)$ is dense in $\mathcal{H}_K$ and $\widetilde{\mathcal{H}}_K$, respectively, the isometry U_0 extends to a unitary isomorphism $U\colon \mathcal{H}_K \to \widetilde{\mathcal{H}}_K$ that satisfies $U \circ e = \widetilde{e}$.

2.7.8 Tensor products

In order to define the tensor product of two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, we define the kernel K on the set $S = \mathcal{H}_A \times \mathcal{H}_B$ by

$$K\big((\psi_A, \psi_B), (\psi'_A, \psi'_B)\big) := \langle \psi_A, \psi'_A\rangle\langle \psi_B, \psi'_B\rangle.$$

To see that this is a positive definite kernel, let $z_i \in \mathbb{C}$, let $(\psi^i_A, \psi^i_B) \in S$ for $i = 1, \ldots n$, and let $V_A \subseteq \mathcal{H}_A$ and $V_B \subseteq \mathcal{H}_B$ be the *finite dimensional* subspaces spanned by ψ^i_A and ψ^i_B, respectively. Then

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \overline{z}^i z_j K\big((\psi^i_A, \psi^i_B); (\psi^j_A, \psi^j_B)\big) = \Bigg\|\sum_{i=1^n} z_i \psi^i_A \otimes \psi^i_B\Bigg\|^2$$

is the norm squared of the vector $\sum_{i=1^n} z_i\psi^i_A \otimes \psi^i_B$, considered as an element of the *finite dimensional* Hilbert space $V_A \otimes V_B$. In particular, it is nonnegative. The RKHS for this particular kernel is the *completed* tensor product $\mathcal{H}_A \overline{\otimes} \mathcal{H}_B$ of the two Hilbert spaces. The map e is bilinear, and denote by $\overline{\otimes}\colon \mathcal{H}_A \times \mathcal{H}_B \to \mathcal{H}_A \overline{\otimes} \mathcal{H}_B$.

Remark 2.10. The completed tensor product $\mathcal{H}_A \overline{\otimes} \mathcal{H}_B$ of two Hilbert spaces is again a Hilbert space, something which is not always true for the ordinary tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$. For instance, $L^2(\mathbb{R}) \overline{\otimes} L^2(\mathbb{R}) = L^2(\mathbb{R}^2)$, whereas $L^2(\mathbb{R}) \otimes L^2(\mathbb{R}) \subset L^2(\mathbb{R}^2)$ is the proper subspace of functions that can be expressed as a *finite* sum $\psi(x, y) = \sum_{i=1}^{n} \psi_A(x)\psi_B(y)$. In the litarature, it is common to denote the *completed* tensor product by $\mathcal{H}_A \otimes \mathcal{H}_B$, because the ordinary tensor product is rarely used in this context.

2.8 Symmetries and conserved quantities

Quantum systems often come with a group of symmetries. We investigate how to model symmetry in quantum systems, and we will see how continuous groups of symmetries (*Lie groups*) give rise to conserved quantities. This can be used to decompose the system into smaller subsystems, which can drastically simplify the Schrödinger equation.

2.8.1 Conserved quantities

An observable $A = A^\dagger$ on a finite dimensional Hilbert space is called a *conserved quantity* if it commutes with the Hamilton operator,

$$[A, H] = 0. \tag{41}$$

This is the case if and only if $[A, U_t] = 0$ for all $t \in \mathbb{R}$, since the time evolution is given by $U_t = \exp(-i\frac{t}{\hbar}H)$. Since $[A, U_t] = 0$ if and only if $[P_a, U_t] = 0$ for all spectral projections P_a of A, we conclude that the eigenspaces $V_a = P_a\mathcal{H}$ are invariant under the time evolution,

$$U_t V_a = U_t P_a \mathcal{H} = P_a U_t \mathcal{H} = P_a \mathcal{H} = V_a.$$

The time evolution $U_t \colon \mathcal{H} \to \mathcal{H}$ therefore decomposes into 'blocks' $U_t^a \colon V_a \to V_a$ corresponding to the various eigenspaces $V_a \subseteq \mathcal{H}$,

$$U_t = \left(\begin{array}{c|c|c} U_t^{a_1} & \mathbf{0} & \\ \hline \mathbf{0} & \ddots & \mathbf{0} \\ \hline & \mathbf{0} & U_t^{a_n} \end{array}\right)$$

The smaller the blocks, the easier it is to solve Schrödinger's equation, numerically as well as analytically. If A and B are *commuting* conserved quantities, then one can of course apply the same trick to the joint eigenspaces $V_{a,b} = \{\psi \in \mathcal{H}\,;\, A\psi = a\psi \text{ and } B\psi = b\psi\}$.

The name *conserved quantity* comes from the fact that the probability of a measurement outcome a is time-independent;

$$\langle \psi_t, P_a \psi_t \rangle = \langle U_t \psi, P_a U_t \psi \rangle = \langle \psi_0, U_t^\dagger P_a U_t \psi_0 \rangle = \langle \psi_0, U_t^\dagger U_t P_a \psi_0 \rangle = \langle \psi_0, P_a \psi_0 \rangle.$$

Problem 2.62. For the simple case of 2 particles, the Hamiltonian for the Heisenberg XXX-model on $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ is

$$H = -\frac{J}{2}(\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z).$$

a) Using Problem 2.21 or otherwise, show that the angular momentum $S_z := \sigma_z \otimes \mathbf{1} + \mathbf{1} \otimes \sigma_z$ is conserved.

b) Show that $\mathrm{spec}(S_z) = \{-2, 0, 2\}$, give the corresponding eigenspaces V_{-2}, V_0 and V_2, and express H as a matrix with respect to an orthonormal basis of eigenvectors for S_z. Use this to solve the Schrödinger equation.

Problem 2.63. The Heisenberg XXX model on the n-fold tensor product $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ is given by the Hamiltonian

$$H = -\frac{J}{2}\sum_{i=1}^{n-1} \sigma_{i,x}\sigma_{i+1,x} + \sigma_{i,y}\sigma_{i+1,y} + \sigma_{i,z}\sigma_{i+1,z},$$

where $\sigma_{i,x}$ is short for $\mathbf{1} \otimes \ldots \otimes \mathbf{1} \otimes \sigma_x \otimes \mathbf{1} \otimes \ldots \otimes \mathbf{1}$ with σ_x on position i.

a) The three angular momenta $S_x = \sum_{i=1}^n \sigma_{i,x}$, $S_y = \sum_{i=1}^n \sigma_{i,y}$ and $S_z = \sum_{i=1}^n \sigma_{i,z}$ are conserved.

b) The total angular momentum $J^2 := S_x^2 + S_y^2 + S_z^2$ commutes with S_z.

c) Conclude that the time evolution respects the joint eigenspace decomposition for J^2 and S_z.

2.8.2 Symmetries and unitary representations

If the symmetries of a quantum system are modelled by an abstract group G, then it seems reasonable to model the action of G on the Hilbert space $\mathcal{H}$ by a *unitary representation.*

Definition 2.17 (Unitary representation)**.** A unitary representation of a group G on a Hilbert space $\mathcal{H}$ is a group homomorphism $\pi\colon G \to \mathrm{U}(\mathcal{H})$ into the group $\mathrm{U}(\mathcal{H})$ of unitary operators on $\mathcal{H}$.

Unravelling the definition, this means that:

1) The identity $e \in G$ acts trivially, $\pi(e)\psi = \psi$ for all $\psi \in \mathcal{H}$.

2) The group acts unitarily, $\langle \pi(g)\psi, \pi(g)\psi \rangle = \langle \psi, \psi \rangle$ for all $g, h \in G$ and $\psi \in \mathcal{H}$.

3) Multiplication is respected, $\pi(gh)\psi = \pi(g)\pi(h)\psi$ for all $g, h \in G$ and $\psi \in \mathcal{H}$.

The representation is said to be a *symmetry* if the unitary transformations commute with the Hamilton operator, $[\pi(g), H] = 0$ for all $g \in G$. Equivalently, we can ask that $\pi(g)$ commutes with the time translations,

$$U_t\pi(g) = \pi(g)U_t.$$

This means that *first* transforming the system by the symmetry operation $\pi(g)$ and *then* running the time evolution for time t is equivalent to *first* running the time evolution and *then* acting with the symmetry transformation.

For some applications – including the important case of the group of rotations acting on a qubit – the above definition of a unitary representation is too restrictive. Recall (§2.3.2, §2.4) that two unitary operators U and $U' = e^{i\phi}U$ that differ by a phase represent the same physical transformation. The group of physical transformations is therefore not the group $\mathrm{U}(\mathcal{H})$ of unitary operators, but its quotient $\mathrm{PU}(\mathcal{H}) = \mathrm{U}(\mathcal{H})/\mathbb{T}$ by the normal subgroup $\mathbb{T} = \{e^{i\phi}\mathbf{1}\,;\,\phi \in [0, 2\pi)\}$ of unitary operators that act by a phase factor. A *projective unitary representation* assigns to every $g \in G$ a physical transformation $\overline{\pi}(g) \in \mathrm{PU}(\mathcal{H})$.

Definition 2.18 (Projective unitary representation)**.** A *projective* unitary representation of a group G on a Hilbert space $\mathcal{H}$ is a group homomorphism $\overline{\pi}\colon G \to \mathrm{PU}(\mathcal{H})$ from G into the group $\mathrm{PU}(\mathcal{H})$ of *projective* unitary operators.

Needless to say, every unitary representation $\pi\colon G \to \mathrm{U}(\mathcal{H})$ gives rise to a projective unitary representation $\overline{\pi}\colon G \to \mathrm{PU}(\mathcal{H})$ by $\overline{\pi}(g) = [\pi(g)]$. However, if we start with a projective unitary transformation and choose a representative $\pi(g) \in \mathrm{U}(\mathcal{H})$ for every class $\overline{\pi}(g) \in \mathrm{PU}(\mathcal{H})$, then in general $\pi(gh) \neq \pi(g)\pi(h)$.

2.8.3 Lie theory

It appears to be a fundamental property of Nature that continuous symmetries give rise to conserved quantities. In *Lie theory*, continuous groups of symmetries are modelled by *Lie groups*, and the corresponding algebra of conserved quantities is a *Lie algebra.*

If one is willing to use a little differential geometry, then the definitions of Lie groups and Lie algebras are not hard to state. A Lie group G is a group which is at the same time a smooth manifold, and one requires that the group multiplication $(g, h) \mapsto gh$ is smooth. The corresponding *Lie algebra* $\mathfrak{g}$ is the tangent space to G at the identity. Unlike the Lie group G, the Lie algebra $\mathfrak{g}$ is a *vector space*, which makes it considerably easier to handle. It comes with a bilinear map $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ called the *Lie bracket*, denoted $(X, Y) \mapsto [X, Y]$. It is skew-symmetric, $[Y, X] = -[X, Y]$, and satisfies the Jacobi identity

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Rather than developing Lie theory in general (which would be the topic for an entire course), we will illustrate some of the main points fot the concrete example of the rotation group $G = \mathrm{SO}(3)$.

2.8.4 The Lie group $\mathrm{SO}(3)$ and its Lie algebra $\mathfrak{so}(3)$

Let $\mathrm{SO}(3)$ be the group of real orthogonal transformations of $\mathbb{R}^3$ with determinant one,

$$\mathrm{SO}(3) = \{g \in M_3(\mathbb{R})\,;\, g^T g = \mathbf{1}, \det(g) = 1\}.$$

The following proposition explains why SO(3) is sometimes called the *rotation group.*

Proposition 2.16. *For every $g \in \mathrm{SO}(3)$, there exists a unit vector $\vec{v} \in \mathbb{R}^3$ and an angle θ such that g is the rotation around the axis $\mathbb{R}\vec{v}$ over angle θ.*

Proof. Since g is orthogonal with $\det(g) = 1$, its eigenvalues λ_1, λ_2 and λ_3 are complex numbers on the unit circle with $\lambda_1\lambda_2\lambda_3 = 1$. Since the characteristic polynomial of g has real coefficients, the complex conjugate of every eigenvalue is an eigenvalue again. It follows that at least one of these eigenvalues is 1. Indeed, if g has an eigenvalue (say λ_2) which is not real, then $\overline{\lambda}_2$ is a different eigenvalue (say λ_3), so the remaining eigenvalue is

$$\lambda_1 = \lambda_1|\lambda_2|^2 = \lambda_1\lambda_2\lambda_3 = 1.$$

Alternatively, if all eigenvalues of g are real, then they are all in $\{\pm 1\}$. Since their product is 1, at least one of them is $+1$. Summarizing, we may assume that $\lambda_1 = 1$, and $\lambda_2 = e^{i\theta}$, $\lambda_3 = e^{-i\theta}$ for some angle $\theta \in [0, 2\pi)$. In fact, we may assume that $\theta \in [0, \pi]$ by interchanging λ_2 and λ_3 if necessary.

Since $\lambda_1 = 1$ is real, we can choose a real eigenvector $\vec{v}_1 \in \mathbb{R}^3$ with $g\vec{v}_1 = \vec{v}_1$. Then g fixes the axis $\mathbb{R}\vec{v}_1$. It restricts to an orthogonal linear transformation of the orthogonal complement $\vec{v}_1^{\perp}$ because $\vec{w} \perp \vec{v}_1$ implies $g\vec{w} \perp g\vec{v}_1 = \vec{v}_1$.

If λ_2 and λ_3 are real, then either $\lambda_2 = \lambda_3 = 1$ and g is the identity, or $\lambda_2 = \lambda_3 = -1$ and g is the rotation around $\mathbb{R}\vec{v}_1$ over an angle $\theta = \pi$. We may therefore assume without loss of generality that $\lambda_2 = e^{i\theta}$ and $\lambda_3 = e^{-i\theta}$ with $e^{-i\theta} \notin \mathbb{R}$. Let $\vec{v}_+ \in \mathbb{C}^3$ be a *complex* unit eigenvector with eigenvalue $e^{i\theta}$. Then its complex conjugate $\vec{v}_-$ is a complex eigenvector with eigenvalue $e^{-i\theta}$. Since these eigenvalues are distinct, the eigenvectors $\vec{v}_+$ and $\vec{v}_-$ are orthogonal. It follows that the *real* vectors $\vec{v}_2 = \frac{1}{\sqrt{2}}(\vec{v}_+ + \vec{v}_-)$ and $\vec{v}_3 = \frac{i}{\sqrt{2}}(\vec{v}_+ - \vec{v}_-)$ form an orthonormal basis of the 2-dimensional vector space $\vec{v}_1^{\perp}$.

One verifies that the matrix of g with respect to the basis $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix},$$

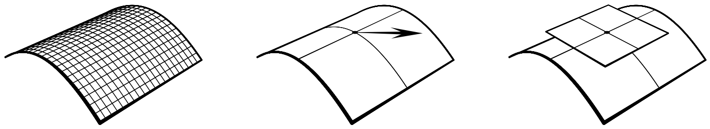
so g represents rotation around the axis $\mathbb{R}\vec{v}_1$ over an angle θ. □

Problem 2.64. Verify that SO(3) is a group. Conclude that the composition of two rotations (around possibly different axes) is again a rotation.

To find the Lie algebra of the Lie group SO(3), we consider SO(3) as a 3-dimensional hypersurface in the 9-dimensional vector space $M_3(\mathbb{R})$. In general, let $\Sigma \subseteq V$ be a (possibly nonlinear) hypersurface in a vector space V. Then its *tangent space* $T_p\Sigma$ at the point $p \in \Sigma$ is the set of all tangent vectors at p to smooth curves $\gamma\colon \mathbb{R} \to \Sigma$ that pass through p,

$$T_p\Sigma := \{\gamma'(0) \in V\ ;\ \gamma \text{ smooth with } \gamma(0) = p\}.$$

The dimension of Σ is the same as the dimension of its tangent spaces.



Proposition 2.17. *The Lie algebra of* SO(3) *is the vector space* $\mathfrak{so}(3) \subseteq M_3(\mathbb{R})$ *of skew-symmetric linear transformations,*

$$\mathfrak{so}(3) = \{X \in M_3(\mathbb{R})\,;\, X + X^T = 0\}.$$

Proof. By definition, the Lie algebra $\mathfrak{so}(3)$ of SO(3) is the tangent space to $\mathrm{SO}(3) \subseteq M_3(\mathbb{R})$ at the identity $\mathbf{1} \in \mathrm{SO}(3)$. So a matrix $X \in M_3(\mathbb{R})$ is in $\mathfrak{so}(3)$ if it is the tangent vector $X = g'(0)$ to some smooth curve $t \mapsto g(t)$ with $g(0) = \mathbf{1}$, and with $g(t) \in \mathrm{SO}(3)$ for all $t \in \mathbb{R}$.

Since $g(t)^T g(t) = \mathbf{1}$, the product rule yields

$$0 = \big(g(t)^T g(t)\big)' = g'(0)^T g(0) + g(0)^T g'(0) = g'(0)^T + g'(0),$$

so every tangent vector $X := g'(0)$ satisfies $X + X^T = 0$. Conversely, if $X + X^T = 0$, then the smooth curve $g(t) := \exp(tX)$ satisfies $g(0) = \mathbf{1}$ and $g(t)^T g(t) = \exp(tX^T)\exp(tX) = \exp(-tX)\exp(tX) = \mathbf{1}$, so $g(t)$ is a curve of orthogonal transformations with $g'(0) = \frac{d}{dt}\exp(tX)|_0 = X$. Since $\mathbf{tr}(X) = 0$, we have $\det(g(t)) = \det(\exp(tX)) = \exp(\mathbf{tr}(tX)) = 1$, so $g(t) \in \mathrm{SO}(3)$ for all t. □

Problem 2.65. If $A \in M_n(\mathbb{C})$ is normal, then $\det(\exp(A)) = \exp(\mathbf{tr}(A))$.

Note that $\mathfrak{so}(3)$ is a vector subspace of $M_3(\mathbb{R})$ that is closed under the commutator bracket; if $X, Y \in \mathfrak{so}(3)$, then $[X, Y] = XY - YX$ is again an element of $\mathfrak{so}(3)$. Every element $X \in \mathfrak{so}(3)$ can be written as

$$X = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} \tag{42}$$

for some $x, y, z \in \mathbb{R}$. Since $X\vec{v} = 0$ for $\vec{v} = (x, y, z)$, we have $\exp(tX)\vec{v} = \vec{v}$ for all $t \in \mathbb{R}$. The Lie algebra element X therefore generates a one-parameter group $g_t = \exp(tX)$ of rotations around the axis $\vec{v} = (x, y, z)$.

Proposition 2.18. *The exponential map* $\exp\colon \mathfrak{so}(3) \to \mathrm{SO}(3)$ *is surjective.*

In other words, for every $g \in \mathrm{SO}(3)$, there exists an $X \in \mathfrak{so}(3)$ such that $g = \exp(X)$. It turns out that this is a general property of compact Lie groups.

Problem 2.66. Prove Proposition 2.18. One way to do this is by showing that $X \in \mathfrak{so}(3)$ as in (42) satisfies $X\vec{w} = \vec{v} \times \vec{w}$ for $\vec{v} = (x, y, z)$. Find an orthonormal basis $\vec{v}_1, \vec{v}_2, \vec{v}_3$ of $\mathbb{R}^3$ such that $X\vec{v}_1 = 0$, $X\vec{v}_2 = \|\vec{v}\|\vec{v}_3$ and $X\vec{v}_3 = -\|\vec{v}\|\vec{v}_2$, and infer that $\exp(tX)$ is the rotation around $\mathbb{R}\vec{v}$ over an angle $\theta = t\|\vec{v}\|$.

2.8.5 Conserved quantities for unitary representations of SO(3)

A unitary representation $\pi\colon \mathrm{SO}(3) \to \mathrm{U}(\mathcal{H})$ is called *continuous* if $g \mapsto \pi(g)\psi$ is a continuous map $\mathrm{SO}(3) \to \mathcal{H}$ for every $\psi \in \mathcal{H}$. For every $X \in \mathfrak{so}(3)$, the one-parameter group of unitary transformations

$$U_\theta = \pi(\exp(\theta X))$$

is then continuous as well. By Stone's theorem, there exists a unique self-adjoint operator L_X on $\mathcal{H}$ that generates the one-parameter group $V_\theta = \pi(\exp(\theta X))$,

$$\pi(\exp(\theta X)) = \exp(-i\theta L_X)$$

for all $\theta \in \mathbb{R}$. If X generates rotations around the axis $\mathbb{R}\vec{v}$, then the observable L_X is interpreted as *angular momentum* in the $\vec{v}$-direction.

The following simple result provides the link between symmetries and conserved quantities for the rotation group SO(3).

Theorem 2.19. *If the continuous unitary representation $\pi\colon \mathrm{SO}(3) \to \mathrm{U}(\mathcal{H})$ is a symmetry of H, then the angular momenta L_X are conserved.*

Proof. If the representation π is a symmetry of the Hamilton operator H, then $[\pi(g), U_t] = 0$ for all $g \in \mathrm{SO}(3)$, so in particular $[\pi(\exp(\theta X)), U_t] = 0$. It follows that $[\exp(-i\theta L_X), U_t] = 0$ for all θ, and differentiating in θ we conclude that $[L_X, U_t] = 0$ for all $t \in \mathbb{R}$. □

In the following series of problems, we investigate the angular momenta for the continuous unitary representation of SO(3) on the state space $\mathcal{H} = L^2(\mathbb{R}^3)$ for a single particle moving in $\mathbb{R}^3$.

Problem 2.67. For every $g \in \mathrm{SO}(3)$, the linear map

$$\pi(g)\colon L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3), \qquad (\pi(g)\psi)(\vec{x}) = \psi(g^{-1}\vec{x}) \tag{43}$$

is unitary, and $\pi\colon \mathrm{SO}(3) \to U(L^2(\mathbb{R}^3))$ is a unitary representation of SO(3).

We first determine the angular momentum operator L_z corresponding to the Lie algebra element

$$X_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Problem 2.68. The element $X_z \in \mathfrak{so}(3)$ generates the 1-parameter group $g(\theta) = \exp(\theta X_z)$ of rotations around the z-axis,

$$g(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It follows that $L_z\psi(x, y, z) = i\frac{d}{d\theta}|_0\psi(g(-\theta)(x, y, z)^T)$ is given by

$$L_z\psi = -i(x\tfrac{\partial}{\partial y} - y\tfrac{\partial}{\partial x})\psi.$$

Problem 2.69. Similarly, for the generators

$$X_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad X_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad X_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

of rotation around the x, y and z axis, the angular momenta are

$$\begin{aligned} L_x &= -i(y\tfrac{\partial}{\partial z} - z\tfrac{\partial}{\partial y}) && (44) \\ L_y &= -i(z\tfrac{\partial}{\partial x} - x\tfrac{\partial}{\partial z}) && (45) \\ L_z &= -i(x\tfrac{\partial}{\partial y} - y\tfrac{\partial}{\partial x}). && (46) \end{aligned}$$

For a particle moving in $\mathbb{R}^3$ under the influence of a potential $V(x, y, z)$, the Hamilton operator on $L^2(\mathbb{R}^3)$ is given by

$$H\psi = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi + V\psi. \tag{47}$$

Problem 2.70. If the potential $V(x, y, z)$ depends only on the radius $r = \sqrt{x^2 + y^2 + z^2}$, then the unitary representation (43) is a symmetry for the Hamiltonian (47). In that case, the angular momenta L_x, L_y and L_z are conserved.

Problem 2.71. If $L^2 := L_x^2 + L_y^2 + L_z^2$, then $[L_z, L^2] = 0$. So L^2 and L_z are *commuting* conserved quantities.

2.8.6 The Lie group SU(2) and its Lie algebra $\mathfrak{su}(2)$

Let SU(2) be the group of complex unitary transformations of $\mathbb{C}^2$ with determinant 1,

$$\mathrm{SU}(2) = \{U \in M_2(\mathbb{C})\,;\, U^\dagger U = \mathbf{1},\, \det(U) = 1\}.$$

To find its Lie algebra, we consider SU(2) as a hypersurface in the 4-dimensional complex vector space $M_2(\mathbb{C})$.

Proposition 2.20. *The Lie algebra of* SU(2) *is the real vector space*

$$\mathfrak{su}(2) = \{X \in M_2(\mathbb{C})\,;\, X + X^\dagger = 0,\, \mathbf{tr}(X) = 0\}.$$

Proof. We proceed as in the proof of Proposition 2.17. Suppose that $X = U'(0)$ for a smooth curve $t \mapsto U(t)$ in SU(2) with $U(0) = \mathbf{1}$. Then $\mathbf{1} = U(t)^\dagger U(t)$ for all $t \in \mathbb{R}$. Differentiation at $t = 0$ yields $0 = U'(0)^\dagger U(0) + U(0)^\dagger U'(0) = X^\dagger \mathbf{1} + \mathbf{1}X$, so $X + X^\dagger = 0$. Since $\det(U(t)) = 1$, we derive from $\frac{d}{dt}\det(U(t))|_{t=0} = \mathbf{tr}(U'(0))$ that $\mathbf{tr}(X) = 0$.

Conversely, suppose that $X^\dagger + X = 0$ and $\mathbf{tr}(X) = 0$. Since iX is Hermitian, the curve $U(t) = \exp(tX)$ is a unitary one-parameter group with generator $U'(0) = X$ by Stone's Theorem. Since $\det(U(t)) = \det(\exp(tX)) = \exp(t\mathbf{tr}(X)) = 1$, the smooth curve $t \mapsto U_t$ lies entirely within SU(2). □

Problem 2.72. If $t \mapsto A(t)$ is a differentiable curve in $M_2(\mathbb{C})$ with $A(0) = \mathbf{1}$, then $\frac{d}{dt}|_0 \det(A(t)) = \mathbf{tr}(\frac{d}{dt}|_0 A(t))$.

Note that $\mathfrak{su}(2)$ is a real linear subspace of $M_2(\mathbb{C})$ that is closed under the commutator bracket; if $X, Y \in \mathfrak{su}(2)$, then $[X, Y] = XY - YX$ is again an element of $\mathfrak{su}(2)$.

Problem 2.73. Verify that SU(2) is a group, and that $\mathfrak{su}(2)$ is closed under the commutator bracket.

Lemma 2.21. *The exponential map* $\exp\colon \mathfrak{su}(2) \to \mathrm{SU}(2)$ *is surjective.*

Proof. Since u is unitary its eigenvalues are complex numbers of modulus 1, $\lambda_1 = e^{i\phi}$ and $\lambda_2 = e^{i\theta}$. Since $\det(u) = 1$, we have $\lambda_1\lambda_2 = 1$, so $\lambda_1 = e^{i\phi}$ and $\lambda_2 = e^{-i\phi}$. Since $u = e^{i\phi}P_1 + e^{-i\phi}P_2$, the skew-Hermitian operator $X = i\phi P_1 - i\phi P_2$ satisfies $\exp(X) = u$. We can always choose P_1 and P_2 to be orthogonal projections of rank 1, so that $\mathbf{tr}(X) = 0$. □

2.8.7 Rotation of qubits: the spin homomorphism

The natural action of the rotation group SO(3) on the state space $\mathcal{H} = \mathbb{C}^2$ of a qubit is not by a unitary representation $\pi\colon \mathrm{SO}(3) \to \mathrm{U}(\mathbb{C}^2)$, but by a *projective* unitary transformation $\overline{\pi}\colon \mathrm{SO}(3) \to \mathrm{PU}(\mathbb{C}^2)$.

The key to understanding this action is the *spin homomorphism*, a continuous group homomorphism $s\colon \mathrm{SU}(2) \to \mathrm{SO}(3)$ which is surjective with kernel $\{\pm\mathbf{1}\}$. By the first isomorphism theorem, the spin homomorphism yields an *isomorphism* $\mathrm{SU}(2)/\{\pm\mathbf{1}\} \to \mathrm{SO}(3)$. The inverse of this isomorphism yields the required projective representation: for $g \in \mathrm{SO}(3)$, we define $\overline{\pi}(g) = [u]$ if $u \in \mathrm{SU}(2)$ and $s(u) = g$. Note that although u is not uniquely determined by g, it is uniquely determined *up to sign*. Indeed, if $s(u) = g$ and $s(u') = g$, then $s(u^{-1}u') = \mathbf{1}$, so $u^{-1}u'$ is an element of the kernel $\{\pm\mathbf{1}\}$ of s. Since $u^{-1}u' = \pm\mathbf{1}$, we have $u' = \pm u$, and both unitaries define the same class $[u] \in \mathrm{PU}(\mathbb{C}^2)$.

For the construction of the spin homomorphism $s\colon \mathrm{SU}(2) \to \mathrm{SO}(3)$, we require a number of lemmas. First, it is convenient to identify $\mathbb{R}^3$ with the 3-dimensional real vector space

$$\mathrm{Herm}_2^0 := \{A \in M_2(\mathbb{C})\,;\, A^\dagger = A, \mathbf{tr}(A) = 0\}$$

of observables with vanishing trace. Recall that any observable $A = A^\dagger$ can be expressed as $A = t\mathbf{1} + x\sigma_x + y\sigma_y + z\sigma_z$. Since the Pauli matrices have trace zero, $\mathbf{tr}(A) = 0$ implies that $t = 0$, so every $A \in \mathrm{Herm}_2^0$ can be written as

$$A = x\sigma_x + y\sigma_y + z\sigma_z$$

for a vector $(x, y, z) \in \mathbb{R}^3$. The resulting identification of $\mathbb{R}^3$ with Herm_2^0 respects the natural inner product

$$(A, B) := \tfrac{1}{2}\mathbf{tr}(AB)$$

on Herm_2^0. Indeed, if $A = x_A\sigma_x + y_A\sigma_y + z_A\sigma_z$ and $B = x_B\sigma_x + y_B\sigma_y + z_B\sigma_z$, then

$$(A, B) = x_A x_B + y_A y_B + z_A z_B. \tag{48}$$

Problem 2.74. Verify that $\mathbf{tr}(\sigma_i\sigma_j) = 2\delta_{i,j}$, and conclude that (48) holds.

For $U \in \mathrm{SU}(2)$, we define the linear transformation $s(U)\colon \mathrm{Herm}_2^0 \to \mathrm{Herm}_2^0$ by conjugation with U,

$$s(U)\colon A \mapsto UAU^\dagger. \tag{49}$$

To see that $s(U)$ maps Herm_2^0 to itself, note that if A is Hermitian, then $UAU^\dagger$ is again Hermitian, and if $\mathbf{tr}(A) = 0$ then also $\mathbf{tr}(UAU^\dagger) = 0$ because the trace depends only on the eigenvalues of A, and those are invariant under conjugation.

Lemma 2.22. *The linear map $s(U)\colon \mathrm{Herm}_2^0 \to \mathrm{Herm}_2^0$ is an orthogonal transformation of determinant 1.*

Proof. To see that it is orthogonal, we use $U^\dagger U = \mathbf{1}$ together with the cyclic property $\mathbf{tr}(XY) = \mathbf{tr}(YX)$ of the trace. For all $A, B \in \mathrm{Herm}_2^0$, we have

$$(s(U)A, s(U)B) = \tfrac{1}{2}\mathbf{tr}(UAU^\dagger UBU^\dagger) = \tfrac{1}{2}\mathbf{tr}(UABU^\dagger) = \tfrac{1}{2}\mathbf{tr}(AB) = (A, B).$$

To see that $s(U)$ has determinant 1, note first that $\det(s(U)) = \pm 1$ because $s(U)$ is an orthogonal transformation. Recall from Lemma 2.21 that every $U \in \mathrm{SU}(2)$ can be written as $U = \exp(X)$ for some $X \in \mathfrak{su}(2)$. Since $t \mapsto \det\big(s(\exp(tX))\big)$ is a continuous function that takes only the values 1 and -1, its value for $t = 0$ is the same as its value at $t = 1$. So $\det(s(U)) = \det(s(\mathbf{1})) = 1$ for all $U \in \mathrm{SU}(2)$. □

If we identify Herm_2^0 with $\mathbb{R}^3$, we can therefore consider $s(U)$ as an element of $\mathrm{SO}(3)$. This allows us to define the spin homomorphism as follows.

Lemma 2.23. *The map $U \mapsto s(U)$ is a group homomorphism $s\colon \mathrm{SU}(2) \to \mathrm{SO}(3)$ with kernel $\{\pm\mathbf{1}\}$.*

Proof. The previous proposition shows that under the identification of Herm_2^0 with $\mathbb{R}^3$, the linear map $s(U)\colon \mathrm{Herm}_2^0 \to \mathrm{Herm}_2^0$ corresponds to an element of $\mathrm{SO}(3)$. To see that the map $U \mapsto s(U)$ is a group homomorphism, note that

$$s(U)s(V)A = U(VAV^\dagger)U^\dagger = (UV)A(UV)^\dagger = s(UV)A$$

for all $A \in \mathrm{Herm}_2^0$. It remains to prove that $\mathrm{Ker}(s) = \{\pm\mathbf{1}\}$. Note that $U \in \mathrm{Ker}(s)$ if and only if $s(U)\colon \mathrm{Herm}_2^0 \to \mathrm{Herm}_2^0$ is the identity transformation, i.e., if $UAU^\dagger = A$ for all $A \in \mathrm{Herm}_2^0$. Equivalently, $U \in \mathrm{Ker}(s)$ if and only if

$$UA = AU \tag{50}$$

for all $A \in \mathrm{Herm}_2^0$. Then U commutes with all Hermitian operators $A = A^\dagger$ because (50) is trivially satisfied for $A = \mathbf{1}$. In fact, U even commutes with arbitrary complex 2×2-matrices, since any $A \in M_2(\mathbb{C})$ is a complex linear combination $A = \frac{1}{2}(A + A^\dagger) + i\frac{1}{2i}(A - A^\dagger)$ of two Hermitian matrices. Since $[U, A] = 0$ for every $A \in M_2(\mathbb{C})$, it follows that $U = \lambda\mathbf{1}$ is a multiple of the identity, with $|\lambda| = 1$ because U is unitary and $\lambda^2 = 1$ because $\det U = 1$. So $\lambda = \pm 1$, and $U = \pm\mathbf{1}$. □

Problem 2.75. If $X \in M_n(\mathbb{C})$ satisfies $[X, Y] = 0$ for all $Y \in M_n(\mathbb{C})$, then $X = \lambda\mathbf{1}$ for some $\lambda \in \mathbb{C}$. *Hint: take* $Y = |\chi\rangle\langle\psi|$ *and consider* $[X, Y]\psi$.

For $X \in \mathfrak{su}(2)$, we define the linear map $\mathrm{ad}_X \colon \mathrm{Herm}_2^0 \to \mathrm{Herm}_2^0$ by

$$\mathrm{ad}_X(A) = [X, A]. \tag{51}$$

Note that if A is Hermitian, then $\mathrm{ad}_X(A)$ is Hermitian as well, since $[X, A]^\dagger = [A^\dagger, X^\dagger] = [A, -X] = [X, A]$. Further, $\mathbf{tr}(\mathrm{ad}_X(A)) = 0$ because $\mathbf{tr}(XA) = \mathbf{tr}(AX)$. The linear map ad_X is skew-symmetric with respect to the inner product on Herm_2^0, since

$$(\mathrm{ad}_X(A), B) = \tfrac{1}{2}\mathbf{tr}(XAB - AXB) = \tfrac{1}{2}\mathbf{tr}(ABX - AXB) = -(A, \mathrm{ad}_X(B)).$$

It follows that $X \mapsto \mathrm{ad}_X$ is a linear map from $\mathfrak{su}(2)$ to $\mathfrak{so}(3)$.

A linear map between two Lie algebras is called a *Lie algebra homomorphism* if it respects the commutator bracket, and a *Lie algebra isomorphism* if it is also bijective. It is not hard to verify that $\mathrm{ad} \colon \mathfrak{su}(2) \to \mathfrak{so}(3)$ is a Lie algebra isomorphism.

Lemma 2.24. *The map* $\mathrm{ad} \colon \mathfrak{su}(2) \to \mathfrak{so}(3)$ *is a linear isomorphism that satisfies* $\mathrm{ad}_{[X,Y]} = [\mathrm{ad}_X, \mathrm{ad}_Y]$ *for all* $X, Y \in \mathfrak{su}(2)$.

Proof. Evaluated on A, the equality $\mathrm{ad}_{[X,Y]}(A) = [\mathrm{ad}_X, \mathrm{ad}_Y](A)$ is precisely the Jacobi identity $[[X, Y], A] = [X, [Y, A]] - [Y, [X, A]]$, which can be verified in a straightforward manner.

To see that $\mathrm{ad} \colon \mathfrak{su}(2) \to \mathfrak{so}(3)$ is injective, suppose that $\mathrm{ad}_X = 0$. Then $[X, A] = 0$ for all $A \in \mathrm{Herm}_2^0$, and hence for all $A \in M_2(\mathbb{C})$ by the same argument as in Lemma 2.23. So $X = \lambda\mathbf{1}$ is a multiple of the identity, and $\lambda = 0$ because $\mathbf{tr}(X) = 0$. Since both $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are 3-dimensional, injectivity of the linear map $X \mapsto \mathrm{ad}_X$ implies that it is surjective as well. □

It turns out that $\mathrm{ad} \colon \mathfrak{su}(2) \to \mathfrak{so}(3)$ is precisely the derivative at $\mathbf{1} \in \mathrm{SU}(2)$ of the continuous group homomorphism $s \colon \mathrm{SU}(2) \to \mathrm{SO}(3)$. Indeed, for $X \in \mathfrak{su}(2)$ we have

$$\tfrac{d}{dt}|_0 s(e^{tX})A = \tfrac{d}{dt}|_0 e^{tX} A e^{-tX} = XA - AX = \mathrm{ad}_X(A). \tag{52}$$

This is a general feature in Lie theory: the derivative at the identity of a continuous homomorphism of Lie groups is always a Lie algebra homomorphism.

Theorem 2.25. *The spin homomorphism* $s \colon \mathrm{SU}(2) \to \mathrm{SO}(3)$ *is surjective with kernel* $\{\pm\mathbf{1}\}$.

Proof. The only thing left to show is that s is surjective. We claim that $s(\exp(X)) = \exp(\mathrm{ad}_X)$ for all $X \in \mathfrak{su}(2)$. From this the surjectivity of s immediately follows. Indeed, by Proposition 2.18 every $g \in \mathrm{SO}(3)$ is of the form $g = \exp(Y)$ for some $Y \in \mathfrak{so}(3)$. By Lemma 2.24, there exists an $X \in \mathfrak{su}(2)$ with $\mathrm{ad}_X = Y$, so the claim implies that $g = \exp(Y) = s(\exp(X))$ is in the image of s.

To prove the claim, note that both $U_t := \exp(t\,\mathrm{ad}_\xi)$ and $V_t := s(\exp(tX))$ are continuous 1-parameter groups of orthogonal transformations of $\mathbb{R}^3$. For U_t this is immediate, and for V_t this follows from the fact that s is a continuous group homomorphism, $V_tV_{t'} = s(\exp(tX))s(\exp(t'X)) = s(\exp((t+t')X)) = V_{t+t'}$. By (52), the 1-parameter groups U_t and V_t have the same generator

$$\tfrac{d}{dt}\Big|_{t=0} \exp(t\,\mathrm{ad}_X)(A) = \mathrm{ad}_X(A) = \tfrac{d}{dt}\Big|_{t=0} s(e^{tX})(A). \tag{53}$$

Since a 1-parameter group of orthogonal transformation of $\mathbb{R}^3$ is in particular a 1-parameter group of unitary transformations of $\mathbb{C}^3$, it follows from Stone's Theorem that $\exp(t\,\mathrm{ad}_X) = s(\exp(tX))$ for all $t \in \mathbb{R}$, so in particular $\exp(\mathrm{ad}_X) = s(\exp(X))$. □

The projective unitary representation of SO(3) on the Hilbert space $\mathbb{C}^2$ is sometimes called the *spin-$\frac{1}{2}$ representation.* The reason is that if one rotates continuously around the z-axis, then a full rotation results in the linear transformation $|\psi\rangle \mapsto -|\psi\rangle$ of the Hilbert space $\mathbb{C}^2$, and it takes *two* full rotations to arrive at the identity transformation $|\psi\rangle \mapsto |\psi\rangle$.

Problem 2.76. Show that although the curve $U(t) = \exp(t\sigma_z)$ in SU(2) satisfies $U(t) = \mathbf{1}$ for $t \in 2\pi\mathbb{Z}$, its image $s(U(t))$ in SO(3) under the spin homomorphism satisfies $s(U(t)) = \mathbf{1}$ for $t \in \pi\mathbb{Z}$. Infer that continuous rotation around the z-axis over 360° takes $|\psi\rangle$ to $-|\psi\rangle$.

Problem 2.77. The projective unitary representation of SU(2) on the n-qubit Hilbert space $\mathcal{H} = \mathbb{C}^2 \otimes \ldots \otimes \mathbb{C}^2$ is given by

$$\pi(U) = U \otimes \ldots \otimes U.$$

Show that this yields a unitary representation of SO(3) if n is even, and a projective unitary representation of SO(3) if n is odd.

3 Open systems

Since quantum information theory describes the flow of information in and out of a quantum system, it is essential to model *open* quantum systems. In order to do so, we have to modify our postulates for observables, states and transformations.

3.1 Postulates for open systems

We write $\mathcal{L}(\mathcal{H})$ for the algebra of bounded operators on the Hilbert space $\mathcal{H}$. In the following we restrict attention to the finite dimensional setting, where $\mathcal{H} \simeq \mathbb{C}^n$ and $\mathcal{L}(\mathcal{H}) \simeq M_n(\mathbb{C})$ is the algebra of $n \times n$ matrices.

.

Definition 3.1 ($*$-subalgebras)**.** A $*$-subalgebra $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H})$ is a linear subspace which contains the unit and is closed under multiplication and adjoints:

1) $\mathbf{1} \in \mathcal{A}$

2) $A, B \in \mathcal{A} \;\Rightarrow\; AB \in \mathcal{A}$

3) $A \in \mathcal{A} \Rightarrow A^\dagger \in \mathcal{A}$.

The observables that correspond to an open system are characterized by a $*$-subalgebra of the algebra $\mathcal{L}(\mathcal{H})$ of all observables.

Postulate 1

An open quantum system is modelled by a $*$-subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{H})$. Its observables are the Hermitian elements of $\mathcal{A}$.

We will take the point of view that a *state* assigns an expectation $\rho(A)$ to every observable A in $\mathcal{A}$. The expectation of the identity operator should be 1, and nonnegative observables should have nonnegative expectation.

Definition 3.2 (States)**.** A state on $\mathcal{A}$ is a linear functional $\rho\colon \mathcal{A} \to \mathbb{C}$ that is

1) Normalized: $\rho(\mathbf{1}) = 1$ and

2) Positive: $\rho(X^\dagger X) \geq 0$ for all $X \in \mathcal{A}$.

The set of all states on $\mathcal{A}$ is denoted by $\mathcal{S}(\mathcal{A})$.

Postulate 2

The set of physical states of an open quantum system $\mathcal{A}$ is modelled by $\mathcal{S}(\mathcal{A})$.

The above description encompasses both classical probability theory and closed quantum systems, and allows them to interact in a natural way.

Example 3.1 (Classical probability space)**.** Let $\mathcal{C}_n \subseteq M_n(\mathbb{C})$ be the commutative $*$-algebra of diagonal $n \times n$ matrices,

$$\mathcal{C}_n = \left\{ \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} ;\, a_i \in \mathbb{C} \right\}.$$

Equivalently, $\mathcal{C}_n$ is the algebra of random variables $A\colon \Omega \to \mathbb{C}$ on the probability space $\Omega = \{1, \ldots, n\}$. Every linear functional $\rho\colon \mathcal{C}_n \to \mathbb{C}$ is of the form $\rho(A) =$

$p_1 a_1 + \ldots + p_n a_n$ for some $p_i \in \mathbb{C}$. If $X = \operatorname{diag}(x_1, \ldots, x_n)$, then $X^\dagger X = \operatorname{diag}(|x_1|^2, \ldots, |x_n|^2)$ has nonnegative coefficients, so $\rho(X^\dagger X) = \sum_{i=1}^n p_i |x_i|^2 \geq 0$ if and only if $p_i \geq 0$ for all i, and $\rho(\mathbf{1}) = 1$ if $\sum_{i=1}^n p_i = 1$. It follows that a *state* on $\mathcal{C}_n$ corresponds precisely to a *probability density* on Ω.

Example 3.2 (Quantum system)**.** Let $\mathcal{A} = M_n(\mathbb{C})$ be the algebra of all linear operators on $\mathcal{H} = \mathbb{C}^n$. Then every unit vector $\psi \in \mathcal{H}$ gives rise to a state $\rho_\psi \colon \mathcal{A} \to \mathbb{C}$ by $\rho_\psi(A) := \langle \psi, A\psi \rangle$. These *vector states* correspond with the notion of physical states that we encountered previously: if ψ and ψ' differ by a phase, they give rise to the same state $\rho_\psi = \rho_{\psi'}$.

Note that the state space $\mathcal{S}(\mathcal{A})$ is *convex*. That is, if ρ_1 and ρ_2 are states on $\mathcal{A}$, then for any $p_0, p_1 \geq 0$ with $p_0 + p_1 = 1$, the convex combination $\rho = p_0\rho_0 + p_1\rho_1$ is again a state. This *mixed state* corresponds to a system which is in state ρ_0 with probability p_0, and in state ρ_1 with probability p_1.

This definition can be rather elegantly recast in terms of convex geometry. Recall that an *extreme point* of a convex set $\mathcal{S}$ is an element $\rho \in \mathcal{S}$ that does not lie on any open line segment in $\mathcal{S}$; if $\rho = p\rho_1 + (1-p)\rho_2$ for $p \in (0,1)$ and $\rho_1, \rho_2 \in \mathcal{S}$, then $\rho_1 = \rho_2$.

Definition 3.3 (Mixed and pure states)**.** A state is called *pure* if it is an extreme point of the convex set $\mathcal{S}(\mathcal{A})$, and *mixed* otherwise.

Note that even for $\mathcal{A} = M_n(\mathbb{C})$, our new definition of states is more general than the old one. If ψ_0 and ψ_1 are orthogonal unit vectors in $\mathcal{H}$, then the mixed state $\rho = p_0\rho_{\psi_0} + p_1\rho_{\psi_1}$ is in general *not* a vector state.

Problem 3.1. Let $\psi = \alpha\psi_0 + \beta\psi_1$ for complex numbers α, β with $|\alpha|^2 + |\beta|^2 = 1$ and $|\alpha|^2 = p_0 \neq 0$ and $|\beta|^2 = p_1 \neq 0$. Then ρ_ψ is *not* the same as $p_0\rho_{\psi_0} + p_1\rho_{\psi_1}$.

Problem 3.2 (Pure states on commutative $*$-algebras)**.** The pure states on the algebra $\mathcal{C}_n \subseteq M_n(\mathbb{C})$ of diagonal matrices are given by probability distributions that are concentrated in a single point, $(p_1, \ldots, p_n) = (0, \ldots, 1, \ldots, 0)$.

So the state space of the commutative $*$-algebra $\mathcal{C}_n$ is the standard $(n-1)$-simplex

$$\Delta_{n-1} = \{(p_1, \ldots p_n) \in \mathbb{R}^n \,;\, p_i \geq 0, \textstyle\sum_{i=1}^n p_i = 1\},$$

and the pure states are the n vertices of this polygon.

Problem 3.3. Sketch $\Delta_2 \subseteq \mathbb{R}^3$.

3.1.1 The cone $\mathcal{A}_+$ of positive semidefinite elements of $\mathcal{A}$

As in the case of closed systems, there is a third postulate that governs *transformations* between open quantum systems. This requires a little preparation, so we interrupt our discussion of the basic postulates for a moment to study the cone [3] $\mathcal{A}_+ \subseteq \mathcal{A}$ of positive semidefinite elements of a $*$-algebra $\mathcal{A}$.

[3] If V is a (real or complex) vector space, then $C \subseteq V$ is called a *cone* if it is stable under multiplication by positive real numbers, $\lambda C \subseteq C$ for all $\lambda \in \mathbb{R}^{>0}$.

Definition 3.4. An element $A \in \mathcal{A}$ is *positive semidefinite*, denoted $A \geq 0$, if $\langle \psi, A\psi \rangle \geq 0$ for all $\psi \in \mathcal{H}$. The convex cone of positive semidefinite elements of $\mathcal{A}$ is denoted

$$\mathcal{A}_+ := \{A \in \mathcal{A}\, ;\, A \geq 0\}.$$

Recall that a matrix $A \in M_n(\mathbb{C})$ is positive semidefinite if and only if it is of the form $A = X^\dagger X$ for some $X \in M_n(\mathbb{C})$. We now prove that if $A \in \mathcal{A}$, then we can also choose X to be in $\mathcal{A}$. This requires a few preparations.

Lemma 3.1. *A $*$-subalgebra $\mathcal{A} \subseteq M_n(\mathbb{C})$ contains all spectral projections of its Hermitian elements.*

Proof. Let $a_1, \ldots, a_r$ be the eigenvalues of a Hermitian element $A \in \mathcal{A}$, and let P_{a_i} be the orthogonal projection onto the eigenspace V_{a_i}. Then

$$P_{a_i} = \frac{(A - a_1\mathbf{1})}{a_i - a_1} \cdots \frac{(A - a_{i-1}\mathbf{1})}{a_i - a_{i-1}} \cdot \frac{(A - a_{i+1}\mathbf{1})}{a_i - a_{i+1}} \cdots \frac{A - a_r\mathbf{1}}{a_i - a_r}.$$

Indeed, the r.h.s. acts as zero on the eigenspaces V_{a_j} with $j \neq i$, and as the identity on V_{a_i}. If A is in $\mathcal{A}$, then each of the factors on the r.h.s. is in $\mathcal{A}$ as well, so $P_{a_i} \in \mathcal{A}$. □

Problem 3.4. Let $A \in \mathcal{A}$, and let V_a be an eigenspace of A. If $\dim(V_a) > 1$, then it is *not* necessarily true that the projection $|\psi_a\rangle \langle\psi_a|$ onto an eigenvector $\psi_a \in V_a$ is again an element of $\mathcal{A}$. Give a counterexample.

Lemma 3.2. *An element $A \in \mathcal{A}$ is Hermitian if and only if $\langle \psi, A\psi \rangle \in \mathbb{R}$ for all $\psi \in \mathcal{H}$.*

Proof. The imaginary part of $\langle \psi, A\psi \rangle$ is zero if and only if

$$\begin{aligned} 2i\mathrm{Im}\langle \psi, A\psi \rangle &= \langle \psi, A\psi \rangle - \overline{\langle \psi, A\psi \rangle} \\ &= \langle \psi, A\psi \rangle - \langle A\psi, \psi \rangle \\ &= \langle \psi, (A - A^\dagger)\psi \rangle \end{aligned}$$

is zero for all $\psi \in \mathcal{H}$. This is the case if and only if $A = A^\dagger$. □

Proposition 3.3. *For $A \in \mathcal{A}$, the following are equivalent:*

1) $A \geq 0$.

2) $A = X^\dagger X$ for some $X \in \mathcal{A}$.

3) A is Hermitian with $a \geq 0$ for all $a \in \mathrm{spec}(A)$.

Proof. The main point here is that if $A \in \mathcal{A}$, then X is in $\mathcal{A}$ again.

2) ⇒ 1) This follows from $\langle \psi, X^\dagger X\psi \rangle = \|X\psi\|^2 \geq 0$.

1) ⇒ 3) The operator A is Hermitian by Lemma 3.2, and $a = \langle \psi_a, A\psi_a \rangle \geq 0$ for every eigenvector ψ_a of unit length.

3) ⇒ 2) In the spectral decomposition $A = \sum_{a \in \mathrm{spec}(A)} aP_a$, we have $a \geq 0$ by assumption and $P_a \in \mathcal{A}$ by Lemma 3.1. So $X := \sum_{a \in \mathrm{spec}(A)} \sqrt{a}P_a$ is an element of $\mathcal{A}$, and $A = X^\dagger X$. □

The cone $\mathcal{A}_+$ of positive semidefinite elements can therefore be expressed as

$$\mathcal{A}_+ = \{X^\dagger X \,;\, X \in \mathcal{A}\}.$$

Similarly, the requirement in Definition 3.2 that $\rho(X^\dagger X) \geq 0$ for all $X \in \mathcal{A}$ can be reformulated as $\rho(A) \geq 0$ for all $A \in \mathcal{A}_+$. In other words, the expectation of every positive semidefinite observable is nonnegative.

The following result tells us that every $*$-algebra $\mathcal{A}$ has a 'large' supply of positive semidefinite elements. More precisely, $\mathcal{A}$ is spanned (as a complex vector space) by its cone $\mathcal{A}_+$ of positive definite elements.

Proposition 3.4. *Every $A \in \mathcal{A}$ can be written as a complex linear combination $A = A_1 - A_2 + iA_3 - iA_4$ with $A_j \in \mathcal{A}$ and $A_j \geq 0$ for $j = 1, 2, 3, 4$.*

Proof. Every $A \in \mathcal{A}$ can be written as $A = X + iY$ for the Hermitian operators $X := \frac{1}{2}(A + A^\dagger)$ and $Y := \frac{1}{2i}(A - A^\dagger)$. The Hermitian element $X \in \mathcal{A}$ has spectral decomposition $X = \sum_{x \in \mathrm{spec}(X)} xP_x$. We therefore have $X = A_1 - A_2$ for the two nonnegative operators

$$A_1 := \sum_{x \geq 0,\, x \in \mathrm{spec}(X)} xP_x, \quad A_2 := \sum_{x < 0,\, x \in \mathrm{spec}(X)} (-x)P_x,$$

both of which are in $\mathcal{A}$ by Lemma 3.1. Similarly we have $Y = A_3 - A_4$, so that $A = X + iY$ can be written as $X + iY = A_1 - A_2 + iA_3 - iA_4$. □

3.1.2 Trace pairing and self-duality of the cone $\mathcal{A}_+$

Recall that if a finite dimensional Hilbert space $\mathcal{H}$ admits an orthonormal basis $e_1, \ldots, e_n$, then the *trace* of $A \in \mathcal{L}(\mathcal{H})$ is defined as

$$\mathbf{tr}(A) := \textstyle\sum_{i=1}^n \langle e_i, Ae_i \rangle. \tag{54}$$

It is the sum of the diagonal elements of the matrix $A_{ij} = \langle e_i, Ae_j \rangle$.

Remark 3.1. The *normalized trace* $\frac{1}{n}\mathbf{tr} \colon \mathcal{A} \to \mathbb{C}$ is a convex combination of vector states, and hence a state on $\mathcal{A}$.

Note that for all $A, B \in \mathcal{L}(\mathcal{H})$, we have

$$\mathbf{tr}(AB) = \mathbf{tr}(BA). \tag{55}$$

Indeed, since $(AB)_{ij} = \sum_{k=1}^n A_{ik}B_{kj}$, we have $\mathbf{tr}(AB) = \sum_{i=1}^n \sum_{k=1}^n A_{ik}B_{ki}$. Interchanging A and B, we find $\mathbf{tr}(BA) = \sum_{i=1}^n \sum_{k=1}^n B_{ik}A_{ki}$, which is the same expression. We will often use the cyclic property of the trace

$$\mathbf{tr}(A_1A_2 \cdots A_n) = \mathbf{tr}(A_2 \cdots A_nA_1), \tag{56}$$

which follows easily from (55). For example, the cyclic property of the trace allows one to see right away that definition 54 is independent of the choice of orthonormal basis; if a basis $\psi_1, \ldots, \psi_n$ is related to $e_1, \ldots, e_n$ by $\psi_i = Ue_i$, then

$$\textstyle\sum_{i=1}^n \langle \psi_i, A\psi_i \rangle = \sum_{i=1}^n \langle e_i, U^\dagger A U e_i \rangle = \mathbf{tr}(U^\dagger A U).$$

Using (55) we find $\mathbf{tr}(U^\dagger AU) = \mathbf{tr}(AUU^\dagger) = \mathbf{tr}(A)$, and we conclude that

$$\mathbf{tr}(A) = \textstyle\sum_{i=1}^n \langle \psi_i, A\psi_i \rangle$$

for *any* orthonormal basis ψ_i of $\mathcal{H}$.

The trace induces a sesquilinear form $\langle \cdot, \cdot \rangle_{\mathbf{tr}} \colon \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ on $\mathcal{A}$ called the *trace pairing.* It is defined on $A, B \in \mathcal{A}$ by

$$\langle A, B \rangle_{\mathbf{tr}} := \mathbf{tr}(A^\dagger B). \tag{57}$$

Lemma 3.5. *The trace pairing is an inner product on $\mathcal{A}$.*

Proof. The trace pairing is clearly linear in B. To show that it is sesquilinear, note that $\mathbf{tr}(X^\dagger) = \overline{\mathbf{tr}(X)}$ for all $X \in \mathcal{A}$. We thus have $\langle B, A \rangle_{\mathbf{tr}} = \mathbf{tr}(B^\dagger A) = \mathbf{tr}\big((AB^\dagger)^\dagger\big) = \overline{\mathbf{tr}(AB^\dagger)} = \overline{\langle A, B \rangle}_{\mathbf{tr}}$. To see that the inner product is positive definite, note that

$$\langle A, A \rangle_{\mathbf{tr}} = \mathbf{tr}(A^\dagger A) = \sum_{a \in \mathrm{spec}(A^\dagger A)} a \dim(V_a)$$

is the sum of the (nonnegative!) eigenvalues of $A^\dagger A$ counted with multiplicity. So $\langle A, A \rangle_{\mathbf{tr}} \geq 0$, and $\langle A, A \rangle_{\mathbf{tr}} = 0$ if and only if all eigenvalues of $A^\dagger A$ are zero. But then $A^\dagger A = 0$, so $A = 0$ since $\langle \psi, A^\dagger A\psi \rangle = \|A\psi\|^2 = 0$ for all $\psi \in \mathcal{H}$. □

Problem 3.5. In fact, the trace pairing is the usual inner product under the identification $\mathcal{L}(\mathcal{H}) \simeq M_n(\mathbb{C}) \simeq \mathbb{C}^{n^2}$.

Problem 3.6. Show that $\mathbf{tr}(A \,|\psi\rangle\langle\psi|) = \langle \psi, A\psi \rangle$ for any unit vector ψ.

The cone $\mathcal{A}_+ \subseteq \mathcal{A}$ of positive semidefinite elements is *self-dual*[4] with respect to the trace pairing.

Theorem 3.6 (Self-duality of $\mathcal{A}_+$)**.** *For $A \in \mathcal{A}$, we have $A \in \mathcal{A}_+$ if and only if $\langle A, B \rangle_{\mathbf{tr}} \geq 0$ for all $B \in \mathcal{A}_+$.*

Proof. If $A, B \in \mathcal{A}^+$, then $A = X^\dagger X$ and $B = Y^\dagger Y$ for some $X, Y \in \mathcal{A}$. So

$$\langle A, B \rangle_{\mathbf{tr}} = \mathbf{tr}(X^\dagger X Y^\dagger Y) = \mathbf{tr}(XY^\dagger Y X^\dagger) = \langle YX^\dagger, YX^\dagger \rangle_{\mathbf{tr}} \geq 0.$$

So if $A \in \mathcal{A}^+$, then $\langle A, B \rangle_{\mathbf{tr}} \geq 0$ for all $B \in \mathcal{A}^+$.

Conversely, suppose that $A \in \mathcal{A}$ is such that $\langle A, B \rangle_{\mathbf{tr}} \geq 0$ for all $B \in \mathcal{A}_+$. Then A is Hermitian. Indeed, since the imaginary part of $\langle A, B \rangle_{\mathbf{tr}}$ is zero, we have $0 = \langle A, B \rangle_{\mathbf{tr}} - \langle B, A \rangle_{\mathbf{tr}}$. It follows that

$$0 = \mathbf{tr}(A^\dagger B) - \mathbf{tr}(B^\dagger A) = \mathbf{tr}\big((A^\dagger - A)B\big) = \langle A - A^\dagger, B \rangle_{\mathbf{tr}}$$

[4] If $\mathcal{H}$ is a Hilbert space, then the *dual* of $C \subseteq \mathcal{H}$ is the cone $C^* := \{\psi \in \mathcal{H}\,;\, \langle \psi, C \rangle \subseteq \mathbb{R}^{\geq 0}\}$, and C is called *self-dual* if $C = C^*$.

for all $B \in \mathcal{A}^+$. Since $A - A^\dagger$ is perpendicular to $\mathcal{A}^+$, and since $\mathcal{A}$ is spanned (as a complex vector space) by $\mathcal{A}^+$ (Proposition 3.4), we conclude that $A - A^\dagger$ is perpendicular to $\mathcal{A}$. So $A - A^\dagger = 0$, and A is Hermitian. If we decompose A into spectral projections as $A = \sum_{a \in \mathrm{spec}(A)} aP_a$, then from $P_a \geq 0$ we find that $\mathbf{tr}(AP_a) = a \dim(V_a) \geq 0$. So all eigenvalues $a \in \mathrm{spec}(A)$ are nonnegative, and $A \geq 0$. □

3.1.3 Density matrices

A state on $\mathcal{A}$ can be conveniently described in terms of a *density matrix*, which is an element $R \in \mathcal{A}$ with $R \geq 0$ and $\mathbf{tr}(R) = 1$. The following result characterizes states on $\mathcal{A}$ in terms of density matrices.

Theorem 3.7 (density matrices)**.** *Every state $\rho \in \mathcal{S}(\mathcal{A})$ can be written as*

$$\rho(A) = \mathbf{tr}(RA)$$

for a unique density matrix $R \in \mathcal{A}$ with $R \geq 0$ and $\mathbf{tr}(R) = 1$.

To show this, we will need the *Riesz Representation Theorem*. It holds for arbitrary Hilbert spaces, but here we only prove the finite dimensional version (which is much easier). We denote by $\mathcal{H}^*$ the linear dual of $\mathcal{H}$,

$$\mathcal{H}^* := \{\phi \colon \mathcal{H} \to \mathbb{C}\,;\, \phi \text{ is linear}\}.$$

Lemma 3.8 (Riesz)**.** *Let $\mathcal{H}$ be a finite dimensional Hilbert space. Then every $\phi \in \mathcal{H}^*$ can be represented as $\phi(\psi) = \langle \chi, \psi \rangle$ for a unique $\chi \in \mathcal{H}$.*

Proof. The map $\mathcal{H} \to \mathcal{H}^*$ defined by $\psi \mapsto \langle \psi, \cdot \, \rangle$ is $\mathbb{R}$-linear, and injective because the inner product is nondegenerate. Since $\mathcal{H}$ and $\mathcal{H}^*$ have the same dimension over $\mathbb{R}$, it is also surjective. □

Proof of Theorem 3.7. By Lemma 3.8 applied to $\mathcal{A}$ with the trace pairing, we have $\rho(A) = \langle R, A \rangle_{\mathbf{tr}}$ for a unique element $R \in \mathcal{A}$. Since $\rho(A) \geq 0$ for all $A \in \mathcal{A}$ with $A \geq 0$, we have $\langle R, A \rangle_{\mathbf{tr}} \geq 0$ for all $A \in \mathcal{A}_+$, and it follows from Theorem 3.6 that $R \in \mathcal{A}_+$. Since $\rho(\mathbf{1}) = 1$, we have $\langle R, \mathbf{1} \rangle_{\mathbf{tr}} = \mathbf{tr}(R^\dagger) = 1$, which yields $\mathbf{tr}(R) = 1$ because R is Hermitian. □

Remark 3.2. In particular, the eigenvalues p of R are nonnegative and sum to one if counted with multiplicity. Every state ρ on $\mathcal{A}$ therefore induces a probability distribution on $\mathrm{spec}(R)$.

Example 3.3. On the algebra $\mathcal{C}_n$ of diagonal $n \times n$ matrices, the state $\rho(A) = \sum_{i=1}^n p_i a_i$ is represented by the density matrix $R = \mathrm{diag}(p_1, \ldots, p_n)$.

3.1.4 The Bloch sphere: pure and mixed states of a qubit

For $\mathcal{A} = M_n(\mathbb{C})$, the pure states are the vector states $\rho_\psi(A) = \langle \psi, A\psi \rangle$, and the corresponding density matrices are precisely the Hermitian matrices R with $\mathrm{spec}(R) = \{0, 1\}$.

Problem 3.7. In order to prove this, let R be a density matrix in $M_n(\mathbb{C})$.

a) If $\mathrm{spec}(R) \neq \{0, 1\}$, then R is a convex combination of two distinct density matrices.

b) If $\mathrm{spec}(R) = \{0, 1\}$, then $R = |\psi\rangle \langle\psi|$ for a unit vector $\psi \in \mathbb{C}^n$.

c) If $|\psi\rangle \langle\psi| = p_1 R_1 + p_2 R_2$ is a convex combination of two density matrices R_1 and R_2, then $p_1 R_1$ and $p_2 R_2$ are zero on $\psi^\perp$, and map $\mathbb{C}\psi$ to $\mathbb{C}\psi$.

d) Conclude that the pure states on $\mathcal{A} = M_n(\mathbb{C})$ are precisely the vector states $\rho(A) = \langle \psi, A\psi \rangle$.

So the state space $\mathcal{S}(\mathcal{A})$ of $\mathcal{A} = M_n(\mathbb{C})$ is the set of all density matrices, and its extremal points are parameterized by the complex projective space $\mathbb{C}\mathrm{P}^{n-1}$.

For the qubit $\mathcal{A} = M_2(\mathbb{C})$, this can be made even more explicit. Recall that every Hermitian 2×2 matrix R can be written as

$$R = \frac{1}{2}\begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix}$$

for some $t, x, y, z \in \mathbb{R}$. If R is a density matrix, then $\mathbf{tr}(R) = 1$ implies that $t = 1$. Since the eigenvalues of R are two numbers $p_0, p_1 \in [0, 1]$ that sum to 1, we have $\det(R) = p_0 p_1 \in [0, 1/4]$ with $\det(R) = 0$ if and only if R represents a pure state. Since $\det(R) = \frac{1}{4}(1 - x^2 - y^2 - z^2)$, the density matrices correspond to vectors $\vec{x} = (x, y, z)$ with norm at most one, and $\|\vec{x}\| = 1$ if and only if the state is pure.

Definition 3.5. The *Bloch ball* is the set of density matrices

$$R = \tfrac{1}{2}(\mathbf{1} - x\sigma_x - y\sigma_y - z\sigma_z)$$

on $M_2(\mathbb{C})$, parameterized by *Bloch vectors* $\vec{x} = (x, y, z)$ in the closed unit ball $\{\vec{x} \in \mathbb{R}^3\,;\, \|\vec{x}\| \leq 1\}$. The pure states correspond to Bloch vectors of norm one.

So the extremal points of the state space of $M_2(\mathbb{C})$ are parameterized by the unit sphere in $\mathbb{R}^3$. By contrast, recall from Problem 3.2 that the extremal points of the state space of the commutative algebra $\mathcal{C}_n$ constitute a finite set with n elements.

3.1.5 Combined systems

If system A is modelled by the $*$-algebra $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H}_A)$, and system B is modelled by $\mathcal{B} \subseteq \mathcal{L}(\mathcal{H}_B)$, then the combined system is modelled by the $*$-algebra

$$\mathcal{A} \otimes \mathcal{B} \subseteq \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B).$$

Example 3.4. With respect to the basis $e_i \otimes f_k$ of $\mathbb{C}^n \otimes \mathbb{C}^m$ obtained from the canonical bases $\{e_i\,;\, i = 1, \ldots, n\}$ and $\{f_k\,;\, k = 1, \ldots, m\}$ of $\mathbb{C}^n$ and $\mathbb{C}^m$, respectively, the tensor product $A \otimes B$ of $A \in M_n(\mathbb{C})$ and $B \in M_m(\mathbb{C})$ is

$$\begin{pmatrix} a_{11}B & a_{12}B & \ldots \\ a_{21}B & a_{22}B & \ldots \\ \vdots & & \ddots \end{pmatrix}$$

if we order the basis elements lexicographically in blocks of size m as $e_1 \otimes f_1$, $e_1 \otimes f_2$, …, $e_1 \otimes f_m$; $e_2 \otimes f_1$, … $e_n \otimes f_m$. (If we order anti-lexicographically in blocks of size n, then the roles of A and B are interchanged.)

Example 3.5. If we combine a quantum system $M_m(\mathbb{C})$ with a classical probability space $\mathcal{C}_n$, then elements of $\mathcal{C}_n \otimes M_m(\mathbb{C})$ can be described as $nm \times nm$ matrices with nonzero entries only on the diagonal blocks of size m,

$$A = \begin{pmatrix} A_1 & 0 & \ldots \\ 0 & A_2 & \ldots \\ \vdots & & \ddots \end{pmatrix}$$

for $A_i \in M_m(\mathbb{C})$. Alternatively, we can view them as matrix-valued functions $A \colon \Omega \to M_m(\mathbb{C})$ on the probability space $\Omega = \{1, \ldots, m\}$.

3.2 Transformations

The *stochastic equivalence principle* states that a system $\mathcal{A}$ that is in state ρ_0 with probability p_0 and in state ρ_1 with probability p_1 is physically indistinguishable from a system in state $\rho = p_0\rho_0 + p_1\rho_1$.

Problem 3.8. There are two containers in the room. Alice enters, carrying two bags with an equal amount of qubits in state $|0\rangle$ (first bag) and $|1\rangle$ (second bag). She empties her bags into one of the two containers. Then Bob enters with two bags of qubits in state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ (first bag) and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ (second bag). He empties his bags into the other container. Alice and Bob leave the room. You enter, and find two containers filled with qubits. Is it possible to determine which container belongs to Alice and which one belongs to Bob? Suppose instead that Alice and Bob would have simply put their bags of qubits on the floor. Would it then be possible to determine which of the four bags belong to Alice and which belong to Bob? If so, how would you do this?

Starting from the stochastic equivalence principle, we now deduce that transformations of an open quantum system are modelled by *completely positive maps.*

3.2.1 Completely positive maps

Since the state of a system captures everything there is to know about this system, a transformation from $\mathcal{A}$ to $\mathcal{B}$ is modelled by a map $\tau \colon \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{B})$

from the state space of $\mathcal{A}$ to the state space of $\mathcal{B}$. The stochastic equivalence principle implies that this must be an *affine* map:

$$\tau(p_0\rho_0 + p_1\rho_1) = p_0\tau(\rho_0) + p_1\tau(\rho_1) \tag{58}$$

for all $\rho_0, \rho_1 \in \mathcal{S}(\mathcal{A})$ and $p_0, p_1 \in [0, 1]$ with $p_0 + p_1 = 1$.

Since every state on $\mathcal{A}$ is in particular a linear map $\mathcal{A} \to \mathbb{C}$, the state space $\mathcal{S}(\mathcal{A})$ is a convex subset of the dual space $\mathcal{A}^* = \{\phi\colon \mathcal{A} \to \mathbb{C}\,;\, \phi \text{ is linear}\}$.

Proposition 3.9. *Every $\phi \in \mathcal{A}^*$ can be decomposed into 4 states $\rho_j \in \mathcal{S}(\mathcal{A})$ as*

$$\phi = \lambda_1\rho_1 - \lambda_2\rho_2 - i\lambda_3\rho_3 + i\lambda_4\rho_4 \tag{59}$$

for 4 nonnegative real numbers $\lambda_j \geq 0$. The decomposition is unique if we omit the terms with $\lambda_j = 0$.

Proof. By trace duality, $\phi(A) = \langle F, A\rangle_{\mathbf{tr}} = \mathbf{tr}(F^\dagger A)$ for a unique element $F \in \mathcal{A}$. Using Proposition 3.4, we can decompose $F = F_1 - F_2 + iF_3 - iF_4$ into $F_j \in \mathcal{A}$ with $F_j \geq 0$. Set $\lambda_j := \mathbf{tr}F_j$. If $\lambda_j = 0$ then $F_j = 0$ and we omit the term. If $\lambda_j > 0$ then we we normalize F_j to a density matrix $R_j := \frac{1}{\lambda_j}F_j$, yielding (59) with $\rho_j(A) = \mathbf{tr}(R_jA)$. The uniqueness follows from the nondegeneracy of the inner product $\langle F, A\rangle_{\mathbf{tr}}$ and the uniqueness of the decomposition of F. □

Corollary 3.10. *Every* affine *transformation $\tau\colon \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{B})$ extends to a* linear *transformation $T^*\colon \mathcal{A}^* \to \mathcal{B}^*$.*

Proof. Set $T^*(\phi) := \lambda_1\tau(\rho_1) - \lambda_2\tau(\rho_2) - i\lambda_3\tau(\rho_3) + i\lambda_4\tau(\rho_4)$. This is well defined, linear in ϕ, and $T(\rho) = \tau(\rho)$ on states. □

If V and W are complex vector spaces, then a linear map $L\colon W \to V$ gives rise to a dual linear map $L^*\colon V^* \to W^*$, defined by $L^*\phi = \phi \circ L$. Moreover, if V and W are finite dimensional, then every linear map $V^* \to W^*$ is of this form. Since $\tau\colon \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{B})$ extends to a linear map $T^*\colon \mathcal{A}^* \to \mathcal{B}^*$, it can be expressed as the dual of a linear map $T\colon \mathcal{B} \to \mathcal{A}$ in the *other* direction,

$$(T^*\phi)(B) = \phi(T(B)) \text{ for all } \phi \in \mathcal{A}^* \text{ and } B \in \mathcal{B}.$$

Definition 3.6. A linear map $T\colon \mathcal{B} \to \mathcal{A}$ is called *positive* if $T(B) \geq 0$ for all $B \geq 0$, and *normalized* if $T(\mathbf{1}_{\mathcal{B}}) = \mathbf{1}_{\mathcal{A}}$

Proposition 3.11. *Let $T\colon \mathcal{B} \to \mathcal{A}$ be a linear map. Then $T^*\colon \mathcal{A}^* \to \mathcal{B}^*$ maps $\mathcal{S}(\mathcal{A}) \subseteq \mathcal{A}^*$ to $\mathcal{S}(\mathcal{B}) \subseteq \mathcal{B}^*$ if and only if T is positive and normalized.*

Problem 3.9. Prove this.

Proposition 3.12. *If $T\colon \mathcal{B} \to \mathcal{A}$ is positive, then $T(B^\dagger) = T(B)^\dagger$ for all $B \in \mathcal{B}$.*

Problem 3.10. Prove this, for example using Proposition 3.4.

Interestingly, the requirement that T is positive and normalized is not quite sufficient in order to interpret it as a physical transformation. The reason for this is the following funny phenomenon. Suppose $T\colon \mathcal{B} \to \mathcal{A}$ is a positive, normalized map that describes a physical transformation from $\mathcal{A}$ to $\mathcal{B}$. If we extend $\mathcal{A}$ by a quantum system $M_n(\mathbb{C})$, then the trivial extension $T \otimes \mathrm{Id}_n \colon \mathcal{B} \otimes M_n(\mathbb{C}) \to \mathcal{A} \otimes M_n(\mathbb{C})$ is again a physical transformation: it describes what happens if we perform the transformation T from $\mathcal{A}$ to $\mathcal{B}$ while doing nothing at all to the quantum system $M_n(\mathbb{C})$. But although both $T\colon \mathcal{B} \to \mathcal{A}$ and $\mathrm{Id}_n \colon M_n(\mathbb{C}) \to M_n(\mathbb{C})$ are positive and normalized, their tensor product $T \otimes \mathrm{Id}_n$ is *not* automatically positive. For this reason, physical transformations from $\mathcal{A}$ to $\mathcal{B}$ are modelled by *completely positive maps.*

Definition 3.7 (CP maps)**.** A linear map $T\colon \mathcal{B} \to \mathcal{A}$ is completely positive if $T \otimes \mathrm{Id}_n \colon \mathcal{B} \otimes M_n(\mathbb{C}) \to \mathcal{A} \otimes M_n(\mathbb{C})$ is positive for all $n \in \mathbb{N}$.

Postulate 3

Operations from a system $\mathcal{A}$ to a system $\mathcal{B}$ are modelled by normalized, completely positive maps $T\colon \mathcal{B} \to \mathcal{A}$.

Problem 3.11. Show that compositions of completely positive maps are completely positive. Use this to show that tensor products of completely positive maps are completely positive as well.

One can show that if either $\mathcal{A}$ or $\mathcal{B}$ is commutative, then positivity implies complete positivity. The following example shows that this does not hold in general.

Example 3.6 (A positive map which is not CP)**.** The transposition

$$T\colon M_2(\mathbb{C}) \to M_2(\mathbb{C}), \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

is clearly normalized. It is positive because it preserves the trace and the determinant, and hence the spectrum. On the other hand, the trivial extension $T \otimes \mathrm{Id}_2 \colon M_2(\mathbb{C}) \otimes M_2(\mathbb{C}) \to M_2(\mathbb{C}) \otimes M_2(\mathbb{C})$ to a two-qubit system satisfies

$$T \otimes \mathrm{Id}_2 \colon \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The former is a positive matrix (it is twice the projection onto the Bell state $|00\rangle + |11\rangle$), whereas the latter has spectrum $\{1, -1\}$.

3.2.2 Schrödinger and Heisenberg picture

This gives us two ways to describe a physical transformation. In the *Schrödinger picture* we have a linear map $T^* \colon \mathcal{A}^* \to \mathcal{B}^*$ that maps *states* on $\mathcal{A}$ to *states* on $\mathcal{B}$. In the *Heisenberg picture* we have a completely positive map $T \colon \mathcal{B} \to \mathcal{A}$ that maps *observables* in $\mathcal{B}$ to *observables* in $\mathcal{A}$.

Schrödinger picture	Heisenberg picture
$T^* \colon \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{B})$	$T \colon \mathcal{B} \to \mathcal{A}$

By trace duality, we get a *third* description in terms of density matrices. Denote by T_* the *adjoint* of the linear map $T \colon \mathcal{B} \to \mathcal{A}$, so that

$$\langle A, T(B)\rangle_{\mathbf{tr}_{\mathcal{A}}} = \langle T_*(A), B\rangle_{\mathbf{tr}_{\mathcal{B}}} \tag{60}$$

for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. If $\rho(A) = \mathbf{tr}(RA)$ for a density matrix $R \in \mathcal{A}$, then $T^*\rho(B) = \mathbf{tr}(T_*(R)B)$ for a density matrix $T_*(R) \in \mathcal{B}$. Indeed, since

$$(T^*\rho)(B) = \rho(T(B)) = \mathbf{tr}(RT(B)) = \langle R, T(B)\rangle_{\mathbf{tr}_{\mathcal{A}}},$$

equation (60) yields

$$(T^*\rho)(B) = \langle T_*(R), B\rangle_{\mathbf{tr}_{\mathcal{B}}} = \mathbf{tr}(T_*(R)B).$$

In other words: if $\rho \in \mathcal{S}(\mathcal{A})$ is described by the density matrix $R \in \mathcal{A}$, then $T^*\rho \in \mathcal{B}$ is described by the density matrix $T_*(R) \in \mathcal{B}$.

Problem 3.12. A linear map $T_* \colon \mathcal{A} \to \mathcal{B}$ maps density matrices to density matrices if and only if it is *positive*, $T_*(A) \geq 0$ for $A \geq 0$ and *trace preserving*, $\mathbf{tr}(T_*(A)) = \mathbf{tr}(A)$ for all $A \in \mathcal{A}$.

Schrödinger picture	Heisenberg picture
$T^* \colon \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{B})$ $T_* \colon \mathcal{A} \to \mathcal{B}$	$T \colon \mathcal{B} \to \mathcal{A}$

Note that although both $T \colon \mathcal{B} \to \mathcal{A}$ and $T_* \colon \mathcal{A} \to \mathcal{B}$ are completely positive, the map T preserves the identity whereas T_* preserves the trace.

Problem 3.13. The properties of $T \colon \mathcal{B} \to \mathcal{A}$ and $T_* \colon \mathcal{A} \to \mathcal{B}$ are related as follows.

a) Show that T_* is trace preserving if and only if T is normalized.

b) Using Theorem 3.6 or otherwise, show that T_* is positive if and only if T is positive.

c) Show that $(T \otimes \mathrm{Id}_n)_* = T_* \otimes \mathrm{Id}_n$, and conclude that T_* is CP if and only if T is CP.

3.3 Information transfer

The point of studying open systems is that they can interact with their environment. First one couples the system $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H}_A)$ of interest to a second system $\mathcal{B} \subseteq \mathcal{L}(\mathcal{H}_B)$. Then there is a unitary time evolution on the *closed* system $\mathcal{A} \otimes \mathcal{B} \subseteq \mathcal{L}(\mathcal{H}_A \otimes \mathcal{B})$. Finally, one restricts attention to the subsystem $\mathcal{A} \otimes \mathbf{1}_B \subseteq \mathcal{A} \otimes \mathcal{B}$. We describe these 3 steps separately, and then investigate the implications of information transfer from $\mathcal{A}$ to $\mathcal{B}$.

Coupling to a second system In the Schrödinger picture, coupling a system $\mathcal{A}$ to a second system $\mathcal{B}$ in state Φ is modelled by

$$T^* \colon \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{A} \otimes \mathcal{B}) \colon \quad T^*\rho = \rho \otimes \Phi.$$

On density matrices this yields $T_* \colon \mathcal{A} \to \mathcal{A} \otimes \mathcal{B}$ with $T_* R_A = R_A \otimes R_\Phi$, where R_Φ is the density matrix of Φ. In the Heisenberg picture, this yields $T \colon \mathcal{A} \otimes \mathcal{B} \to \mathcal{A}$ with $T(A \otimes B) = \Phi(B)A$, so $T = \mathrm{Id}_{\mathcal{A}} \otimes \Phi$.

Problem 3.14. Show that every vector state $\rho_\phi(B) = \langle \phi, B\phi \rangle$ is a completely positive map $\rho_\phi \colon \mathcal{B} \to \mathbb{C}$. Show that convex combinations of completely positive maps are completely positive, and conclude that every state $\Phi \colon \mathcal{B} \to \mathbb{C}$ is completely positive. Conclude that $T = \mathrm{Id}_{\mathcal{A}} \otimes \Phi$ is completely positive.

Unitary transformations Let $\mathcal{A} = \mathcal{L}(\mathcal{H})$, and let $U \colon \mathcal{H} \to \mathcal{H}$ be a unitary transformation. On *observables*, the corresponding CP map $T \colon \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ is

$$T(B) := U^\dagger B U. \tag{61}$$

On *states* it is given by

$$T^*\rho(B) = \rho(U^\dagger B U), \tag{62}$$

and on *density matrices* it is given by

$$T_*(R_A) = U R_A U^\dagger. \tag{63}$$

Note the difference between (61) and (63)! To see that (63) is indeed correct, we use the cyclic property of the trace to see that

$$T^*\rho(B) = \mathbf{tr}\big(R(U^\dagger B U)\big) = \mathbf{tr}\big((URU^\dagger)B\big).$$

Since this is equal to $\mathbf{tr}(T_*(R)B)$, the density matrix of $T^*\rho$ is indeed $T_*(R) = URU^\dagger$.

By way of sanity check, let us verify that this agrees with the transformation $|\psi\rangle \mapsto U|\psi\rangle$ on vector states that we used for closed systems. For $\rho(B) = \langle \psi, B\psi \rangle$, we find $T^*\rho(B) = \langle \psi, U^\dagger B U \psi \rangle$, which is indeed the expectation $\langle (U\psi), B(U\psi) \rangle$ of B with respect to $U\psi$. And for the corresponding density matrix $R = |\psi\rangle\langle\psi|$, we find $T_*(R) = U|\psi\rangle\langle\psi|U^\dagger = |U\psi\rangle\langle U\psi|$ as expected.

Problem 3.15. $T(B) = U^\dagger B U$ is normalized and completely positive.

Restricting attention to a subsystem If $\mathcal{B} \subseteq \mathcal{A}$ is a subsystem, then the inclusion $T\colon \mathcal{B} \to \mathcal{A}$ gives rise to the restriction $T^*\colon \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{B})$ with $T^*\rho = \rho|_{\mathcal{B}}$. This allows us to *restrict attention* from a system $\mathcal{A}$ to a subsystem $\mathcal{B}$.

In the important special case where $\mathcal{A} = \mathcal{L}(\mathcal{H}_A) \otimes \mathcal{L}(\mathcal{H}_B)$ and $\mathcal{B} = \mathbf{1}_A \otimes \mathcal{L}(\mathcal{H}_B)$, the corresponding map $T_*\colon \mathcal{A} \to \mathcal{B}$ is called the *partial trace* over $\mathcal{H}_A$. It satisfies

$$T_*(R_A \otimes R_B) = \mathbf{tr}_{\mathcal{H}_A}(R_A) R_B$$

for all $R_A \in \mathcal{L}(\mathcal{H}_A)$, $R_B \in \mathcal{L}(\mathcal{H}_B)$, so $T_* = \mathbf{tr}_{\mathcal{H}_A} \otimes \mathrm{Id}_{\mathcal{B}}$.

3.3.1 Example: information transfer between qubits

Here is a very simple model, due to von Neumann, for information transfer from one qubit $\mathcal{A} = M_2(\mathbb{C})$ to a second qubit $\mathcal{B} = M_2(\mathbb{C})$.

First we couple the system $\mathcal{A}$ to the system $\mathcal{B}$, which is initially in a pure state $|0\rangle$. Then we perform the unitary CNOT operation

$$U\,|ij\rangle = |i(j-i)\rangle$$

on the joint Hilbert space $\mathbb{C}^2 \otimes \mathbb{C}^2$ of the algebra $\mathcal{A} \otimes \mathcal{B}$. If the system $\mathcal{A}$ is initially in an unknown state ρ with density matrix $R_A \in \mathcal{A}$, then the resulting state on $\mathcal{A} \otimes \mathcal{B}$ has a density matrix $T_*(R_A)$ that is given by the CP map

$$T_*\colon \mathcal{A} \to \mathcal{A} \otimes \mathcal{B}, \quad T_*(R_A) = U(R_A \otimes |0\rangle\langle 0|)U^\dagger.$$

One checks that $T_*(|i\rangle\langle j|) = |ii\rangle\langle jj|$.

Problem 3.16. Check that $T_*(|i\rangle\langle j|) = |ii\rangle\langle jj|$.

At this point, we can do either one of two things:

1) Restrict attention to the commutative subalgebra $\mathbf{1}_A \otimes \mathcal{C}_2 \subseteq \mathcal{A} \otimes \mathcal{B}$. This models what happens if we measure $\sigma_z \in \mathcal{B}$ and forget about the original system.

2) Restrict attention to $\mathcal{A} \otimes \mathbf{1}_B \subseteq \mathcal{A} \otimes \mathcal{B}$. This models what happens if we forget about the system $\mathcal{B}$.

The first option yields the completely positive map $M_*\colon M_2(\mathbb{C}) \to \mathcal{C}_2$ with

$$M_*\colon \begin{pmatrix} r_{00} & r_{01} \\ r_{10} & r_{11} \end{pmatrix} \mapsto \mathrm{diag}(r_{00}, r_{11}).$$

In other words: the probability distribution (p_0, p_1) on the spectrum of $\mathbf{1} \otimes \sigma_z \in \mathcal{C}_2 \subseteq \mathcal{B}$ that arises *after* the operation from the density matrix $M_*(R_A)$ is precisely the same as the the probability distribution on the spectrum of $\sigma_z \in \mathcal{A}$ that arises from the density matrix $R_A \in \mathcal{A}$ *before* the operation. We conclude that information has been transferred from $\mathcal{A}$ to $\mathcal{B}$.

The second option allows us to infer what the consequence of this information transfer is on the system $\mathcal{A}$. Since $|ii\rangle\langle jj| = |i\rangle\langle j| \otimes |i\rangle\langle j|$, the partial trace

$\mathrm{Id}_2 \otimes \mathbf{tr}\colon M_2(\mathbb{C}) \otimes M_2(\mathbb{C}) \to M_2(\mathbb{C})$ of yields $|i\rangle\langle j|\,\delta_{ij}$. So the second option yields the CP map $R_*\colon M_2(\mathbb{C}) \to M_2(\mathbb{C})$ with

$$R_*\colon \begin{pmatrix} r_{00} & r_{01} \\ r_{10} & r_{11} \end{pmatrix} \mapsto \begin{pmatrix} r_{00} & 0 \\ 0 & r_{11} \end{pmatrix}.$$

So we have succeeded in transferring information about σ_z from $\mathcal{A}$ to $\mathcal{B}$, but at a high price: if we throw away the second qubit and agree to work only with $\mathcal{A}$ from now on, then we can never recover any information on the off-diagonal components r_{01} and r_{10}, so all information on σ_x and σ_y in the system $\mathcal{A}$ is lost.

In particular, suppose that the system $\mathcal{A}$ was initially in a pure state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$, corresponding to the density matrix

$$\begin{pmatrix} |\alpha|^2 & \alpha\overline{\beta} \\ \overline{\alpha}\beta & |\beta|^2 \end{pmatrix}.$$

Then after the information transfer, the system $\mathcal{A}$ is in the mixed state

$$\begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}.$$

If we agree to work with the system $\mathcal{A}$ from now on and discard the system $\mathcal{B}$ altogether, then for all practical intents and purposes it is as if the system $\mathcal{A}$ had collapsed into either the state $|0\rangle$ (with probability $p_0 = |\alpha|^2$) or into the state $|1\rangle$ (with probability $p_1 = |\beta|^2$).

Problem 3.17. A qubit $\mathcal{A} = M_2(\mathbb{C})$ is coupled to a second qubit $\mathcal{B} = M_2(\mathbb{C})$ in an initial state with density matrix Φ. Then a time evolution U takes place on $\mathbb{C}^2 \otimes \mathbb{C}^2$, resulting in a CP map $T_*\colon \mathcal{A} \to \mathcal{A} \otimes \mathcal{B}$ given (on density matrices) by

$$T_*(R_A) = U(R_A \otimes \Phi)U^\dagger.$$

a) If we restrict attention to the second system $\mathcal{B}$, the resulting CP map $M_*\colon \mathcal{A} \to \mathcal{B}$ on density matrices is given by $M_* = (\mathbf{tr}_{\mathcal{A}} \otimes \mathrm{Id}_{\mathcal{B}}) \circ T_*$. Find U and Φ so that[5] $M_*(|+\rangle\langle +|) = |0\rangle\langle 0|$ and $T_*(|-\rangle\langle -|) = |1\rangle\langle 1|$, and explain what this means in terms of information transfer.

b) If we restrict attention to the first system $\mathcal{A}$, then the resulting CP map $R_*\colon \mathcal{A} \to \mathcal{A}$ on density matrices is $R_* = (\mathrm{Id}_{\mathcal{A}} \otimes \mathbf{tr}_{\mathcal{B}}) \circ T_*$. Determine $M_*(|\psi\rangle\langle\psi|)$ for $|\psi\rangle = \alpha\,|+\rangle + \beta\,|-\rangle$.

3.3.2 Information transfer implies state collapse

This phenomenon is sometimes called the 'collapse of the wave function'. The following result shows that it is an unavoidable consequence of information transfer out of a system.

[5] As usual, $|+\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle := \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$.

Let $T_*\colon \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_A) \otimes \mathcal{C}_2$ be a trace-preserving, completely positive map. Identifying $\mathcal{L}(\mathcal{H}_A) \otimes \mathcal{C}_2$ with $\mathcal{L}(\mathcal{H}_A) \oplus \mathcal{L}(\mathcal{H}_A)$, we decompose T_* as

$$T_*(R_A) = \begin{pmatrix} T^0_*(R_A) & 0 \\ 0 & T^1_*(R_A) \end{pmatrix}.$$

Here $T^0_*\colon \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_A)$ and $T^1_*\colon \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_A)$ are completely positive but not trace-preserving: if we identify $\mathcal{C}_2$ with the algebra of random variables on $\Omega = \{\omega_0, \omega_1\}$, then ω_0 occurs with probability $p_0 = \mathbf{tr}(T^0_*(R_A))$, and ω_1 occurs with probability $p_1 = \mathbf{tr}(T^1_*(R_A))$. So T_* distinguishes ψ_0 from ψ_1 if the input state $|\psi_0\rangle\langle\psi_0|$ yields ω_0 with certainty, $p_0 = \mathbf{tr}(T^0_*(|\psi_0\rangle\langle\psi_0|)) = 1$, whereas the input state $|\psi_1\rangle\langle\psi_1|$ yields ω_1 with certainty, $p_1 = \mathbf{tr}(T^1_*(|\psi_1\rangle\langle\psi_1|)) = 1$.

Proposition 3.13 (Collapse of the wave function)**.** *Let*

$$T_*\colon \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_A) \otimes \mathcal{C}_2$$

be an operation with $\mathbf{tr}(T^0_*(|\psi_0\rangle\langle\psi_0|)) = 1$ *and* $\mathbf{tr}(T^1_*(|\psi_1\rangle\langle\psi_1|)) = 1$*. Then*

$$T_*\big(\,|\alpha\psi_0 + \beta\psi_1\rangle\langle\alpha\psi_0 + \beta\psi_1|\,\big) = \begin{pmatrix} |\alpha|^2 T^0_*(|\psi_0\rangle\langle\psi_0|) & 0 \\ 0 & |\beta|^2 T^1_*(|\psi_1\rangle\langle\psi_1|) \end{pmatrix}$$

for all $\alpha, \beta \in \mathbb{C}$ with $|\alpha|^2 + |\beta|^2 = 1$.

From this we can draw two conclusions. First of all, note that

$$T_*(|\alpha\psi_0 + \beta\psi_1\rangle\langle\alpha\psi_0 + \beta\psi_1|) = |\alpha|^2 T_*(|\psi_0\rangle\langle\psi_0|) + |\beta|^2 T_*(|\psi_1\rangle\langle\psi_1|).$$

So *after* the information transfer, it is no longer possible to distinguish a system that *started* in a pure state $\rho_{\mathrm{pure}} = |\alpha\psi_0 + \beta\psi_1\rangle\langle\alpha\psi_0 + \beta\psi_1|$ from one that *started* in a mixed state $\rho_{\mathrm{mixed}} = |\alpha|^2\,|\psi_0\rangle\langle\psi_0| + |\beta|^2\,|\psi_1\rangle\langle\psi_1|$. For all intents and purposes, we may therefore pretend that the system had instantaneously jumped from the pure state ρ_{pure} to the mixed state ρ_{mixed} at the start of the information transfer.

The second conclusion is that conditioned on the outcome ω_0, the system $\mathcal{L}(\mathcal{H}_A)$ will further behave as if it had been in state ψ_0 before measurement, and conditioned on the outcome ω_1, it will behave as if it had been in state ψ_1.

Proof. Since $T_*(\big|\varepsilon e^{i\phi}\psi_0 + \psi_1\big\rangle\big\langle\varepsilon e^{i\phi}\psi_0 + \psi_1\big|) \geq 0$ for all $\varepsilon, \phi \in \mathbb{R}$, we also have $T^0_*(\big|\varepsilon e^{i\phi}\psi_0 + \psi_1\big\rangle\big\langle\varepsilon e^{i\phi}\psi_0 + \psi_1\big|) \geq 0$. Expanding into $|\psi_i\rangle\,\langle\psi_j|$ yields

$$\big|\varepsilon e^{i\phi}\psi_0 + \psi_1\big\rangle\big\langle\varepsilon e^{i\phi}\psi_0 + \psi_1\big| = \varepsilon^2\,|\psi_0\rangle\langle\psi_0| + \varepsilon\left(e^{i\phi}\,|\psi_0\rangle\,\langle\psi_1| + e^{-i\phi}\,|\psi_1\rangle\,\psi_0\right) + |\psi_1\rangle\langle\psi_1|\,.$$

If we apply T^0_* to this expression, the result is positive semidefinite. Since $\mathbf{tr}(T^0_*(|\psi_1\rangle\langle\psi_1|)) = 0$ implies $T^0_*(|\psi_1\rangle\langle\psi_1|) = 0$ by positivity, we find

$$\varepsilon^2 T^0_*(|\psi_0\rangle\langle\psi_0|) + \varepsilon T^0_*\left(e^{i\phi}\,|\psi_0\rangle\,\langle\psi_1| + e^{-i\phi}\,|\psi_1\rangle\,\langle\psi_0|\right) \geq 0.$$

But since this holds for all $\varepsilon \in \mathbb{R}$, we find $T^0_*\left(e^{i\phi}\,|\psi_0\rangle\,\langle\psi_1| + e^{-i\phi}\,|\psi_1\rangle\,\langle\psi_0|\right) \geq 0$ for all $\phi \in \mathbb{R}$. Choosing $\phi = 0, \pi/2, \pi, 3\pi/2$ and taking linear combinations, we conclude that $T^0_*(|\psi_0\rangle\,\langle\psi_1|) = 0$ and $T^0_*(|\psi_0\rangle\,\langle\psi_1|) = 0$.

It follows that $T^0_*(|\alpha\psi_0 + \beta\psi_1\rangle\langle\alpha\psi_0 + \beta\psi_1|) = |\alpha|^2 T^0_*(|\psi_0\rangle\langle\psi_0|)$. Similarly, $T^1_*(|\alpha\psi_0 + \beta\psi_1\rangle\langle\alpha\psi_0 + \beta\psi_1|) = |\beta|^2 T^1_*(|\psi_1\rangle\langle\psi_1|)$, so the result follows. □

3.4 Interaction with classical systems

We now investigate the interaction between classical and quantum systems from a more axiomatic point of view.

3.4.1 From quantum to classical: POVMs

Recall that $\mathcal{C}_d$ is the algebra of random variables on a classical probability space Ω with d possible outcomes. We denote by $\delta_\omega \colon \Omega \to \mathbb{C}$ the random variable that is 1 on ω and otherwise zero.

In the Heisenberg picture, a transformation from a quantum system to a classical system is described by a normalized CP map $M \colon \mathcal{C}_d \to \mathcal{L}(\mathcal{H})$. The transformation M is positive if and only if the operators

$$E_\omega := M(\delta_\omega) \tag{64}$$

satisfy $E_\omega \geq 0$, and it is normalized if and only if $\sum_{\omega \in \Omega} E_\omega = \mathbf{1}$.

Definition 3.8 (POVM)**.** A *positive operator valued measure* (POVM) on the set $\Omega = \{\omega_1, \ldots, \omega_d\}$ is a collection $\{E_\omega\,;\, \omega \in \Omega\}$ of operators $E_\omega \in \mathcal{L}(\mathcal{H})$ with $E_\omega \geq 0$ and $\sum_{\omega \in \Omega} E_\omega = \mathbf{1}$.

Since $\mathcal{C}_d$ is commutative, the map $M \colon \mathcal{C}_d \to \mathcal{L}(\mathcal{H})$ is completely positive if and only if it is positive. So normalized CP maps $M \colon \mathcal{C}_d \to \mathcal{L}(\mathcal{H})$ correspond bijectively to POVMs by equation (64); in terms of the POVM, the CP map is given by

$$M(f) = \sum_{\omega \in \Omega} f(\omega) E_\omega.$$

For an input state ρ with density matrix R, the probability that an outcome $\omega \in \Omega$ occurs is $p_\omega = \mathbf{tr}(R E_\omega)$.

Remark 3.3. Note that every PVM (in the sense of Definition 2.4) is a POVM, but not every POVM is a PVM.

3.4.2 Retaining the quantum system

In the Heisenberg picture, a normalized CP map

$$T \colon \mathcal{L}(\mathcal{H}) \otimes \mathcal{C}_d \to \mathcal{L}(\mathcal{H})$$

describes a transformation from a quantum system $\mathcal{L}(\mathcal{H})$ to the same quantum system coupled to a classical system $\mathcal{C}_d$. So we keep track of the quantum system as well as the measurement outcomes.

If we restrict attention to the classical subsystem $\mathcal{C}_d \otimes \mathbf{1} \subseteq \mathcal{C}_d \otimes \mathcal{L}(\mathcal{H})$, we obtain the CP map $M \colon \mathcal{C}_d \to \mathcal{L}(\mathcal{H})$ with $M(f) := T(f \otimes \mathbf{1})$. The corresponding POVM on the set Ω with d outcomes is $E_\omega = T(\delta_\omega \otimes \mathbf{1})$. Conversely, if we restrict attention to the quantum system $\mathcal{L}(\mathcal{H})$, we obtain the CP map $S \colon \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ defined by $S(A) := T(\mathbf{1} \otimes A)$.

But we can also *condition* on a measurement outcome $\omega \in \Omega$. Let $\rho \in \mathcal{S}(\mathcal{L}(\mathcal{H}))$ be the initial state on $\mathcal{L}(\mathcal{H})$. For a projection $P \in \mathcal{L}(\mathcal{H})$, the probability that P *and* ω occur in the final state $T^*\rho$ is $T^*\rho(P \otimes \delta_\omega)$. Since the probability that ω occurs is $T^*\rho(\mathbf{1} \otimes \delta_\omega)$, the quotient

$$\mathbb{P}(P \,|\, \omega) = \frac{T^*\rho(P \otimes \delta_\omega)}{T^*\rho(\mathbf{1} \otimes \delta_\omega)}$$

is the probability that P occurs *conditioned* on the outcome ω. Similarly,

$$\mathbb{E}(A \,|\, \omega) = \frac{T^*\rho(A \otimes \delta_\omega)}{T^*\rho(\mathbf{1} \otimes \delta_\omega)}$$

is the expectation of $A \in \mathcal{L}(\mathcal{H})$ conditioned on ω.

We therefore interpret the CP map

$$T_\omega \colon \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})\,; \quad T_\omega(A) = T(\delta_\omega \otimes A)$$

as *conditioning* on the measurement outcome $\omega \in \Omega$. Note that this is *not* a normalized CP map. In fact, $T_\omega(\mathbf{1}) = T(\mathbf{1} \otimes \delta_\omega)$ is the operator $E_\omega \geq 0$ that represents the outcome ω in the POVM assocated to M. In the Schrödinger picture, this manifests itself in the normalization; if R is a normalized density matrix, then $\mathbf{tr}(T_{\omega *}R)$ is equal to the probability p_ω that ω occurs in the state M_*R on $\mathcal{C}_d$.

Conditional expectation Let $(p_1, \ldots, p_n)$ be a probability distribution on the finite set $\Omega = \{1, \ldots, n\}$. So far we have silently assumed that the σ-algebra is $\Sigma = \mathcal{P}(\Omega)$, the set of all subsets of Ω. Then C_n is the algebra of all complex valued random variables $F \colon \Omega \to \mathbb{C}$. Let Σ' be the coarser σ-algebra generated by a partition of Ω into k subsets, $\Omega = A_1 \sqcup \ldots \sqcup A_k$, and let $\mathcal{A} \subseteq \mathcal{C}_n$ be the subalgebra of all random variables $F \colon \Omega \to \mathbb{C}$ that are *measurable* with respect to Σ. For every $\omega \in \Omega$, let $A(\omega)$ be the unique A_i that contains ω. Then the *conditional expectation* with respect to $\mathbb{P}$ is the map that averages F over the A_i,

$$E \colon \mathcal{C}_n \to \mathcal{A}\,; \quad E(F)(\omega) := \frac{\sum_{\omega' \in A(\omega)} p(\omega') f(\omega')}{\sum_{\omega' \in A(\omega)} p(\omega')}.$$

Problem 3.18. Show that E is positive and normalized. Show that E is a projection, $E^2 = E$. Show that $E(G_1 F G_2) = G_1 E(F) G_2$ for all $G \in \mathcal{A}$. Show that E is the *orthogonal* projection from $\mathcal{C}_n$ onto $\mathcal{A}$ if we equip $\mathcal{C}_n$ with the inner product $\langle F, G \rangle = \sum_{\omega \in \Omega} \overline{F}(\omega) G(\omega) p(\omega)$.

3.5 Dilations

If we restrict attention to a smaller subsystem $\mathcal{A} \subset \mathcal{B}$, then *pure* states on $\mathcal{B}$ can restrict to *mixed* states on $\mathcal{A}$. We now consider the converse problem: for a given mixed state ρ on $\mathcal{A}$, can we find a larger system $\mathcal{B}$ such that ρ extends to a *pure* state on $\mathcal{B}$?

3.5.1 The GNS-representation

Let $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H}_A)$ and $\mathcal{B} \subseteq \mathcal{L}(\mathcal{H}_B)$ be $*$-algebras.

Definition 3.9. A $*$-homomorphism $\pi \colon \mathcal{A} \to \mathcal{B}$ is a linear map such that

1) $\pi(AB) = \pi(A)\pi(B)$ for all $A, B \in \mathcal{A}$.

2) $\pi(A^\dagger) = \pi(A)^\dagger$

3) $\pi(\mathbf{1}_A) = \mathbf{1}_B$.

A $*$-isomorphism is an invertible $*$-homomorphism.

Note that if $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H}_A)$ and $\mathcal{B} \subseteq \mathcal{L}(\mathcal{H}_B)$ are isomorphic as $*$-algebras, then the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ can be quite different! For example, $\mathcal{A} = \mathcal{L}(\mathcal{H})$ is isomorphic to $\mathcal{A} \otimes \mathbf{1} \subseteq \mathcal{L}(\mathcal{H} \otimes \mathcal{K})$.

The GNS-representation (after Israel Gel'fand, Mark Naimark and Irving Segal) allows one to represent *every* state as a vector state, at the cost of enlarging the Hilbert space.

Theorem 3.14 (GNS representation)**.** *For every $\rho \in \mathcal{S}(\mathcal{A})$, there exist:*

1) a Hilbert space $\mathcal{H}_\rho$

2) a $*$-homomorphism $\pi_\rho \colon \mathcal{A} \to \mathcal{L}(\mathcal{H}_\rho)$

3) and a unit vector $\Omega \in \mathcal{H}_\rho$

such that $\rho(A) = \langle \Omega, \pi_\rho(A)\Omega \rangle$ for all $A \in \mathcal{A}$.

Physically, this means that for every *mixed* state ρ on $\mathcal{A}$, there exists a (fictitious) larger quantum system $\mathcal{L}(\mathcal{H}_\rho)$ in a *pure* state $\rho_\Omega(B) = \langle \Omega, B\Omega \rangle$ such ρ is the restriction of ρ_Ω to the subsystem $\mathcal{A} \subseteq \mathcal{L}(\mathcal{H}_\rho)$.

We start with two basic results on positivity.

Lemma 3.15. *If $X \in \mathcal{L}(\mathcal{H})$ with $X \geq 0$, then $B^\dagger X B \geq 0$ for all $B \in \mathcal{L}(\mathcal{H})$.*

Proof. As $X \geq 0$, we have $\langle \psi, B^\dagger X B \psi \rangle = \langle B\psi, X(B\psi) \rangle \geq 0$ for all $\psi \in \mathcal{H}$. □

So if X and Y are Hermitian with $X \leq Y$, then the the above applied to $Y - X$ shows that $B^\dagger X B \leq B^\dagger Y B$. Recall that the operator norm of $X \in \mathcal{L}(\mathcal{H})$ is defined by

$$\|X\| := \sup\{\|X\psi\| \,;\, \psi \in \mathcal{H}, \|\psi\| = 1\}.$$

Lemma 3.16. *If $X \in \mathcal{L}(\mathcal{H})$ is Hermitian, then $B^\dagger X B \leq \|X\| B^\dagger B$.*

Proof. Since $\langle \psi, X\psi \rangle \leq \|X\| \langle \psi, \psi \rangle$ by definition, we have $X \leq \|X\|\mathbf{1}$. Now apply Lemma 3.15 to $\|X\|\mathbf{1} - X$. □

A state $\rho \in \mathcal{S}(\mathcal{A})$ is *faithful* if $\rho(X^\dagger X) = 0$ implies $X = 0$. If $\rho(A) = \mathbf{tr}(RA)$, then ρ is faithful if and only if $0 \notin \mathrm{spec}(R)$. For faithful states ρ, the GNS representation is particularly straightforward.

Proof of Theorem 3.14 for faithful states. Simply take

$$\mathcal{H}^0_\rho = \mathcal{A} \quad \text{with} \quad \langle X, Y\rangle_\rho := \rho(X^\dagger Y).$$

This is linear on the right hand side, and $\langle Y, X\rangle_\rho = \overline{\langle X, Y\rangle_\rho}$ because

$$\rho(Y^\dagger X) = \rho((X^\dagger Y)^\dagger) = \overline{\rho(X^\dagger Y)}.$$

For any state ρ this is positive semidefinite, $\langle X, X\rangle_\rho = \rho(X^\dagger X) \geq 0$. But to show that it is nondegenerate, we need ρ to be faithful; $\langle X, X\rangle_\rho = 0$ if and only if $\rho(X^\dagger X) = 0$, which implies $X = 0$ if ρ is faithful.

For the $*$-homomorphism we take $\pi^0_\rho(A)X := AX$. This clearly satisfies $\pi^0_\rho(AB)X = \pi^0_\rho(A)\pi^0_\rho(B)X$, and $\pi^0_\rho(A^\dagger) = \pi^0_\rho(A)^\dagger$ because

$$\langle \pi^0_\rho(A^\dagger)X, Y\rangle = \rho((A^\dagger X)^\dagger Y) = \rho(X^\dagger A Y) = \langle X, \pi^0_\rho(A)Y\rangle.$$

If we choose the unit vector $\Omega = \mathbf{1}$ in $\mathcal{H}^0_\rho$, we have

$$\langle \Omega, \pi_\rho(A)\Omega\rangle = \langle \mathbf{1}, A\rangle_\rho = \rho(A)$$

and we are done. □

If ρ is not faithful, then the Hermitian form on $\mathcal{H}^0_\rho$ has a nonzero kernel

$$\mathcal{N}_\rho = \{X \in \mathcal{H}^0_\rho\,;\, \langle X, X\rangle_\rho = 0\}.$$

To get an inner product, we consider the quotient vector space $\mathcal{H}_\rho := \mathcal{H}^0_\rho / \mathcal{N}_\rho$.

The Hermitian form

$$\langle [X], [Y]\rangle_\rho = \langle X, Y\rangle_\rho \tag{65}$$

is well defined: if $Y \in \mathcal{N}_\rho$, then $\langle Y, Y\rangle_\rho = 0$, so $|\langle X, Y\rangle_\rho|^2 \leq \langle X, X\rangle_\rho \langle Y, Y\rangle_\rho$ is zero by the Cauchy-Schwarz inequality. So (65) does not depend on the representative Y of $[Y]$, and in the same way one sees that it is independent of the representative X of $[X]$. Since we removed all the vectors with zero norm, the result is of course nondegenerate: $\langle [X], [X]\rangle_\rho = 0$ implies $\langle X, X\rangle_\rho = 0$, so $X \in \mathcal{N}_\rho$ and $[X] = 0$.

Proof of Theorem 3.14 for general states. It remains to show that $\pi_\rho(A)[X] := [AX]$ yields a well-defined $*$-homomorphism $\pi_\rho \colon \mathcal{A} \to \mathcal{L}(\mathcal{H}_\rho)$. Since

$$\langle \pi^0_\rho(A)X, \pi^0_\rho(A)X\rangle_\rho = \langle AX, AX\rangle_\rho = \rho(X^\dagger A^\dagger A X) \leq \|A\|^2 \langle X, X\rangle_\rho \tag{66}$$

by Lemma 3.16, we have $\langle \pi^0_\rho(A)X, \pi^0_\rho(A)X\rangle_\rho = 0$ whenever $\langle X, X\rangle_\rho = 0$. So $X \in \mathcal{N}_\rho$ implies $\pi^0_\rho(A)X \in \mathcal{N}_\rho$, and $\pi_\rho(A)[X] := [AX]$ does not depend on the representative $X \in \mathcal{H}^0_\rho$ of the class $[X] \in \mathcal{H}^0_\rho/\mathcal{N}_\rho$. □

Remark 3.4. As a byproduct of (66), we find that $\|\pi_\rho(A)\| \leq \|A\|$. The operator norm of $\pi_\rho(A) \colon \mathcal{H}_\rho \to \mathcal{H}_\rho$ on the GNS Hilbert space $\mathcal{H}_\rho$ is bounded by the operator norm of $A \colon \mathcal{H} \to \mathcal{H}$ on the original Hilbert space $\mathcal{H}$.

Problem 3.19. The kernel $\mathcal{N}_\rho$ of the sesquilinear form $\langle X, Y\rangle_\rho = \rho(X^\dagger Y)$ is zero if and only if ρ is faithful.

Problem 3.20. The vector state $\rho(A) = \langle\psi, A\psi\rangle$ on $\mathcal{A} = M_n(\mathbb{C})$ is not faithful. Determine the kernel $\mathcal{N}_\rho$ of the sesquilinear form $\langle X, Y\rangle_\rho = \rho(X^\dagger Y)$, and show that the map $\mathcal{H}_\rho \to \mathbb{C}^n$ defined by $[X] \mapsto X\psi$ is a well-defined isomorphism of Hilbert spaces.

Problem 3.21. The tracial state $\tau(A) = \frac{1}{n}\mathbf{tr}(A)$ on $M_n(\mathbb{C})$ is faithful, so in this case $\mathcal{H}_\tau = M_n(\mathbb{C})$. Determine the inner product $\langle A, B\rangle_\tau$ in terms of the matrix coeficients a_{ij}, b_{kl}.

Problem 3.22. Let $\mathcal{H}_\rho$ be the GNS Hilbert space for the state ρ on $M_3(\mathbb{C})$ with density matrix

$$R = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

What is the dimension of $\mathcal{H}_\rho$?

The GNS Hilbert space $\mathcal{H}_\rho$ depends on the state ρ. The following problem shows that for $M_n(\mathbb{C})$, there exists a *single* Hilbert space $\mathcal{H}_\tau$ on which *every* state $\rho \in \mathcal{S}(M_n(\mathbb{C}))$ can be represented by a vector state.

Problem 3.23. Let ρ be a state on $M_n(\mathbb{C})$ with density matrix R, and let τ be the tracial state $\tau(A) = \frac{1}{n}\mathbf{tr}(A)$. Show that the map

$$V\colon \mathcal{H}_\rho \to \mathcal{H}_\tau, \quad V([A]) := [A\sqrt{nR}]$$

is a well defined isometry, which is an isomorphism if and only if ρ is faithful. Show that $V\pi_\rho(A) = \pi_\tau(A)V$, and conclude that $\rho(A) = \langle\psi, \pi_\tau(A)\psi\rangle$ for the vector $\psi = V\Omega$ in $\mathcal{H}_\tau$.

3.5.2 The Stinespring dilation

An *isometry* $V\colon \mathcal{H} \to \mathcal{K}$ is a linear map such that

$$\langle V\psi, V\chi\rangle_{\mathcal{K}} = \langle\psi, \chi\rangle_{\mathcal{H}} \quad \text{for all} \quad \psi, \chi \in \mathcal{H}. \tag{67}$$

Since V is injective, we can consider $\mathcal{H} \simeq V(\mathcal{H})$ as a subspace of $\mathcal{K}$.

Problem 3.24. Under this identification, the adjoint $V^\dagger\colon \mathcal{K} \to \mathcal{H}$ is the orthogonal projection onto $V(\mathcal{H}) \subseteq \mathcal{K}$.

Problem 3.25. The map $V\colon \mathcal{H} \to \mathcal{K}$ is an isometry if and only if $V^\dagger V = \mathbf{1}_{\mathcal{H}}$.

Problem 3.26. For $\chi \in \mathcal{H}_B$, define $V_\chi\colon \mathcal{H}_A \to \mathcal{H}_A \otimes \mathcal{H}_B$ by $V_\chi\psi = \psi \otimes \chi$.

a) Show that $V_\chi^\dagger(\psi_A \otimes \psi_B) = \langle\chi, \psi_B\rangle\psi_A$.

b) V_χ is an isometry if and only if $\|\chi\| = 1$.

If $V\colon \mathcal{H} \to \mathcal{K}$ is an isometry, then the map

$$T\colon \mathcal{L}(\mathcal{K}) \to \mathcal{L}(\mathcal{H}), \quad T(A) = V^{\dagger}AV$$

is normalized and completely positive. The *Stinespring Dilation Theorem* asserts that at the cost of enlarging the Hilbert space, *every* CP map is essentially of this form.

Theorem 3.17 (Stinespring)**.** *For every CP map $T\colon \mathcal{A} \to \mathcal{L}(\mathcal{H})$ with $T(\mathbf{1}) = \mathbf{1}$, there exist:*

1) a Hilbert space $\mathcal{H}_T$

2) a $*$-homomorphism $\pi_T\colon \mathcal{A} \to \mathcal{L}(\mathcal{H}_T)$

3) and an isometry $V\colon \mathcal{H} \to \mathcal{H}_T$

such that $T(A) = V^{\dagger}\pi(A)V$.

Remark 3.5. A normalized CP map $T\colon \mathcal{A} \to \mathbb{C}$ is the same as a state on $\mathcal{A}$. It is instructive to check that in this special case, the Stinespring dilation agrees with the GNS construction.

Proof. Equip the complex vector space $\mathcal{H}_T^0 := \mathcal{A} \otimes \mathcal{H}$ with the Hermitian form

$$\langle X \otimes \psi, X' \otimes \psi' \rangle_T := \langle \psi, T(X^{\dagger}X')\psi' \rangle_{\mathcal{H}}. \tag{68}$$

Define the linear map $V_0\colon \mathcal{H} \to \mathcal{H}_T^0$ by

$$V_0\psi = \mathbf{1} \otimes \psi,$$

and note that $\langle V_0\psi, V_0\psi' \rangle_T = \langle \psi, \psi' \rangle_{\mathcal{H}}$ because $T(\mathbf{1}) = \mathbf{1}$. Set

$$\pi_0(A)(X \otimes \psi) = AX \otimes \psi.$$

We will show momentarily that (68) is positive semidefinite, but before we do so we can already check that π_0 and V_0 satisfy some of the properties we are after. It is not hard to see that $\pi_0(A)\pi_0(B) = \pi_0(AB)$, and that

$$\langle \pi_0(A^{\dagger})X \otimes \psi, X' \otimes \psi' \rangle_T = \langle X \otimes \psi, \pi_0(A)X' \otimes \psi' \rangle_T. \tag{69}$$

This shows that $\pi_0(A^{\dagger})$ is a *formal adjoint* of $\pi_0(A)$. In the same vein, the linear map $V_0^{\dagger}\colon \mathcal{H}_T^0 \to \mathcal{H}$ with $V_0^{\dagger}(X \otimes \psi) := T(X)\psi$ is a formal adjoint of V_0,

$$\langle V_0^{\dagger} X \otimes \psi, \psi' \rangle_{\mathcal{H}} = \langle X \otimes \psi, V_0\psi' \rangle_T. \tag{70}$$

With these definitions, it immediately follows that

$$T(A) = V_0^{\dagger}\pi^0(A)V_0.$$

It remains to check that (68) is positive semidefinite, and that the kernel

$$\mathcal{N} := \{F \in \mathcal{H}_T^0\,;\, \langle F, F\rangle_T = 0\}$$

is annihilated by $\pi_T^0(A)$ and $V_0^\dagger$. The quotient $\mathcal{H}_T := \mathcal{H}_T^0/\mathcal{N}$ is then an inner product space, and the linear maps $\pi_T(A)\colon \mathcal{H}_T \to \mathcal{H}_T$ and $V^\dagger\colon \mathcal{H}_T \to \mathcal{H}$ are well defined by $\pi_T(A)([X\otimes\psi]) = [\pi_0(A)(X\otimes\psi)]$ and $V^\dagger([X\otimes\psi]) = [V_0^\dagger(X\otimes\psi)]$.

Step 1

We show that the Hermitian form is positive semidefinite. For $F = \sum_{i=1}^N X_i\otimes\psi_i$, we have

$$\langle F, F\rangle_T = \sum_{i=1}^N\sum_{j=1}^N \langle \psi_i, T(X_i^\dagger X_j)\psi_j\rangle_{\mathcal{H}}. \tag{71}$$

To see that this is nonnegative, consider the vector $\Psi := \sum_{i=1}^N \psi_i \otimes e_i$ in $\mathcal{H}\otimes\mathbb{C}^N$ and the operator $\Xi := \sum_{i=1}^N X_i \otimes |e_1\rangle\langle e_i|$ in $\mathcal{A}\otimes M_N(\mathbb{C})$. Since $\Xi^\dagger\Xi = \sum_{i=1}^N\sum_{j=1}^N X_i^\dagger X_j \otimes |e_i\rangle\langle e_j|$, complete positivity of T. guarantees that

$$\langle F, F\rangle_T = \langle \Psi, (T\otimes \mathrm{Id})(\Xi^\dagger\Xi)\Psi\rangle \geq 0,$$

so (68) is indeed positive semidefinite.

Step 2.

Next, we show that $\langle \pi_0(A)F, \pi_0(A)F\rangle_T \leq \|A\|^2\langle F, F\rangle_T$. This ensures that $\pi_0(A)\mathcal{N} \subseteq \mathcal{N}$, and that $\pi_T(A)[F] = [\pi_0(A)F]$ defines an operator on $\mathcal{H}_T := \mathcal{H}_T^0/\mathcal{N}$ of norm at most $\|A\|$.

Since $\pi_0(A)F = \sum_{i=1}^N AX_i\otimes\psi_i$, we can calculate $\langle \pi_0(A)F, \pi_0(A)F\rangle_T$ by repeating step 1 with X_i replaced by AX_i. Then Ξ is replaced by $(A\otimes\mathbf{1})\Xi$, and we find

$$\langle \pi_0(A)F, \pi_0(A)F\rangle_T = \big\langle \Psi\,,\, (T\otimes\mathrm{Id})(\Xi^\dagger\big(A^\dagger A\otimes\mathbf{1})\Xi\big)\Psi\big\rangle\,.$$

Since $\Xi^\dagger\big(A^\dagger A\otimes\mathbf{1})\Xi \leq \|A\|^2\Xi^\dagger\Xi$, we find $\langle \pi_0(A)F, \pi_0(A)F\rangle_T \leq \|A\|^2\langle F, F\rangle_T$ as required.

Step 3.

Finally, we show that $V_0^\dagger\mathcal{N} \subseteq \mathcal{N}$. By the Cauchy-Schwarz inequality for positive semidefinite forms, $\langle\psi,\psi\rangle_T = 0$ implies $\langle\psi,\psi'\rangle_T = 0$ for all $\psi' \in \mathcal{H}_T^0$. So if $\psi\in\mathcal{N}$ then$\langle V_0^\dagger\psi, \chi\rangle = \langle\psi, V\chi\rangle = 0$ for all $\chi\in\mathcal{H}$, and $V_0^\dagger\psi\in\mathcal{N}$ as required. □

Problem 3.27. Verify equations (69) and (70).

Problem 3.28. Formulate (and prove) a version of Stinespring's Theorem for CP maps without the requirement that $T(\mathbf{1}) = \mathbf{1}$.

3.5.3 Choi's Theorem and Kraus operators

It is enlightening to consider the special case of a normalized CP map

$$T\colon M_n(\mathbb{C}) \to M_k(\mathbb{C}).$$

Recall from §2.6.2 that $\mathbb{C}^n \otimes \mathbb{C}^n \simeq M_n(\mathbb{C})$, where the vector $\xi_1 \otimes \xi_2 \in \mathbb{C}^n \otimes \mathbb{C}^n$ corresponds to the rank one operator $|\xi_1\rangle \langle\overline{\xi}_2| \in M_n(\mathbb{C})$. If we identify

$$\mathcal{H}^0_T = M_n(\mathbb{C}) \otimes \mathbb{C}^k \quad \text{with} \quad \mathcal{H}^0_T = \mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^k,$$

then the map $\pi_0(A)$ acts only the *first* tensor leg of $(\mathbb{C}^n) \otimes (\mathbb{C}^n \otimes \mathbb{C}^k)$. Indeed,

$$\pi_0(A)\,|\xi_1\rangle \langle\overline{\xi}_2| \otimes \psi = |A\xi_1\rangle \langle\overline{\xi}_2| \otimes \psi$$

becomes

$$\pi_0(A)(\xi_1 \otimes \xi_2 \otimes \psi) = (A\xi_1) \otimes \xi_2 \otimes \psi$$

if we identify $M_n(\mathbb{C}) \otimes \mathbb{C}^k$ with $\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^k$.

On the other hand, the twist in the inner product affects only the *last* two tensor legs of $(\mathbb{C}^n) \otimes (\mathbb{C}^n \otimes \mathbb{C}^k)$. Indeed, the (degenerate) inner product on $\mathcal{H}^0_T$ is given by

$$\langle \xi_1 \otimes \xi_2 \otimes \psi, \xi_1 \otimes \xi_2 \otimes \psi \rangle_T = \langle \psi, T\big((|\xi_1\rangle \langle\overline{\xi_2}|)^\dagger (|\xi_1'\rangle \langle\overline{\xi}_2'|)\big)\psi\rangle_{\mathbb{C}^k}.$$

Since $(|\xi_1\rangle\langle\overline{\xi_2}|)^\dagger\, |\xi_1'\rangle\langle\overline{\xi}_2'| = |\overline{\xi}_2\rangle \langle\xi_1, \xi_1'\rangle \langle\overline{\xi}_2'|$, we can take the complex number $\langle\xi_1, \xi_1'\rangle$ out of the equation and obtain

$$\langle \xi_1 \otimes \xi_2 \otimes \psi, \xi_1 \otimes \xi_2 \otimes \psi \rangle_T = \langle\xi_1, \xi_1'\rangle_{\mathbb{C}^n} \langle \psi, T(|\overline{\xi}_2\rangle \langle\overline{\xi}_2'|)\psi\rangle_{\mathbb{C}^k}.$$

Apparently, we can view $\mathcal{H}^0_T = (\mathbb{C}^n) \otimes (\mathbb{C}^n \otimes \mathbb{C}^k)$ as the tensor product with in the first leg $\mathbb{C}^n$ with the *ordinary* inner product, and on the second and third leg $\mathbb{C}^n \otimes \mathbb{C}^k$ with the (degenerate) inner product

$$\langle \xi_2 \otimes \psi, \xi_2' \otimes \psi' \rangle_T = \big\langle \psi', T\big(\,|\overline{\xi}_2\rangle \langle\overline{\xi}_2'|\,\big)\psi \big\rangle\,. \tag{72}$$

With these observations, it is not hard to prove Choi's Theorem.

Theorem 3.18 (Choi)**.** *For every CP map $T\colon M_n(\mathbb{C}) \to M_k(\mathbb{C})$ with $T(\mathbf{1}) = \mathbf{1}$, there exist operators $V_i\colon \mathbb{C}^k \to \mathbb{C}^n$ such that*

$$T(A) = \sum_{i=1}^{d} V_i^\dagger A V_i \quad \textit{for all} \quad A \in M_n(\mathbb{C}).$$

The CP map $\widehat{T}\colon \mathcal{C}_d \otimes M_n(\mathbb{C}) \to M_k(\mathbb{C})$ with $\widehat{T}(\delta_i \otimes A) = V_i^\dagger A V_i$ is then a *dilation* of T, in the sense that $T(A) = \widehat{T}(\mathbf{1} \otimes \mathcal{A})$. For the case $n = k$, this has the following interpretation. In the Schrödinger picture, we start with the system $M_n(\mathbb{C})$ in the initial state ρ. Then $\widehat{T}^*$ produces an outcome ω_i with probability $p_i = \rho(V_i^\dagger V_i)$, in which case the system $M_n(\mathbb{C})$ is left in the final state

$$\rho_i(A) = \frac{\rho(V_i^\dagger A V_i)}{\rho(V_i^\dagger V_i)}$$

Proof. Let Ψ_i be an orthonormal basis of $\mathbb{C}^n \otimes \mathbb{C}^k/\mathcal{N}$ with respect to the inner product (72), and let $W_i \colon \mathbb{C}^n \to \mathbb{C}^n \otimes (\mathbb{C}^n \otimes \mathbb{C}^k/\mathcal{N})$ be the isometry defined by $W_i\psi = \psi \otimes \Psi_i$. Then since $W_i^\dagger \psi \otimes \Phi = \langle \Psi_i, \Phi\rangle\psi$, we have

$$\pi_T(A) = \sum_{i=1}^{d} W_i A W_i^\dagger. \tag{73}$$

So by Stinespring's Theorem $T(A) = \sum_{i=1}^{d} V W_i^\dagger A W_i^\dagger V^\dagger$, and Choi's Theorem holds with $V_i = V W_i^\dagger$. □

Problem 3.29. Verify (73), for example by using $\sum_{i=1}^{d} |\Psi_i\rangle \langle\Psi_i| = \mathbf{1}$.

Problem 3.30. We first couple a system $\mathcal{A} = \mathcal{L}(\mathcal{H}_A)$ to an auxilliary system $\mathcal{B} = \mathcal{L}(\mathcal{H}_B)$ in a state $\phi \in \mathcal{S}(\mathcal{B})$, then perform unitary time evolution on $\mathcal{H}_A \otimes \mathcal{H}_B$ and finally restrict attention to the subsystem $\mathcal{A} \subseteq \mathcal{A} \otimes \mathcal{B}$. In the Heisenberg picture the corresponding CP map $T \colon \mathcal{A} \to \mathcal{A}$ is

$$T(A) = \mathrm{Id} \otimes \phi(U^\dagger A \otimes \mathbf{1}_B U).$$

We explicitly construct $V_i \in \mathcal{A}$ such that $T(A) = \sum_{i=1}^{n} V_i^\dagger A V_i$.

a) Without loss of generality we may assume that ϕ is a vector state, $\phi(B) = \langle \Omega, A\Omega\rangle$ for some $\Omega \in \mathcal{H}_B$.

b) Pick an orthonormal basis $|e_i\rangle$ of $\mathcal{H}_B$. Define $\mathbf{1}_A \otimes |e_i\rangle$ as the linear map $\mathcal{H}_A \to \mathcal{H}_A \otimes \mathcal{H}_B$ with $\psi \mapsto \psi \otimes e_i$, and $\mathbf{1}_A \otimes \langle e_i|$ as the linear map $\mathcal{H}_A \otimes \mathcal{H}_B$ with $\psi \otimes \chi \mapsto \langle e_i, \chi\rangle\psi$. Show that

$$A \otimes \mathbf{1} = \sum_{i=1}^{n} (\mathbf{1}_A \otimes |e_i\rangle) A (\mathbf{1}_A \otimes \langle e_i|).$$

c) For $\phi(B) = \langle \Omega, B\Omega\rangle$ we thus have

$$T(A) = \sum_{i=1}^{n} \Big((\mathbf{1}_A \otimes \langle\Omega|)\, U^\dagger\, (\mathbf{1}_A \otimes |e_i\rangle) \Big) A \Big((\mathbf{1}_A \otimes \langle e_i|)\, U\, (\mathbf{1}_A \otimes |\Omega\rangle) \Big).$$

d) So the result follows with $V_i = (\mathbf{1}_A \otimes \langle e_i|)\, U\, (\mathbf{1}_A \otimes |\Omega\rangle)$.

Problem 3.31. The following two operations $T_{a,b} \colon M_2(\mathbb{C}) \otimes \mathcal{C}_2 \to M_2(\mathbb{C})$ are distinct:

$$T_a(A_1, A_2) = V_1^\dagger A_1 V_1 + V_2^\dagger A_2 V_2, \quad T_b(A_1, A_2) = W_1^\dagger A_1 W_1 + W_2^\dagger A_2 W_2,$$

where

$$V_1^a = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, V_2^a = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad W_1^a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, W_2^a = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

However, concatenated with the inclusion $\iota \colon M_2(\mathbb{C}) \to M_2(\mathbb{C}) \otimes \mathbb{C}_2$ defined by $\iota(A) = A \otimes \mathbf{1}$ they give rise to the *same* operation $T_{a,b} \circ \iota$. Show that this is the case, and explain what this means.

3.5.4 Kadison–Schwarz inequality

For a normalized CP map $T\colon \mathcal{A} \to \mathcal{B}$, we define the sesquilinear form $\mathcal{A}\times\mathcal{A} \to \mathcal{B}$ by

$$(A,B)_T := T(A^\dagger B) - T(A)^\dagger T(B).$$

Note that this is linear in the second variable, and $(A,B)_T^\dagger = (B,A)_T$. This $\mathcal{B}$-valued form satisfies the following version of the Cauchy–Schwarz inequality.

Lemma 3.19 (Kadison–Schwarz)**.** *This is positive semidefinite,* $(A,A)_T \geq 0$ *for all* $A \in \mathcal{A}$*. If* $(A,A)_T = 0$*, then* $(A,B)_T = 0$ *for all* $B \in \mathcal{A}$*.*

Proof. Using Stinespring, we find that

$$\begin{aligned} (A,B)_T &= T(A^\dagger B) - T(A)^\dagger T(B) \\ &= V^\dagger \pi_T(A)^\dagger \pi_T(B) V - V^\dagger \pi_T(A)^\dagger V V^\dagger \pi_T(B) V \\ &= V^\dagger \pi_T(A)^\dagger (\mathbf{1} - VV^\dagger) \pi_T(B) V. \end{aligned}$$

Since $VV^\dagger\colon \mathcal{H}_T \to \mathcal{H}_T$ is the orthogonal projection onto $V(\mathcal{H}) \subseteq \mathcal{H}_T$, the operator $(\mathbf{1} - VV^\dagger)$ is the projection onto $V(\mathcal{H})^\perp$. In particular it is positive semidefinite. Setting $X_A := \sqrt{\mathbf{1} - VV^\dagger}\pi_T(A)V$ and $X_B := \sqrt{\mathbf{1} - VV^\dagger}\pi_T(B)V$, we thus obtain

$$(A,B)_T = X_A^\dagger X_B.$$

In particular, $(A,A)_T = X_A^\dagger X_A$ is a positive semidefinite element of $\mathcal{B}$, and $(A,A)_T = 0$ implies $X_A = 0$, and hence $(A,B) = X_A^\dagger X_B = 0$. □

Problem 3.32. A normalized CP map $T\colon \mathcal{A} \to \mathcal{B}$ is a $*$-homomorphism if and only if $(A,A)_T = 0$ for all $A \in \mathcal{A}$.

Problem 3.33. Let $T\colon \mathcal{A} \to \mathcal{B}$ and $S\colon \mathcal{B} \to \mathcal{A}$ be normalized CP maps.

a) Show that
$$(X,Y)_{S\circ T} = S\big((X,Y)_T\big) + \big(T(X),T(Y)\big)_S.$$

b) Using Kadison-Schwartz or otherwise, show that if S and T are each other's inverse, then both S and T are $*$-homomorphisms.

Problem 3.34. Prove the operator-valued Cauchy-Schwarz inequality

$$(A,B)_T(B,A)_T \leq \|B\|^2 (A,A)_T.$$

Hint: review the proof of the Kadison–Schwarz inquality

3.5.5 No-cloning theorem

A *cloning map* is a normalized CP map $T\colon \mathcal{A}\otimes\mathcal{A} \to \mathcal{A}$ such that

$$T(A \otimes \mathbf{1}) = A = T(\mathbf{1} \otimes A).$$

In the Schrödinger picture, this is equivalent to $T^*\rho(A\otimes\mathbf{1}) = \rho(A) = T^*\rho(\mathbf{1}\otimes A)$ for all $A \in \mathcal{A}$. In other words: the output of the state ρ on $\mathcal{A}$ is a state $T^*\rho$ on $\mathcal{A}\otimes\mathcal{A}$ that restricts to ρ on each of the two copies of $\mathcal{A}$.

Theorem 3.20 (No cloning Theorem)**.** *If a cloning map exists, then $\mathcal{A}$ is commutative.*

Proof. Since $(A \otimes \mathbf{1}, A \otimes \mathbf{1})_T = T(A^\dagger A \otimes \mathbf{1}) - T(A^\dagger \otimes \mathbf{1})T(A \otimes \mathbf{1}) = 0$ for all $A \in \mathcal{A}$, the Kadison–Schwarz inequality implies that

$$\begin{aligned} 0 &= (A^\dagger \otimes \mathbf{1}, \mathbf{1} \otimes B)_T = T(A \otimes B) - AB \text{ and} \\ 0 &= (\mathbf{1} \otimes B^\dagger, A \otimes \mathbf{1})_T = T(A \otimes B) - BA \end{aligned}$$

for all $A, B \in \mathcal{A}$. So $AB = BA$ for all $A, B \in \mathcal{A}$. □

Problem 3.35. Construct a cloning map $T \colon \mathcal{C}_n \otimes \mathcal{C}_n \to \mathcal{C}_n$.

3.6 Measures of distance

There are various ways to quantify the distance between quantum states. Here we focus on the *trace distance*, which is a generalization of the L_1-distance between probability measures.

3.6.1 The L^1-distance

Definition 3.10. The L^1-distance between two probability measures $\mathbb{P}$ and $\mathbb{P}'$ on a finite measure set Ω is defined as

$$D(\mathbb{P}, \mathbb{P}') := \sup\{|\mathbb{P}(E) - \mathbb{P}'(E)|\,;\, E \subseteq \Omega\}. \tag{74}$$

In other words, it is the worst possible difference between the probabilities $\mathbb{P}(E)$ and $\mathbb{P}'(E)$ that are assigned to a single event E. If $E^c = \Omega \setminus E$ is the complement of E, then $\mathbb{P}(E^c) - \mathbb{P}'(E^c) = -(\mathbb{P}(E) - \mathbb{P}'(E))$. In equation (74) we can therefore safely remove the absolute value:

$$D(\mathbb{P}, \mathbb{P}') = \sup\{\mathbb{P}(E) - \mathbb{P}'(E)\,;\, E \subseteq \Omega\}. \tag{75}$$

It is not hard to see that the L^1-distance is a *metric* on the set of probability distributions:

1) $D(\mathbb{P}, \mathbb{P}') = D(\mathbb{P}', \mathbb{P})$

2) $D(\mathbb{P}, \mathbb{P}') \geq 0$, and $D(\mathbb{P}, \mathbb{P}') = 0$ implies $\mathbb{P} = \mathbb{P}'$

3) $D(\mathbb{P}, \mathbb{P}'') \leq D(\mathbb{P}, \mathbb{P}') + D(\mathbb{P}', \mathbb{P}'')$

Problem 3.36. Show that the L^1-distance is a metric.

The name L^1-distance comes from the following proposition. We denote by p_ω the probability that ω occurs under the probability distribution $\mathbb{P}$.

Proposition 3.21. $D(\mathbb{P}, \mathbb{P}') = \frac{1}{2} \sum_{\omega \in \Omega} |p_\omega - p'_\omega|$.

Proof. The quantity $\mathbb{P}(E) - \mathbb{P}'(E)$ is maximal for $E_+ := \{\omega \in \Omega\,;\, p_\omega \geq p'_\omega\}$, so by (75) we have

$$D(\mathbb{P}, \mathbb{P}') = \mathbb{P}(E_+) - \mathbb{P}'(E_+) = \textstyle\sum_{\omega \in E_+} |p_\omega - p'_\omega|. \tag{76}$$

Similarly, $\mathbb{P}'(E) - \mathbb{P}(E)$ is maximal for $E_- := \{\omega \in \Omega\,;\, p_\omega < p'_\omega\}$, so

$$D(\mathbb{P}', \mathbb{P}) = \mathbb{P}'(E_-) - \mathbb{P}(E_-) = \textstyle\sum_{\omega \in E_-} |p'_\omega - p_\omega|. \tag{77}$$

Adding (76) to (77), we find

$$2D(\mathbb{P}, \mathbb{P}') \quad = \quad \textstyle\sum_{\omega \in \Omega} |p_\omega - p'_\omega|$$

as required. □

3.6.2 The trace distance

Analogous to the L^1-distance, we define the distance between two states $\rho, \sigma \in \mathcal{S}(\mathcal{A})$ as the largest possible difference between $\rho(P)$ and $\sigma(P)$, where P ranges over the orthogonal projections (events) in the $*$-algebra $\mathcal{A}$.

Definition 3.11. The trace distance between ρ and σ is defined as

$$D(\rho, \sigma) := \sup\{|\rho(P) - \sigma(P)|\,;\, P \in \mathcal{A}, P^2 = P^\dagger = P\}. \tag{78}$$

Since the complementary projection $P^c := \mathbf{1} - P$ gives $\rho(P^c) - \sigma(P^c) = -(\rho(P) - \sigma(P))$, we can omit the absolute value signs in (78),

$$D(\rho, \sigma) := \sup\{\rho(P) - \sigma(P)\,;\, P \in \mathcal{A}, P^2 = P^\dagger = P\}. \tag{79}$$

The trace distance is a metric on the state space $\mathcal{S}(\mathcal{A})$. For all $\rho, \sigma, \tau \in \mathcal{S}(\mathcal{A})$ we have

1) $D(\rho, \sigma) = D(\sigma, \rho)$
2) $D(\rho, \sigma) \geq 0$, and $D(\rho, \sigma) = 0$ implies $\rho = \sigma$
3) $D(\rho, \tau) \leq D(\rho, \sigma) + D(\sigma, \tau)$.

Problem 3.37. Show that the trace distance is a metric on $\mathcal{S}(\mathcal{A})$.

The name 'trace distance' comes from the following analogue of Proposition 3.21. Recall that for a Hermitian operator $A = \sum_{a \in \mathrm{spec}(A)} aP_a$, the absolute value is the nonnegative operator $|A| = \sum_{a \in \mathrm{spec}(A)} |a| P_a$.

Proposition 3.22. *Let R and S be the density matrices of ρ and σ. Then*

$$D(\rho, \sigma) = \frac{1}{2}\mathbf{tr}|R - S|$$

The proof requires the following lemma.

Lemma 3.23. *If A and B are positive semidefinite, then* $\mathbf{tr}(AB) \geq 0$.

Proof. Since $B \geq 0$, we have $B = B^{1/2}B^{1/2}$. So $\mathbf{tr}(AB) = \mathbf{tr}(B^{1/2}AB^{1/2})$, which is nonnegative by Lemma 3.15. □

Proof of Proposition 3.22. If $R - S = \sum_{a \in \mathrm{spec}(R-S)} aP_a$ is the spectral decomposition of the Hermitian operator $R - S$, then we set $P_+ := \sum_{a \geq 0} P_a$, and $P_- := \mathbf{1} - P_+$.

We show that $\rho(P_+) - \sigma(P_+)$ is the largest difference in probability that can occur. Let $P \in \mathcal{A}$ be any other orthogonal projection. Then

$$\begin{aligned} \rho(P) - \sigma(P) &= \mathbf{tr}(P(R-S)) \\ &\leq \mathbf{tr}(P(R-S)P_+) \\ &\leq \mathbf{tr}((R-S)P_+) = \rho(P_+) - \sigma(P_+). \end{aligned}$$

Here both inequalities are due to Lemma 3.23. For the first one,

$$\mathbf{tr}(\underbrace{P}_{\geq 0}\underbrace{(R-S)(P_+ - \mathbf{1})}_{\geq 0}) \geq 0$$

because both P and $(R-S)(P_+ - \mathbf{1})$ are positive semidefinite. For the second one,

$$\mathbf{tr}(\underbrace{\mathbf{1} - P}_{\geq 0}\underbrace{(R-S)P_+}_{\geq 0}) \geq 0$$

because both $\mathbf{1} - P$ and $(R-S)P_+$ are positive semidefinite.

We conclude that $D(\rho, \sigma) = \mathbf{tr}(P_+(R-S))$. Changing the roles of ρ and σ, we similarly find $D(\sigma, \rho) = \mathbf{tr}(P_-(S-R))$. Adding these two equations, we find

$$2D(\rho, \sigma) = \mathbf{tr}((P_+ - P_-)(R-S)) = \mathbf{tr}|R-S|$$

as required. □

Note that the above argument for $\rho(P) - \sigma(P) \leq \rho(P_+) - \sigma(P_+)$ only uses that $P \geq 0$ and $\mathbf{1} - P \geq 0$. As a corollary of the proof, we therefore find the following characterization of the trace distance:

$$D(\rho, \sigma) = \sup\{\rho(A) - \sigma(A)\,;\, A \in \mathcal{A}, 0 \leq A \leq \mathbf{1}\}. \tag{80}$$

Problem 3.38. If two qubit states ρ and σ have density matrices $R = \frac{1}{2}(\mathbf{1} + \vec{r} \cdot \vec{\sigma})$ and $S = \frac{1}{2}(\mathbf{1} + \vec{s} \cdot \vec{\sigma})$, then $D(\rho, \sigma) = \frac{1}{2}\|\vec{r} - \vec{s}\|$.

3.6.3 Operations decrease trace distance

The following theorem shows that operations always decrease the trace distance between states.

Theorem 3.24. *Let $T\colon \mathcal{B} \to \mathcal{A}$ be a normalized completely positive map. Then*

$$D(T^*\rho, T^*\sigma) \leq D(\rho, \sigma)$$

for all $\rho, \sigma \in \mathcal{S}(\mathcal{A})$.

Proof. Since $T(\mathbf{1}) = \mathbf{1}$, the image $T(B) \in \mathcal{A}$ of any $B \in \mathcal{B}$ with $0 \leq B \leq \mathbf{1}$ is also between 0 and $\mathbf{1}$. Using (80), we thus find

$$\begin{aligned} D(T^*\rho, T^*\sigma) &= \sup\{T^*\rho(B) - T^*\sigma(B)\,;\, B \in \mathcal{B}, 0 \leq B \leq \mathbf{1}\} \\ &= \sup\{\rho(T(B)) - \sigma(T(B))\,;\, B \in \mathcal{B}, 0 \leq B \leq \mathbf{1}\} \\ &\leq \sup\{\rho(A) - \sigma(A)\,;\, A \in \mathcal{A}, 0 \leq A \leq \mathbf{1}\} = D(\rho, \sigma). \end{aligned}$$

□

In particular this holds for operations $T\colon \mathcal{C}_d \to \mathcal{L}(\mathcal{H})$, which correspond to POVMs as we have seen in §3.4.1. For any state $\rho \in \mathcal{S}(\mathcal{L}(\mathcal{H}))$, the image $T^*\rho$ is a state on $\mathcal{C}_d$, which we identify with a probability measure $\mathbb{P}^T_\rho$ on the probability space $\Omega = \{\omega_1, \ldots, \omega_d\}$ with d outcomes.

In this setting, Theorem 3.24 tells us that the (classical) L^1-distance between the probability measures $\mathbb{P}^T_\rho$ and $\mathbb{P}^T_\sigma$ is bounded by the (quantum) trace distance between ρ and σ,

$$D(\mathbb{P}^T_\rho, \mathbb{P}^T_\sigma) \leq D(\rho, \sigma). \tag{81}$$

This bound turns out to be sharp:

Theorem 3.25. *Let $\rho, \sigma \in \mathcal{S}(\mathcal{L}(\mathcal{H}))$. Then*

$$D(\rho, \sigma) = \sup_{T\colon \mathcal{C} \to \mathcal{L}(\mathcal{H})} D(\mathbb{P}^T_\rho, \mathbb{P}^T_\sigma),$$

where the supremum runs over all commutative $*$-algebras $\mathcal{C}$, and all normalized CP maps from $\mathcal{C}$ to $\mathcal{L}(\mathcal{H})$.

Proof. By (81), it suffices to find an operation $T\colon \mathcal{C} \to \mathcal{L}(\mathcal{H})$ with $D(T^*\rho, T^*\sigma) = D(\rho, \sigma)$. If R and S are the density matrices for ρ and σ, then we take $\mathcal{C} \subseteq \mathcal{L}(\mathcal{H})$ to be the commutative $*$-algebra spanned by the spectral projections of $R - S$, and $T\colon \mathcal{C} \hookrightarrow \mathcal{L}(\mathcal{H})$ the inclusion. One checks that for this choice of $\mathcal{C}$ and T, equality holds as required. □

Problem 3.39. Verify that this is indeed the case.

3.6.4 Norm distance between operations

Quantum gates that are realized in a laboratory will always be imperfect to some degree. In order to quantify the degree of imperfection, and to analyse the effect this has on – say – executing a quantum algorithm, we introduce the *norm distance* between two normalized completely positive maps $T, S\colon \mathcal{B} \to \mathcal{A}$.

Definition 3.12.

$$D(T,S) := \sup_{\rho \in \mathcal{S}(\mathcal{A})} D(T^*\rho, S^*\rho). \tag{82}$$

If S is the gate one tries to model, and T is the gate one actually manages to construct in a laboratory, then $D(T,S)$ is the worst case difference in output measured with respect to the trace distance.

The name 'norm distance' is explained by the following result.

Proposition 3.26.

$$D(T,S) = \tfrac{1}{2}\sup\{\|T(B) - S(B)\| \,;\, B \in \mathcal{B}, B^\dagger = B, \|B\| \leq 1\}.$$

Lemma 3.27. *If $A \in \mathcal{A}$ is Hermitian, then $\|A\| = \sup\{\rho(A)\,;\, \rho \in \mathcal{S}(\mathcal{A})\}$.*

Proof. Since $A \leq \|A\|\mathbf{1}$, we have $\rho(A) \leq \|A\|$. Similarly $-\|A\|\mathbf{1} \leq A$ implies $-\|A\| \leq \rho(A)$, so $|\rho(A)| \leq \|A\|$. It remains to find a state with $|\rho(A)| = \|A\|$. Let $a_{\max} \in \mathrm{spec}(A)$ be an eigenvalue for which absolute value $|a_{\max}|$ is maximal. Then $|a_{\max}| = \|A\|$, and a corresponding unit eigenvector $\psi_{a_{\max}}$ yields a state ρ with the desired properties, $|\rho(A)| = |\langle \psi_{\max}, A\psi_{\max}\rangle| = |a_{\max}|$. □

Proof of Proposition 3.26. Since

$$\begin{aligned} D(T,S) &= \sup_{\rho \in \mathcal{S}(\mathcal{A})} D(T^*\rho, S^*\rho) \\ &= \sup_{\rho \in \mathcal{S}(\mathcal{A})} \sup_{0 \leq B \leq \mathbf{1}} |T^*\rho(B) - S^*\rho(B)| \\ &= \sup_{0 \leq B \leq \mathbf{1}} \Big(\sup_{\rho \in \mathcal{S}(\mathcal{A})} |\rho(T(B) - S(B))| \Big), \end{aligned}$$

Lemma 3.27 allows us to eliminate the supremum over ρ, yielding

$$D(T,S) = \sup_{0 \leq B \leq \mathbf{1}} \|T(B) - S(B)\|.$$

The positive semidefinite operator $0 \leq B \leq \mathbf{1}$ gives rise to the Hermitian operator $B' := 2B - \mathbf{1}$ with $\|B'\| \leq 1$, and every such Hermitian operator is of this form. Since $T(B') - S(B') = 2(T(B) - S(B))$, we find

$$D(T,S) = \frac{1}{2} \sup_{\|B'\| \leq 1} \|T(B') - S(B')\|$$

as required, with the supremum over the Hermitian elements of $\mathcal{B}$ of norm at most one. □

Problem 3.40. A trace-preserving CP map $T^* \colon \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{A})$ is a *contraction* if there exists a number $0 \leq c < 1$ such that $D(T^*\rho, T^*\sigma) < cD(\rho, \sigma)$.

a) Use the Banach fixed point theorem to conclude that every contraction has a unique fixed state $\rho_0 \in \mathcal{S}(\mathcal{A})$ with $T^*\rho_0 = \rho_0$.

b) Starting from any state ρ, repeated application of T^* yields the same limit $\lim_{n\to\infty}(T^*)^n\rho = \rho_0$.

c) The *depolarizing channel* on $M_2(\mathbb{C})$ is defined by

$$T(A) = \tfrac{p}{2}\mathbf{tr}(A)\mathbf{1} + (1-p)A.$$

(Here p is the probability of depolarization.) Show that T^* is contractive, determine the unique fixed state, and explain what the result of repeated depolarization is.

Alternative to the trace distance, another commonly used way to quantify the difference between states ρ and σ with density matrices R and S is the *fidelity*

$$F(\rho,\sigma) := \mathbf{tr}\sqrt{R^{1/2}SR^{1/2}}.$$

This is not quite a metric, since the fidelity is 1 if $R = S$. Rather, one thinks of states with high fidelity (close to one) as similar, whereas a low fidelity (close to zero) means that the states are very different. The fidelity has a number of properties which makes it easier to use than trace distance in some respects. Although fidelity lacks the clear physical motivation of the trace distance, the two quantities are related by the inequalities

$$1 - F(R,S) \leq D(\rho,\sigma) \leq \sqrt{1 - F(R,S)^2}, \tag{83}$$

see e.g. [NC00, §9.2.3].

Problem 3.41. Show that for pure states $R = |\psi\rangle\langle\psi|$ and $S = |\chi\rangle\langle\chi|$, the fidelity is $F(\rho,\sigma) = |\langle\psi,\chi\rangle|$.

3.7 Semidefinite optimization and state discrimination

In semidefinite programming (SDP), one considers a pair of optimization problems parameterized by the following data: a Hermitian operator $A \in \mathcal{A}$, a Hermitian operator $B \in \mathcal{B}$, and a complex linear map $\Phi\colon \mathcal{A} \to \mathcal{B}$ that preserves adjoints, $\Phi(A^\dagger) = \Phi(A)^\dagger$. We denote by $\Phi^*\colon \mathcal{B} \to \mathcal{A}$ is the dual map with respect to the trace pairing, so

$$\mathbf{tr}(\Phi(X)Y) = \mathbf{tr}(X\Phi^*(Y))$$

for all $X \in \mathcal{A}$ and $Y \in \mathcal{B}$.

Primal problem	**Dual problem**
Maximize $\mathbf{tr}(AX)$	Minimize $\mathbf{tr}(BY)$
over $X \geq 0$ in $\mathcal{A}$	over Hermitian Y in $\mathcal{B}$
with $\Phi(X) = B$.	with $\Phi^*(Y) \geq A$.
The supremum is α^*.	The infimum is β^*.

The primal problem is called *strictly feasible* if there exists at least one $X \geq 0$ with $\Phi(X) = B$. Similarly, the dual problem is strictly feasible if there exists at least one Hermitian $Y \in \mathcal{B}$ with $\Phi^*(Y) \geq A$. The primal and the dual problem are related by *weak duality.*

Proposition 3.28 (Weak duality)**.** *We have $\alpha^* \leq \beta^*$.*

Proof. Since $X \geq 0$ and $\Phi^*(Y) - A \geq 0$, we have

$$\mathbf{tr}(AX) \leq \mathbf{tr}(\Phi^*(Y)X) = \mathbf{tr}(Y\Phi(X)) = \mathbf{tr}(YB), \tag{84}$$

where the first inequality is due to Lemma 3.23. Since this holds for all admissible $X \in \mathcal{A}$ and $Y \in \mathcal{B}$, we conclude that $\alpha^* \leq \beta^*$. □

3.7.1 State discrimination (weak form)

Many problems in quantum information theory can be formulated in terms of SDPs. A case in point is the problem of *state discrimination.*

Alice has a collection of n qubits in (possibly) different states ρ_ω with density matrix R_ω, labelled by a set Ω with cardinality n. With probability p_ω, she chooses the state ρ_ω and sends it to Bob. Bob knows the set $\{\rho_\omega\,;\,\omega \in \Omega\}$ from which Alice can choose, and he knows the probabilities p_ω of the various choices. However, he does *not* know the particular choice ω_0 that Alice made. The objective for Bob is to determine ω_0 as well as possible by performing a POVM on the state ρ_{ω_0} that he received.

In more detail, Bob performs a POVM $\{E_\omega\,;\,\omega \in \Omega\}$ on the state ρ_{ω_0}, and the outcome ω is Bob's guess of the state that Alice sent. The overall probability that Bob gets it right is $\sum_{\omega\in\Omega} p_\omega \mathbf{tr}(R_\omega E_\omega)$, so Bob's objective is to find a POVM which maximizes this expression.

Theorem 3.29 (State discrimination)**.** *The* POVM $\{E_\omega\,;\,\omega \in \Omega\}$ *is optimal if and only if the operator $Y = \sum_{\omega\in\Omega} p_\omega E_\omega R_\omega$ is Hermitian and satisfies $Y \geq p_\omega R_\omega$ for all $\omega \in \Omega$.*

Proof of Theorem 3.29, 'if' direction. The idea is to formulate this as a primal problem for the algebra $\mathcal{A} = \mathcal{C}(\Omega, M_2(\mathbb{C}))$ of $M_2(\mathbb{C})$-valued functions on Ω, and for $\mathcal{B} = M_2(\mathbb{C})$. We choose $A \in \mathcal{A}$ to be the function $A_\omega = p_\omega R_\omega$, we choose $B = \mathbf{1}$, and set $\Phi\colon \mathcal{A} \to \mathcal{B}$ to be the map $\Phi(X) = \sum_{\omega\in\Omega} X_\omega$.

The primal problem is then to maximize $\mathbf{tr}(AX) = \sum_{\omega\in\Omega} p_\omega \mathbf{tr}(R_\omega X_\omega)$ over all $X_\omega \geq 0$ that satisfy $\sum_{\omega\in\Omega} X_\omega = \mathbf{1}$. Setting $E_\omega = X_\omega$, we see that a solution to the primal problem is the same thing as an optimal POVM.

The dual of Φ is the map $\Phi^*\colon \mathcal{B} \to \mathcal{A}$ with $\Phi^*(Y)_\omega = Y$. In other words, $\Phi^*(Y)\colon \Omega \to M_2(\mathbb{C})$ is the constant map that takes the value $Y \in M_2(\mathbb{C})$ on every $\omega \in \Omega$. The dual problem, then, is to minimize $\mathbf{tr}(Y)$ under the requirement that Y is Hermitian, and that $Y \geq p_\omega R_\omega$ for all $\omega \in \Omega$.

Suppose that $Y = \sum_{\omega\in\Omega} p_\omega E_\omega R_\omega$ is Hermitian with $Y \geq p_\omega R_\omega$ for all $\omega \in \Omega$. Then Y is a feasible solution for the dual problem, so $\mathbf{tr}(Y) \geq \beta^* \geq \alpha^*$. But since $\mathbf{tr}(Y) = \sum_{\omega\in\Omega} p_\omega E_\omega \mathbf{tr}(R_\omega)$ is the value of $X_\omega = E_\omega$ in the *primal* problem, we also have $\mathbf{tr}(Y) = \mathbf{tr}(AX) \leq \alpha^*$. We conclude that $\alpha^* = \beta^*$, that $Y = Y^*$ is an optimal solution to the dual problem, and that $X^*_\omega = E_\omega$ is an optimal solution to the primary problem. □

3.7.2 Strong duality and state discrimination

The following key result in semidefinite optimization ensures that $\alpha^* = \beta^*$ in many interesting cases.

Theorem 3.30 (Strong Duality). *If the dual problem is strictly feasible and its values are bounded from below, then $\alpha^* = \beta^*$, and there exists a primal feasible solution that attains the maximum. Conversely, if the the primal problem is strictly feasible and its values are bounded from above, then $\alpha^* = \beta^*$ and there exists a dual feasible solution that attains the minimum.*

Unfortunately it would take us too far afield to prove this theorem, so we refer to e.g. [BV04] or [LV23, §2.4]. Together with the following extension of Lemma 3.23, Strong duality is the main ingredient for the 'only if' part of Theorem 3.29.

Lemma 3.31. *Let $A, B \geq 0$. Then $\mathbf{tr}(AB) \geq 0$, and $\mathbf{tr}(AB) = 0$ if and only if $AB = 0$.*

Proof. Since $\mathbf{tr}(AB) = \mathbf{tr}(B^{1/2}AB^{1/2})$ and $B^{1/2}AB^{1/2}$ is positive semidefinite, we have $\mathbf{tr}(AB) \geq 0$. If $\mathbf{tr}(AB) = 0$, then $B^{1/2}AB^{1/2} = 0$ because it is positive semidefinite with zero trace. If B is invertible, then right and left multiplication by $B^{-1/2}$ yields $A = 0$, so we are done.

If B is not invertible, we multiply from the right and from the left by $f(B)$, where the function $f\colon \operatorname{spec}(B) \to \mathbb{R}$ is defined by $f(x) = x^{-1/2}$ for $x > 0$ and $f(x) = 1$ for $x = 0$. Since $B^{1/2}f(B)$ is the projection P_B onto the range of B, we find $P_BAP_B = 0$. Let $P_B^\perp$ be the projection onto the kernel of B. Since $A \geq 0$, we have $P_BAP_B^\perp = P_B^\perp AP_B = 0$, so $A = P_B^\perp AP_B^\perp$ and $AB = P_B^\perp AP_B^\perp B = 0$. □

Problem 3.42. If $A \geq 0$ and $PAP = 0$ for an orthogonal projection P, then $P^\perp AP = PAP^\perp = 0$.

Proof of Theorem 3.29, 'only if' direction. Since the dual problem is bounded from below ($\mathbf{tr}(Y) \geq 0$ because $Y \geq 0$) and strictly feasible ($Y = \mathbf{1}$ would be an example of a non-optimal solution), strong duality tells us that $\alpha^* = \beta^*$, and that the primal problem has an optimal solution X^*. Similarly, the primal problem is bounded from above (by 1) and strictly feasible (because POVMs exist), so the dual problem has an optimal solution Y^* as well.

Since $\alpha^* = \beta^*$, the inequality

$$\alpha^* = \mathbf{tr}(X^*A) \leq \mathbf{tr}(X^*\Phi^*(Y^*)) = \mathbf{tr}(\Phi(X^*)Y^*) = \mathbf{tr}(Y^*) = \beta^*$$

from (84) is actually an equality, and

$$\mathbf{tr}\big(X^*(\Phi^*(Y^*) - A)\big) = 0.$$

Since $X^* \geq 0$ and $(\Phi^*(Y^*)-A) \geq 0$, Lemma 3.31 implies that $X^*(\Phi^*(Y^*)-A) = 0$. It follows that $X^*\Phi^*(Y^*) = X^*A$, so that

$$X^*_\omega Y^* = p_\omega X^*_\omega R_\omega$$

for all $\omega \in \Omega$. Summing over ω and using $\sum_{\omega\in\Omega} X^*_\omega = \mathbf{1}$, we find that the optimal dual solution is

$$Y^* = \sum_{\omega\in\Omega} p_\omega X^*_\omega R_\omega.$$

So for every optimal POVM $\{E_\omega\,;\,\omega \in \Omega\}$, the operator $Y = \sum_{\omega\in\Omega} p_\omega E_\omega R_\omega$ is Hermitian, and $Y \geq p_\omega R_\omega$. □

Problem 3.43. Formulate weak and strong duality in the special case that the algebras $\mathcal{A} = \mathcal{C}_n$ and $\mathcal{B} = \mathcal{C}_m$ are commutative.

a) Characterize hermiticity and positivity of $X \in \mathcal{C}_n$ in terms of the matrix entries.

b) Characterize the linear maps $\Phi\colon \mathcal{C}_n \to \mathcal{C}_m$ that preserve Hermiticity.

c) Describe the adjoint $\Phi^*\colon \mathcal{C}_m \to \mathcal{C}_n$ with respect to the trace pairing. When is $\Phi^*(Y) \geq A$ in $\mathcal{C}_n$?

d) Show that weak and strong duality for Semidefinite Programming (SDP) reduces to weak and strong duality for Linear Programming (LP) in the special case $\mathcal{A} = \mathcal{C}_n$, $\mathcal{B} = \mathcal{C}_m$. (If you do not know what a linear program is, then just simplify the setting as much as possible and you will discover its definition.)

Problem 3.44. Alice sends Bob a qubit in a state with density matrix either R_0 (with probability p_0) or R_1 (with probability p_1). In order to find out what Alice sent him, Bob performs the POVM with operators $E_0 \geq 0$, $E_1 \geq 0$ that satisfy $E_0 + E_1 = \mathbf{1}$.

a) Let $\Delta := p_0R_0 - p_1R_1$. The POVM is optimal if and only if $E_0\Delta \geq 0$ and $E_1\Delta \leq 0$.

b) This is satisfied by $E_0 = \sum_{\delta\geq 0} P_\delta$ and $E_1 = \sum_{\delta<0} P_\delta$, where P_δ is the projection onto the eigenspace of Δ with eigenvalue $\delta \in \operatorname{spec}(\Delta)$.

c) Under which conditions on p_0, p_1, R_0 and R_1 is the POVM unique?

d) Suppose that $R_0 = |\psi_0\rangle\langle\psi_0|$ and $R_1 = |\psi_1\rangle\langle\psi_1|$ for orthonormal vectors $\psi_0, \psi_1 \in \mathbb{C}^2$. Do you recover the POVM you expected?

Unambiguous state discrimination In the problem of quantum state discrimination, we have taken the overall probability of finding the right answer as our measure of success. But there are other possibilities. In the following problem, we will consider Bob's answer to be a success if it is *cerrtifyably* correct, meaning that Bob is not only right, but that he *knows* that he is right.

Problem 3.45. Alice sends Bob a qubit which is either in state ψ_0 (with probability p_0) or ψ_1 (with probability p_1). We will assume that ψ_0 and ψ_1 are linearly independent. We construct a POVM on $\Omega = \{0, 1, 2\}$ that admits the following interpretation. If Bob obtains the outcome 0, then he is certain that Alice sent him the state ψ_0. If he obtains the outcome 1, he is certain that Alice sent him the state ψ_1. And if he finds the outcome 2, then the test is inconclusive.

a) Let ψ'_0 and ψ'_1 be (unnormalized) vectors such that $\langle \psi'_i, \psi_j \rangle = \delta_{ij}$, and let

$$P'_0 = \frac{|\psi'_0\rangle \langle \psi'_0|}{\langle \psi'_0, \psi'_0 \rangle}, \quad P'_1 = \frac{|\psi'_1\rangle \langle \psi'_1|}{\langle \psi'_1, \psi'_1 \rangle}$$

be the corresponding projectors. If $\mu_0 \in [0, 1]$ and $\mu_1 \in [0, 1]$ are such that $\mu_0 P'_0 + \mu_1 P'_1 \leq \mathbf{1}$, then the POVM on $\Omega = \{0, 1, 2\}$ with $E_0 = \mu_0 P'_0$, $E_1 = \mu_1 P'_1$, and $E_2 = \mathbf{1} - \mu_0 P'_0 - \mu_1 P'_1$ has the above interpretation: an input state ψ_0 yields outcome 0 or 2, and an input state ψ_1 yields outcome 1 or 2.

b) Conversely, every POVM with this property is of this form.

c) The probability that Bob obtains certainty about the qubit that Alice sent him is

$$P_{\text{Test}} = \frac{p_0 \mu_0}{\langle \psi'_0, \psi'_0 \rangle} + \frac{p_1 \mu_1}{\langle \psi'_1, \psi'_1 \rangle}.$$

c) By a change of coordinates, we may assume that $\psi_0 = |0\rangle$ and $\psi_1 = \alpha |0\rangle + \beta |1\rangle$ for $\alpha, \beta \in \mathbb{C}$ with $|\alpha|^2 + |\beta|^2 = 1$. Show that $\mu_0 P'_0 + \mu_1 P'_1 \leq \mathbf{1}$ if and only if $0 \leq 1 - \mu_0 - \mu_1 + |\beta|^2 \mu_0 \mu_1$.

d) Conclude that $P_{\text{Test}}(\mu_0, \mu_1) = |\beta^2|(p_0 \mu_0 + p_1 \mu_1)$.

e) The probability $P_{\text{Test}}(\mu_0, \mu_1)$ is optimal on the hyperbola $f(\mu_0, \mu_1) := 1 - \mu_0 - \mu_1 + |\beta|^2 \mu_0 \mu_1 = 0$. Conclude that if $P_{\text{Test}}(\mu_0, \mu_1)$ is maximal, then $p_1(\mu_1 - 1/|\beta|^2) = p_0(\mu_0 - 1/|\beta|^2)$.

f) For $p_0 = p_1 = 1/2$, the optimal value of P_{Test} is $P^*_{\text{Test}} = 1 - |\alpha|$.

g) Discuss the limit $|\alpha| \downarrow 0$ and $|\alpha| \uparrow 1$. Is this what you expected?

4 Problems related to quantum circuits

4.1 Superdense coding

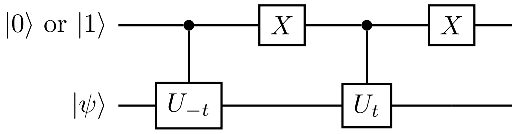
Problem 4.1 (Superdense coding)**.** Show that for any observable A on $\mathcal{H} = \mathbb{C}^2$, the expectation $\langle \psi, A \otimes \mathbf{1}\psi \rangle$ takes the *same* value in each of the four Bell states ψ. Suppose that Alice communicates two classical bits of information to Bob using super dense coding. If an eavesdropper ('Eve') intercepts the qubit that Alice sends to Bob, can she infer which of the bit strings 00, 01, 10, 11 was sent by Alice?

4.2 Circuits

Problem 4.2. Recall that the Pauli matrices are defined by

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

a) Let $U_t = \exp(-itH)$ for a single-qubit Hamiltonian $H \in M_2(\mathbb{C})$. Determine the output of the following diagram for the two inputs $|0\rangle \otimes |\psi\rangle$ and $|1\rangle \otimes |\psi\rangle$.

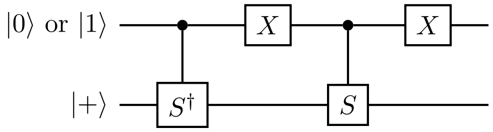


b) Show that the above diagram implements the two-qubit unitary operator $V_t = \exp(-itZ \otimes H)$.

c) Determine $U_t = \exp(-itH)$ for $H = Y$. Determine a time t_0 such that $U_{t_0} = S$, where S is the unitary operator

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

d) Determine the output of the following diagram with input $|0\rangle \otimes |+\rangle$ and $|1\rangle \otimes |+\rangle$, where $|+\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.



e) Let t_0 be the time from c), and let $\psi = \alpha |0\rangle + \beta |1\rangle$ be an arbitrary unit vector. Determine the probability that the event $\mathbf{1} \otimes |0\rangle\langle 0|$ occurs in the state $\Psi = \exp(-it_0 Z \otimes Y) |\psi\rangle \otimes |+\rangle$.

Problem 4.3. In the following problem, $\mathcal{H} = (\mathbb{C}^2)^{\otimes k}$ is an auxilliary register containing k qubits, initialized in $|0\rangle = |0 \dots 0\rangle$. The number k is as big as you like.

a) Construct a unitary gate for adding 2 bits that uses only 2-qubit gates. That is, construct the unitary operator U on

$$\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes (\mathbb{C}^2)^{\otimes 2} \otimes \mathcal{H}$$

that maps $|a\rangle \otimes |b\rangle \otimes |00\rangle \otimes |0\rangle$ to $|a\rangle \otimes |b\rangle \otimes |a+b\rangle \otimes |0\rangle$.
Hint: recall that the Toffoli gate counts for 5 two-qubit gates.

b) Construct a unitary gate for adding 2 two-bit numbers that uses only 2-qubit gates. That is, construct the the unitary operator U on

$$(\mathbb{C}^2)^{\otimes 2} \otimes (\mathbb{C}^2)^{\otimes 2} \otimes (\mathbb{C}^2)^{\otimes 3} \otimes \mathcal{H}$$

that maps $|a\rangle \otimes |b\rangle \otimes |000\rangle \otimes |0\rangle$ to $|a\rangle \otimes |b\rangle \otimes |a+b\rangle \otimes |0\rangle$. (Here $a = a_1 a_2$ and $b = b_1 b_2$ are two-bit numbers.)

c) Estimate how many 2-qubit gates are necessary in order to implement addition on L-bit numbers.

d) Construct a unitary gate for multiplying 2 bits.

e) Estimate the number of 2-qubit operations needed to multiply two L-bit numbers.

f) Are your solutions to c) and e) polynomial in L?

4.3 Error correction

The *Shor code* is the two-dimensional linear subspace $\mathcal{H}_L \subseteq \mathcal{H}_P = (\mathbb{C}^2)^{\otimes 9}$ spanned by

$$\begin{aligned} |0_L\rangle &= 2^{-3/2}(|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \\ |1_L\rangle &= 2^{-3/2}(|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle) \end{aligned}$$

We think of $(\mathbb{C}^2)^{\otimes 9}$ as the Hilbert space that describes 3 'blocks' of 3 physical qubits. We denote by A_i the operator $A_i := \mathbf{1} \otimes \dots \otimes \mathbf{1} \otimes A \otimes \mathbf{1} \otimes \dots \otimes \mathbf{1}$, with $A \in M_2(\mathbb{C})$ at position i.

To protect against *bit* flips $|\psi\rangle \mapsto X_i |\psi\rangle$, one measures $Z_1 Z_2$ and $Z_2 Z_3$ in each of the three blocks of 3 qubits:

$$Z_1 Z_2, Z_2 Z_3; Z_4 Z_5, Z_5 Z_6; Z_7 Z_8, Z_8 Z_9.$$

The syndrome for this consists of 6 signs ±1, corresponding to the outcome of measuring each of these 6 observables.

Problem 4.4. What is the syndrome that results from a single bit flip in position number 5? Show that recovery can be accomplished by applying X_5.

To protect against *phase* flips $|\psi\rangle \mapsto Z_i |\psi\rangle$, one measures the two observables

$$X_1X_2X_3X_4X_5X_6, X_4X_5X_6X_7X_8X_9.$$

The syndrome for this consists of two signs ± 1, one for each observable.

Problem 4.5. What is the syndrome that results from a single phase flip in position number 5? Show that recovery can be accomplished by applying $Z_4Z_5Z_6$.

Problem 4.6. Show that the $6 + 2 = 8$ observables mentioned above commute amongst each other, and act trivially on the logical states $|0_L\rangle$ and $|1_L\rangle$.

References

[Co07] John B. Conway, *A course in functional analysis*, Springer New York, NY 2nd edition, 2007

[JM06] Bas Janssens and Hans Maassen, *Information Transfer Implies State Collapse*, J. Phys. A: Math. Gen. **39** (2006), 9845–9860.

[M04] Hans Maassen, *Quantum Probability, Quantum Information Theory, Quantum Computing*, Katholieke Universiteit Nijmegen, 2004

[NC00] Michael A. Nielsen and Isaac L. Chuang, *Quantum computation and quantum information*, Cambridge University Press, 10th edition, 2010.

[NP22] Scientific background on the Nobel Prize in physics 2022, the Nobel committee for Physics, 2022.
www.nobelprize.org/uploads/2022/10/advanced-physicsprize2022.pdf

[LV23] Monique Laurent and Frank Vallentin. *A course on semidefinite optimization*, Draft lecture notes, January 4, 2023.

[BV04] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[N22] Jan van Neerven, *Functional Analysis*, Cambridge University Press, 2022.