Lecture Notes for Ph219/CS219: Quantum Information Chapter 3

John Preskill
California Institute of Technology

Updated October 2018

PER IL CORSO di INTRODUZIONE ALLA SCIENZA E TECNOLOGIA QUANTISTICA A.A. 2023-24
E' RILEVANTE LA SEZIONE 3.4

Contents

3
Foundations II: Measurement and Evolution

3.1 Orthogonal measurement and beyond

In Chapter 2 we discussed how to describe the state of an *open* quantum system, one which is part of a larger system. In this Chapter we will extend the theory of open quantum systems further. In particular, we will develop two important concepts: *generalized measurements*, which are performed by making use of an auxiliary system, and *quantum channels*, which describe how open systems evolve.

3.1.1 Orthogonal Measurements

An axiom of quantum theory asserts that a measurement may be described as an orthogonal projection operator. But if a measurement of system S is realized by performing an orthogonal measurement on a larger system that contains S, the resulting operation performed on S alone need not be an orthogonal projection. We would like to find a mathematical description of such "generalized measurements" on system S. But first let's recall how measurement of an arbitrary Hermitian operator can be achieved in principle, following the classic treatment of Von Neumann.

To measure an observable $\boldsymbol{M}$, we will modify the Hamiltonian of the world by turning on a coupling between that observable and another variable that represents the apparatus. Depending on the context, we will refer to the auxiliary system used in the measurement as the "pointer," the "meter," or the "ancilla." (The word "ancilla" just means something extra which is used to achieve a desired goal.) The coupling establishes a correlation between the eigenstates of the observable and the distinguishable states of the pointer, so that we can prepare an eigenstate of the observable by "observing" the pointer.

This may not seem like a fully satisfying model of measurement because

we have not explained how to measure the pointer. Von Neumann's attitude was that it is possible in principle to correlate the state of a microscopic quantum system with the value of a macroscopic classical variable, and we may take it for granted that we can perceive the value of the classical variable. A quantum measurement, then, is a procedure for amplifying a property of the microscopic world, making it manifest in the macroscopic world.

We may think of the pointer as a particle of mass m that propagates freely apart from its tunable coupling to the quantum system being measured. Since we intend to measure the position of the pointer, it should be prepared initially in a wavepacket state that is narrow in position space — but not too narrow, because a vary narrow wave packet will spread too rapidly. If the initial width of the wave packet is Δx, then the uncertainty in it velocity will be of order $\Delta v = \Delta p/m \sim \hbar/m\Delta x$, so that after a time t, the wavepacket will spread to a width

$$\Delta x(t) \sim \Delta x + \frac{\hbar t}{\Delta x}, \tag{3.1}$$

which is minimized for $(\Delta x(t))^2 \sim (\Delta x)^2 \sim \hbar t/m$. Therefore, if the experiment takes a time t, the resolution we can achieve for the final position of the pointer is limited by

$$\Delta x \gtrsim (\Delta x)_{SQL} \sim \sqrt{\frac{\hbar t}{m}}, \tag{3.2}$$

the "standard quantum limit." We will choose our pointer to be sufficiently heavy that this limitation is not serious.

The Hamiltonian describing the coupling of the quantum system to the pointer has the form

$$\boldsymbol{H} = \boldsymbol{H}_0 + \frac{1}{2m}\boldsymbol{P}^2 + \lambda(t)\boldsymbol{M} \otimes \boldsymbol{P}, \tag{3.3}$$

where $\boldsymbol{P}^2/2m$ is the Hamiltonian of the free pointer particle (which we will henceforth ignore on the grounds that the pointer is so heavy that spreading of its wavepacket may be neglected), $\boldsymbol{H}_0$ is the unperturbed Hamiltonian of the system to be measured, and λ is a coupling constant that we are able to turn on and off as desired. The observable to be measured, $\boldsymbol{M}$, is coupled to the momentum $\boldsymbol{P}$ of the pointer.

If $\boldsymbol{M}$ does not commute with $\boldsymbol{H}_0$, then we have to worry about how the observable $\boldsymbol{M}$ evolves during the course of the measurement. To simplify the analysis, let us suppose that either $[\boldsymbol{M}, \boldsymbol{H}_0] = 0$, or else the measurement is carried out quickly enough that the free evolution of the system can be neglected during the measurement procedure. Then the

Hamiltonian can be approximated as $\boldsymbol{H} \simeq \lambda(t)\boldsymbol{M} \otimes \boldsymbol{P}$. If the coupling λ switches on suddenly at time zero and switches off suddenly at time T, the resulting time evolution operator is

$$\boldsymbol{U}(T) \simeq \exp\left(-i\lambda T \boldsymbol{M} \otimes \boldsymbol{P}\right). \tag{3.4}$$

Expanding in the basis in which $\boldsymbol{M}$ is diagonal,

$$\boldsymbol{M} = \sum_a |a\rangle M_a \langle a|, \tag{3.5}$$

we express $\boldsymbol{U}(T)$ as

$$\boldsymbol{U}(T) = \sum_a |a\rangle \exp\left(-i\lambda t M_a \boldsymbol{P}\right) \langle a|. \tag{3.6}$$

Now we recall that $\boldsymbol{P}$ generates a translation of the *position* of the pointer: $\boldsymbol{P} = -i\frac{d}{dx}$ in the position representation, so that $e^{-ix_0\boldsymbol{P}} = \exp\left(-x_0\frac{d}{dx}\right)$, and by Taylor expanding,

$$e^{-ix_0\boldsymbol{P}}\psi(x) = \psi(x - x_0); \tag{3.7}$$

In other words $e^{-ix_0\boldsymbol{P}}$ acting on a wavepacket translates the wavepacket by x_0. We see that if our quantum system starts in a superposition of $\boldsymbol{M}$ eigenstates, initially unentangled with the position-space wavepacket $|\psi(x)$ of the pointer, then after time T the quantum state has evolved to

$$\boldsymbol{U}(T)\left(\sum_a \alpha_a |a\rangle \otimes |\psi(x)\rangle\right) = \sum_a \alpha_a |a\rangle \otimes |\psi(x - \lambda T M_a)\rangle; \tag{3.8}$$

the position of the pointer has become correlated with the value of the observable $\boldsymbol{M}$. If the pointer wavepacket is narrow enough for us to resolve all values of the M_a that occur (that is, the width Δx of the packet is small compared to $\lambda T \Delta M_a$, where ΔM_a is the minimal gap between eigenvalues of $\boldsymbol{M}$), then when we observe the position of the pointer (never mind how!) we will prepare an eigenstate of the observable. With probability $|\alpha_a|^2$, we will detect that the pointer has shifted its position by $\lambda T M_a$, in which case we will have prepared the $\boldsymbol{M}$ eigenstate $|a\rangle$. We conclude that the initial state $|\varphi\rangle$ of the quantum system is projected to $|a\rangle$ with probability $|\langle a|\varphi\rangle|^2$. This is Von Neumann's model of orthogonal measurement.

The classic example is the Stern–Gerlach apparatus. To measure $\boldsymbol{\sigma}_3$ for a spin-$\frac{1}{2}$ object, we allow the object to pass through a region of inhomogeneous magnetic field

$$B_3 = \lambda z. \tag{3.9}$$

The magnetic moment of the object is $\mu\vec{\boldsymbol{\sigma}}$, and the coupling induced by the magnetic field is

$$\boldsymbol{H} = -\lambda\mu\boldsymbol{z}\boldsymbol{\sigma}_3. \tag{3.10}$$

In this case $\boldsymbol{\sigma}_3$ is the observable to be measured, coupled to the position $\boldsymbol{z}$ rather than the momentum of the pointer; thus, because $\boldsymbol{z}$ generates a translation of $\boldsymbol{P}_z$, the coupling imparts an *impulse* to the pointer which is correlated with its spin. We can perceive whether the object is pushed up or down, and so project out the spin state $|\uparrow_z\rangle$ or $|\downarrow_z\rangle$. By rotating the magnet, we could measure the observable $\hat{n}\cdot\vec{\boldsymbol{\sigma}}$ instead.

Thinking more abstractly, suppose that $\{\boldsymbol{E}_a, a = 0, 1, 2, \ldots N{-}1\}$ is a complete set of orthogonal projectors satisfying

$$\boldsymbol{E}_a\boldsymbol{E}_b = \delta_{ab}\boldsymbol{E}_a, \quad \boldsymbol{E}_a = \boldsymbol{E}_a^\dagger, \quad \sum_a \boldsymbol{E}_a = \boldsymbol{I}. \tag{3.11}$$

To perform an orthogonal measurement with these outcomes, we introduce an N-dimensional pointer system with fiducial orthonormal basis states $\{|a\rangle, a = 0, 1, 2, \ldots, N{-}1\}$, and, by coupling the system to the pointer, perform the unitary transformation

$$\boldsymbol{U} = \sum_{a,b} \boldsymbol{E}_a \otimes |b+a\rangle\langle b|. \tag{3.12}$$

Thus the pointer advances by an amount a if the state of the system is within the support of the projector $\boldsymbol{E}_a$. (The addition in $|b+a\rangle$ is understood to be modulo N; we may envision the pointer as a circular dial with N uniformly spaced tick marks.) The unitarity of $\boldsymbol{U}$ is easy to verify:

$$\begin{aligned}\boldsymbol{U}\boldsymbol{U}^\dagger &= \left(\sum_{a,b} \boldsymbol{E}_a \otimes |b+a\rangle\langle b|\right)\left(\sum_{c,d} \boldsymbol{E}_c \otimes |d\rangle\langle d+c|\right)\\ &= \sum_{a,b,c,d} \delta_{ac}\boldsymbol{E}_a \otimes \delta_{bd}|b+a\rangle\langle d+c|\\ &= \sum_a \boldsymbol{E}_a \otimes \sum_b |b+a\rangle\langle b+a| = \boldsymbol{I}\otimes\boldsymbol{I}.\end{aligned} \tag{3.13}$$

This unitary transformation acts on an initial product state of system and pointer according to

$$\boldsymbol{U} : |\Psi\rangle = |\psi\rangle\otimes|0\rangle \mapsto |\Psi'\rangle = \sum_a \boldsymbol{E}_a|\psi\rangle\otimes|a\rangle; \tag{3.14}$$

if the pointer is then measured in the fiducial basis, the measurement postulate implies that the outcome a occurs with probability

$$\mathrm{Prob}(a) = \langle\Psi'|\,(\boldsymbol{I}\otimes|a\rangle\langle a|)\,|\Psi'\rangle = \langle\psi|\boldsymbol{E}_a|\psi\rangle, \tag{3.15}$$

and that when this outcome occurs the normalized post-measurement state is

$$\frac{\boldsymbol{E}_a|\psi\rangle}{\|\boldsymbol{E}_a|\psi\rangle\|}. \tag{3.16}$$

If the measurement is performed and its outcome is not known, the initial pure state of the system becomes a mixture of these post-measurement states:

$$\sum_a \mathrm{Prob}(a)\frac{\boldsymbol{E}_a|\psi\rangle\langle\psi|\boldsymbol{E}_a}{\langle\psi|\boldsymbol{E}_a|\psi\rangle} = \sum_a \boldsymbol{E}_a|\psi\rangle\langle\psi|\boldsymbol{E}_a. \tag{3.17}$$

In fact, the system is described by this density operator once it becomes entangled with the pointer, whether we bother to observe the pointer or not. If the initial state of the system before the measurement is a mixed state with density matrix $\boldsymbol{\rho}$, then by expressing $\boldsymbol{\rho}$ as an ensemble of pure states we conclude that the measurement modifies the state according to

$$\boldsymbol{\rho} \mapsto \sum_a \boldsymbol{E}_a\boldsymbol{\rho}\boldsymbol{E}_a. \tag{3.18}$$

We see that if, by coupling the system to our pointer, we can execute suitable unitary transformations correlating the system and the pointer, and if we can observe the pointer in its fiducial basis, then we are empowered to perform any conceivable orthogonal measurement on the system.

3.1.2 Generalized measurements

In this discussion of orthogonal measurement, the fiducial basis of the pointer had two different roles — we assumed that the fiducial pointer states become correlated with the system projectors $\{\boldsymbol{E}_a\}$, and also that the measurement of the pointer projects onto the fiducial basis. In principle we could separate these two roles. Perhaps the unitary transformation applied to system and pointer picks out a different preferred basis than the basis in which the pointer is easily measured. Or perhaps the pointer which becomes entangled with the system is itself microscopic, and we may entangle it with a second macroscopic pointer in order to measure the microscopic pointer in whatever basis we prefer.

Suppose, to be concrete, that the system A is a single qubit, and so is the pointer B. They interact, resulting in the unitary map

$$\boldsymbol{U} : (\alpha|0\rangle + \beta|1\rangle)_A \otimes |0\rangle_B \mapsto \alpha|0\rangle_A \otimes |0\rangle_B + \beta|1\rangle_A \otimes |1\rangle_B. \tag{3.19}$$

Measuring the pointer by projecting onto the basis $\{|0\rangle, |1\rangle\}$ would induce an orthogonal measurement of the system, also in the $\{|0\rangle, |1\rangle\}$ basis. But suppose that we measure the pointer in a different basis instead, such as

$\{|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)\}$. Then the measurement postulate dictates that the two outcomes + and − occur equiprobably, and that the corresponding post-measurement states of the system are

$$\alpha|0\rangle \pm \beta|1\rangle. \tag{3.20}$$

In contrast to an orthogonal measurement of the system, these two post-measurement states are not orthogonal, unless $|\alpha| = |\beta|$. Furthermore, also in contrast to an orthogonal measurement, if two such measurements are performed in rapid succession, the outcomes need not be the same. We use the term *generalized measurement* to mean a measurement, like this one, which is not necessarily an orthogonal projection acting on the system.

It is convenient to describe this measurement procedure by expanding the entangled state of system and pointer in the basis in which the pointer is measured; hence we rewrite eq.(3.19) as

$$\boldsymbol{U} : |\psi\rangle_A \otimes |0\rangle_B \mapsto \boldsymbol{M}_+|\psi\rangle_A \otimes |+\rangle_B + \boldsymbol{M}_-|\psi\rangle_A \otimes |-\rangle_B, \tag{3.21}$$

where

$$\boldsymbol{M}_+ = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \frac{1}{\sqrt{2}}\boldsymbol{I}, \quad \boldsymbol{M}_- = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = \frac{1}{\sqrt{2}}\boldsymbol{\sigma}_3. \tag{3.22}$$

Evidently, by measuring B in the basis $\{|\pm\rangle\}$, we prepare A in one of the states $\boldsymbol{M}_\pm|\psi\rangle$, up to a normalization factor.

Now let's generalize this idea to the case where the pointer system B is N-dimensional, and the measurement of the pointer projects onto an orthonormal basis $\{|a\rangle, a = 0, 1, 2, \ldots, N-1\}$. Again we'll assume that the system A and pointer B are initially in a product state, then an entangling unitary transformation $\boldsymbol{U}$ correlates the system with the pointer. By expanding the action of $\boldsymbol{U}$ in the basis for B we obtain

$$\boldsymbol{U} : |\psi\rangle_A \otimes |0\rangle_B \mapsto \sum_a \boldsymbol{M}_a|\psi\rangle_A \otimes |a\rangle_B. \tag{3.23}$$

Since $\boldsymbol{U}$ is unitary, it preserves the norm of any input, which means that

$$1 = \left\| \sum_a \boldsymbol{M}_a|\psi\rangle \otimes |a\rangle \right\|^2 = \sum_{a,b} \langle\psi|\boldsymbol{M}_a^\dagger \boldsymbol{M}_b|\psi\rangle\langle a|b\rangle = \sum_a \langle\psi|\boldsymbol{M}_a^\dagger \boldsymbol{M}_a|\psi\rangle \tag{3.24}$$

for any $|\psi\rangle$; hence

$$\sum_a \boldsymbol{M}_a^\dagger \boldsymbol{M}_a = \boldsymbol{I}. \tag{3.25}$$

The complete orthogonal measurement projecting onto the pointer basis $\{|a\rangle_B\}$ is equivalent to the incomplete orthogonal measurement on AB with projectors $\{\boldsymbol{I} \otimes |a\rangle\langle a|\}$; the measurement postulate asserts that outcome a occurs with probability

$$\text{Prob}(a) = \|\boldsymbol{M}_a|\psi\rangle\|^2, \tag{3.26}$$

and that if outcome a occurs the post-measurement state of the system is

$$\frac{\boldsymbol{M}_a|\psi\rangle}{\|\boldsymbol{M}_a|\psi\rangle\|}. \tag{3.27}$$

The completeness relation $\sum_a \boldsymbol{M}_a^\dagger \boldsymbol{M}_a = \boldsymbol{I}$ ensures that the probabilities sum to one, but the possible post-measurement states need not be mutually orthogonal, nor are the measurements necessarily repeatable. If we perform the measurement twice in succession and obtain outcome a the first time, the conditional probability of obtaining outcome b in the second measurement is

$$\text{Prob}(b|a) = \frac{\|\boldsymbol{M}_b\boldsymbol{M}_a|\psi\rangle\|^2}{\|\boldsymbol{M}_a|\psi\rangle\|^2}. \tag{3.28}$$

The two measurements agree if $\text{Prob}(b|a) = \delta_{ba}$, which is satisfied for arbitrary initial states of the system only if $\boldsymbol{M}_b\boldsymbol{M}_a = \delta_{ba}\boldsymbol{M}_a$ up to a phase factor, *i.e.* in the case where the measurement is orthogonal.

We see that if the initial state of the system is the density operator $\boldsymbol{\rho}$ (realized as an ensemble of pure states), there is an operator $\boldsymbol{E}_a = \boldsymbol{M}_a^\dagger \boldsymbol{M}_a$ associated with each possible measurement outcome a, such that the probability of outcome a is

$$\text{Prob}(a) = \text{tr}\left(\boldsymbol{\rho}\boldsymbol{E}_a\right). \tag{3.29}$$

The measurement operators $\{\boldsymbol{E}_a\}$ form a complete set of Hermitian nonnegative operators; that is, they satisfy the properties:

1. *Hermiticity.* $\boldsymbol{E}_a = \boldsymbol{E}_a^\dagger$.
2. *Positivity.* $\langle\psi|\boldsymbol{E}_a|\psi\rangle \geq 0$ for any vector $|\psi\rangle$; we abbreviate this property by simply writing $\boldsymbol{E}_a \geq 0$.
3. *Completeness.* $\sum_a \boldsymbol{E}_a = \boldsymbol{I}$.

Such a partition of unity by nonnegative operators is called a *positive operator-valued measure*, or *POVM*. (The term *measure* is a bit heavy-handed in this finite-dimensional context; it becomes more apt when the index a can be continuously varying.)

We have seen how a POVM can arise when an orthogonal measurement is performed on a meter after the meter interacts with the system. In fact any POVM can be realized this way. We need only observe that a nonnegative Hermitian operator $\boldsymbol{E}_a$ has a nonnegative square root $\sqrt{\boldsymbol{E}_a}$; more generally, the operator

$$\boldsymbol{M}_a = \boldsymbol{U}_a \sqrt{\boldsymbol{E}_a} \tag{3.30}$$

obeys $\boldsymbol{M}_a^\dagger \boldsymbol{M}_a = \boldsymbol{E}_a$ where $\boldsymbol{U}_a$ is an arbitrary unitary operator — eq.(3.30) is the *polar decomposition* of the operator $\boldsymbol{M}_a$. Plugging into eq.(3.23) yields the unitary interaction which realizes the POVM $\{\boldsymbol{E}_a\}$. In this formulation, the post-measurement state corresponding to outcome a,

$$\boldsymbol{U}_a \left(\frac{\sqrt{\boldsymbol{E}_a}|\psi\rangle}{\|\sqrt{\boldsymbol{E}_a}|\psi\rangle\|} \right), \tag{3.31}$$

is arbitrary, since we are free to choose the unitary $\boldsymbol{U}_a$ however we please for each possible outcome. The POVM attributes a probability to each measurement outcome, but provides no guidance regarding the state after the measurement. Indeed, after the measurement we have the freedom to discard the state and replace it by whatever freshly prepared state we desire.

3.2 Quantum channels

3.2.1 The operator-sum representation

We now proceed to the next step in our program of understanding the behavior of one part of a bipartite quantum system. We have seen that a pure state of the bipartite system AB may behave like a mixed state when we observe subsystem A alone, and that an orthogonal measurement of the bipartite system can realize a (nonorthogonal) POVM on A alone. Next we ask, if a state of the bipartite system undergoes unitary evolution, how do we describe the evolution of A alone?

In effect, we have already answered this question in our discussion of generalized measurements. If system A starts out in a pure state $|\psi\rangle$ (unentangled with B), and then interacts with B, the joint state of AB has the form eq.(3.23); the resulting density operator for A is found by tracing out B. Equivalently, we may imagine measuring system B in the basis $\{|a\rangle\}$, but failing to record the measurement outcome, so we are forced to average over all the possible post-measurement states, weighted by their probabilities. The result is that the initial density operator $\boldsymbol{\rho} = |\psi\rangle\langle\psi|$ is subjected to a linear map $\mathcal{E}$, which acts as

$$\mathcal{E}(\boldsymbol{\rho}) = \sum_a \boldsymbol{M}_a \boldsymbol{\rho} \boldsymbol{M}_a^\dagger, \tag{3.32}$$

where the operators $\{\boldsymbol{M}_a\}$ obey the completeness relation eq.(3.25). Being linear, $\mathcal{E}$ acts as in eq.(3.32) not just on pure states, but on any density operator.

A linear map of the form eq.(3.32), where the $\{\boldsymbol{M}_a\}$ obey eq.(3.25), is called a *quantum channel*. The word "channel" is drawn from communication theory — we are to imagine a sender who transmits the state $\boldsymbol{\rho}$ though a communication link to another party who receives the modified state $\mathcal{E}(\boldsymbol{\rho})$. Sometimes the word *superoperator* is used as a synonym for quantum channel, where "super" conveys that the map takes operators to operators rather than vectors to vectors. Yet another name for the same object is *trace-preserving completely positive map*, or *TPCP* map for short. The justification for this name will emerge shortly. Eq.(3.32) is said to be an *operator-sum representation* of the quantum channel, and the operators $\{\boldsymbol{M}_a\}$ are called the *Kraus operators* or *operation elements* of the channel.

A quantum channel maps density operators to density operators; that is, has the following easily verified properties:

1. *Linearity.* $\mathcal{E}(\alpha\boldsymbol{\rho}_1 + \beta\boldsymbol{\rho}_2) = \alpha\mathcal{E}(\boldsymbol{\rho}_1) + \beta\mathcal{E}(\boldsymbol{\rho}_2)$.
2. *Preserves Hermiticity.* $\boldsymbol{\rho} = \boldsymbol{\rho}^\dagger$ implies $\mathcal{E}(\boldsymbol{\rho}) = \mathcal{E}(\boldsymbol{\rho})^\dagger$.
3. *Preserves positivity.* $\boldsymbol{\rho} \geq 0$ implies $\mathcal{E}(\boldsymbol{\rho}) \geq 0$.
4. *Preserves trace.* $\mathrm{tr}\,(\mathcal{E}(\boldsymbol{\rho})) = \mathrm{tr}\,(\boldsymbol{\rho})$.

These properties partially explain the locution "trace-preserving completely positive map," except that we are still missing the reason for the modifier "completely." That's coming soon.

We've seen how a quantum channel acting on system A arises from a unitary transformation acting on A and B followed by a partial trace on B. As in our discussion of generalized measurements, we can also run this argument backwards to see that any quantum channel may be realized this way. Given a quantum channel $\mathcal{E}$ acting on A with Kraus operators $\{\boldsymbol{M}_a\}$, we may introduce the auxiliary system B with Hilbert space dimension matching the number of Kraus operators. A unitary transformation may then be constructed whose action on $|\psi\rangle_A \otimes |0\rangle_B$ is as in eq.(3.23), from which the quantum channel $\mathcal{E}$ is obtained by tracing out B.

The operator-sum representation of a given quantum channel $\mathcal{E}$ is not unique, because we can perform the partial trace on B in any basis we please. When expressed in terms of rotated basis states $\{|\mu\rangle\}$ such that

$$|a\rangle = \sum_\mu |\mu\rangle V_{\mu a} \tag{3.33}$$

for unitary V, the joint state of AB becomes

$$\sum_{a,\mu} \boldsymbol{M}_a|\psi\rangle_A \otimes |\mu\rangle_B V_{\mu a} = \sum_{\mu} \boldsymbol{N}_\mu|\psi\rangle_A \otimes |\mu\rangle_B \tag{3.34}$$

where the new Kraus operators are

$$\boldsymbol{N}_\mu = \sum_a V_{\mu a}\boldsymbol{M}_a. \tag{3.35}$$

We will see soon that any two operator-sum representations of the same quantum channel are always related by such a unitary change of basis for the Kraus operators.

Quantum channels are important because they provide us with a formalism for discussing *decoherence*, the evolution of pure states into mixed states. *Unitary* evolution of $\boldsymbol{\rho}_A$ is the special case in which there is only one term in the operator sum. If there are two or more terms, then there are pure initial states of A which become entangled with B under evolution governed by the joint unitary transformation U_{AB}, and therefore the state of A becomes mixed when we trace out B.

Two channels $\mathcal{E}_1$ and $\mathcal{E}_2$ can be composed to obtain another channel $\mathcal{E}_2 \circ \mathcal{E}_1$; if $\mathcal{E}_1$ describes evolution from yesterday to today, and $\mathcal{E}_2$ describes evolution from today to tomorrow, then $\mathcal{E}_2 \circ \mathcal{E}_1$ describes the evolution from yesterday to tomorrow. Specifically, if $\mathcal{E}_1$ has an operator-sum representation with N Kraus operators $\{\boldsymbol{M}_a\}$, and $\mathcal{E}_2$ has an operator-sum representation with M Kraus operators $\{\boldsymbol{N}_\mu\}$, then $\mathcal{E}_2 \circ \mathcal{E}_1$ has an operator-sum representation with NM Kraus operators $\{\boldsymbol{N}_\mu\boldsymbol{M}_a\}$. Because we can compose them in this way, we say that quantum channels form a dynamical *semigroup*.

3.2.2 Reversibility

A unitary transformation $\boldsymbol{U}$ has a unitary inverse $\boldsymbol{U}^\dagger$. Thus if today's quantum state was obtained by applying $\boldsymbol{U}$ to yesterday's state, we can in principle recover yesterday's state by applying $\boldsymbol{U}^\dagger$ to today's state. Unitary time evolution is reversible.

Is the same true for general quantum channels? If channel $\mathcal{E}_1$ with Kraus operators $\{\boldsymbol{M}_a\}$ is inverted by channel $\mathcal{E}_2$ with Kraus operators $\{\boldsymbol{N}_\mu\}$, then for any pure state $|\psi\rangle$ we have

$$\mathcal{E}_2 \circ \mathcal{E}_1(|\psi\rangle\langle\psi|) = \sum_{\mu,a} \boldsymbol{N}_\mu\boldsymbol{M}_a|\psi\rangle\langle\psi|\boldsymbol{M}_a^\dagger\boldsymbol{N}_\mu^\dagger = |\psi\rangle\langle\psi|. \tag{3.36}$$

Since the left-hand side is a sum of positive terms, eq.(3.36) can hold only if each of these terms is proportional to $|\psi\rangle\langle\psi|$, hence

$$\boldsymbol{N}_\mu\boldsymbol{M}_a = \lambda_{\mu a}\boldsymbol{I} \tag{3.37}$$

for each μ and a. Using the completeness relation, we find

$$\boldsymbol{M}_b^\dagger \boldsymbol{M}_a = \boldsymbol{M}_b^\dagger \left(\sum_\mu \boldsymbol{N}_\mu^\dagger \boldsymbol{N}_\mu \right) \boldsymbol{M}_a = \sum_\mu \lambda_{\mu b}^* \lambda_{\mu a} \boldsymbol{I} \equiv \beta_{ba} \boldsymbol{I}. \tag{3.38}$$

where each β_{aa} is real and positive unless $\boldsymbol{M}_a = 0$.

Since the channel $\mathcal{E}_1$ maps system A to itself, each Kraus operator $\boldsymbol{M}_a$ is a square $d \times d$ matrix, where d is the dimension of A. Therefore $\boldsymbol{M}_a$ has a polar decomposition, which yields

$$\boldsymbol{M}_a = \boldsymbol{U}_a \sqrt{\boldsymbol{M}_a^\dagger \boldsymbol{M}_a} = \sqrt{\beta_{aa}}\, \boldsymbol{U}_a \tag{3.39}$$

for some unitary $\boldsymbol{U}_a$, and it then follows that

$$\boldsymbol{M}_b^\dagger \boldsymbol{M}_a = \sqrt{\beta_{aa}\beta_{bb}}\, \boldsymbol{U}_b^\dagger \boldsymbol{U}_a = \beta_{ba} \boldsymbol{I}, \tag{3.40}$$

and hence

$$\boldsymbol{U}_a = \frac{\beta_{ba}}{\sqrt{\beta_{aa}\beta_{bb}}}\, \boldsymbol{U}_b \tag{3.41}$$

for each a and b. We conclude that each Kraus operator $\boldsymbol{M}_a$ is proportional to a single unitary matrix, and hence that $\mathcal{E}_1$ is a unitary map. A quantum channel can be inverted by another quantum channel only if it is unitary.

We have found that decoherence is irreversible. Once system A becomes entangled with system B, we can't undo the damage to A if we don't have access to B. Decoherence causes quantum information to leak to a system's environment, and because we cannot control the environment this information cannot be recovered.

This argument applies to a channel which maps A to A' as long as A and A' have the same dimension. But the conclusion can be evaded if A' has a larger dimension than A, since in that case the Kraus operators are rectangular, and the polar decomposition does not apply. We will take advantage of this exception when developing the theory of quantum error correction in Chapter 7.

3.2.3 Quantum channels in the Heisenberg picture

We have described quantum channels using the *Schrödinger picture* in which the quantum state evolves with time. Sometimes it is convenient to use the *Heisenberg picture*, in which the state is stationary and the operators evolve instead.

When time evolution is unitary, in the Schrödinger picture the state vector at time t is obtained from the state vector at time 0 by

$$|\psi(t)\rangle = \boldsymbol{U}(t)|\psi(0)\rangle \tag{3.42}$$

where $\boldsymbol{U}(t)$ is unitary, and correspondingly a density operator evolves according to

$$\boldsymbol{\rho}(t) = \boldsymbol{U}(t)\boldsymbol{\rho}(0)\boldsymbol{U}(t)^\dagger. \tag{3.43}$$

In the Heisenberg picture the density operator $\boldsymbol{\rho}$ is fixed, and an operator $\boldsymbol{A}$ evolves according to

$$\boldsymbol{A}(t) = \boldsymbol{U}(t)^\dagger \boldsymbol{A}(0)\boldsymbol{U}(t). \tag{3.44}$$

This evolution law is chosen so that the two pictures agree on the expectation values of observables at any time:

$$\langle \boldsymbol{A} \rangle_{t,\mathrm{Schr}} = \mathrm{tr}\left(\boldsymbol{A}(0)\boldsymbol{\rho}(t)\right) = \mathrm{tr}\left(\boldsymbol{A}(t)\boldsymbol{\rho}(0)\right) = \langle \boldsymbol{A} \rangle_{t,\mathrm{Heis}}, \tag{3.45}$$

where we have used the cyclic property of the trace.

Likewise, if the $\mathcal{E}$ is a quantum channel which acts on density operators according to

$$\boldsymbol{\rho}' = \mathcal{E}(\boldsymbol{\rho}) = \sum_a \boldsymbol{M}_a \boldsymbol{\rho} \boldsymbol{M}_a^\dagger, \tag{3.46}$$

we may use an alternative description in which the state is fixed, but operators evolve as

$$\boldsymbol{A}' = \mathcal{E}^*(\boldsymbol{A}) = \sum_a \boldsymbol{M}_a^\dagger \boldsymbol{A} \boldsymbol{M}_a, \tag{3.47}$$

so that

$$\mathrm{tr}\left(\boldsymbol{A}\ \mathcal{E}(\boldsymbol{\rho})\right) = \mathrm{tr}\left(\mathcal{E}^*(\boldsymbol{A})\boldsymbol{\rho}\right). \tag{3.48}$$

We say that $\mathcal{E}^*$ is the *dual* or *adjoint* of $\mathcal{E}$.

Note that the dual of a channel need not be a channel, that is, might not be trace preserving. Instead, the completeness property of the Kraus operators $\{\boldsymbol{M}_a\}$ implies that

$$\mathcal{E}^*(\boldsymbol{I}) = \boldsymbol{I} \tag{3.49}$$

if $\mathcal{E}$ is a channel. We say that a map is *unital* if it preserves the identity operator, and conclude that the dual of a channel is a unital map.

Not all quantum channels are unital, but some are. If the Kraus operators of $\mathcal{E}$ satisfy

$$\sum_a \boldsymbol{M}_a^\dagger \boldsymbol{M}_a = \boldsymbol{I} = \sum_a \boldsymbol{M}_a \boldsymbol{M}_a^\dagger, \tag{3.50}$$

then $\mathcal{E}$ is unital and its dual $\mathcal{E}^*$ is also a unital channel. A unital quantum channel maps a maximally mixed density operator to itself; it is the quantum version of a doubly stochastic classical map, which maps probability distributions to probability distributions and preserves the uniform distribution.

3.2.4 Quantum operations

Generalized measurements and quantum channels are actually special cases of a more general notion called a *quantum operation*. As already noted, a generalized measurement can be realized by entangling a system with a meter and performing an orthogonal measurement on the meter, while a quantum channel arises if we measure the meter but completely forget the measurement outcome. In a quantum operation, we imagine measuring the meter, then retaining some of the information about the outcome while discarding the rest.

We may consider a generalized measurement described by Kraus operators $\{\boldsymbol{M}_{a\mu}\}$ which carry two labels, a and μ. These obey the usual completeness relation

$$\sum_{a,\mu} \boldsymbol{M}_{a\mu}^\dagger \boldsymbol{M}_{a\mu} = \boldsymbol{I}. \tag{3.51}$$

Suppose that, after a measurement that projects onto a definite value for both a and μ, we remember a but forget μ. Then, if the quantum state is $\boldsymbol{\rho}$ before the measurement, the post-measurement state (up to a normalization factor) is

$$\mathcal{E}_a(\boldsymbol{\rho}) \equiv \sum_{\mu} \boldsymbol{M}_{a\mu} \boldsymbol{\rho} \boldsymbol{M}_{a\mu}^\dagger, \tag{3.52}$$

where the outcome a occurs with probability

$$\text{Prob}(a) = \text{tr } \mathcal{E}_a(\boldsymbol{\rho}). \tag{3.53}$$

Eq.(3.52) looks like the operator-sum representation for a quantum channel, except that now instead of the completeness relation the Kraus operators obey an inequality constraint

$$\sum_{\mu} \boldsymbol{M}_{a\mu}^\dagger \boldsymbol{M}_{a\mu} \leq \boldsymbol{I}. \tag{3.54}$$

(We write $\boldsymbol{A} \leq \boldsymbol{I}$ as a shorthand for the statement that $\boldsymbol{I} - \boldsymbol{A}$ is a nonnegative operator; that is, the eigenvalues of the Hermitian operator $\boldsymbol{A}$ are no larger than 1.) Our earlier notion of a generalized measurement is recovered when μ takes just one value (all information about the outcome is retained), and the operation becomes a channel when a takes just one value (all information about the outcome is discarded).

The state needs to be renormalized to restore the unit trace condition; therefore under an operation the state really evolves nonlinearly according to

$$\boldsymbol{\rho} \mapsto \frac{\mathcal{E}_a(\boldsymbol{\rho})}{\text{tr } \mathcal{E}_a(\boldsymbol{\rho})}. \tag{3.55}$$

It is often convenient, though, to regard the operation as a linear map that takes $\boldsymbol{\rho}$ to a subnormalized state. For example, we may want to consider a sequence of n consecutive measurements with outcomes $\{a_1, a_2, \ldots, a_{n-1}, a_n\}$, where the ith measurement transforms the state according to the operation $\mathcal{E}_{a_i}$. Rather than renormalizing the state after each measurement, we can wait until after the final measurement in the sequence before renormalizing. The final state can then be written

$$\boldsymbol{\rho} \mapsto \frac{\mathcal{E}_{a_n} \circ \mathcal{E}_{a_{n-1}} \circ \cdots \circ \mathcal{E}_{a_2} \circ \mathcal{E}_{a_1}(\boldsymbol{\rho})}{\operatorname{tr} \mathcal{E}_{a_n} \circ \mathcal{E}_{a_{n-1}} \circ \cdots \circ \mathcal{E}_{a_2} \circ \mathcal{E}_{a_1}(\boldsymbol{\rho})} \tag{3.56}$$

where the normalizing factor in the denominator is just the probability of the observed sequence of measurement outcomes.

3.2.5 Linearity

A quantum channel specifies how an initial density operator evolves to a final density operator. Why on general grounds should we expect evolution of a quantum state to be described by a linear map? One possible answer is that nonlinear evolution would be incompatible with interpreting the density operator as an ensemble of possible states.

Suppose that $\mathcal{E}$ maps an initial state at time $t = 0$ to a final state at time $t = T$, and suppose that at time $t = 0$ the initial state $\boldsymbol{\rho}_i$ is prepared with probability p_i. Then the time-evolved state at $t = T$ will be $\mathcal{E}(\boldsymbol{\rho}_i)$ with probability p_i.

On the other hand we argued in Chapter 2 that an ensemble in which $\boldsymbol{\sigma}_i$ is prepared with probability q_i can be described by the convex combination of density operators

$$\boldsymbol{\sigma} = \sum_i q_i \boldsymbol{\sigma}_i. \tag{3.57}$$

Therefore the initial state is described by $\sum_i p_i \boldsymbol{\rho}_i$, which evolves to

$$\boldsymbol{\rho}' = \mathcal{E}\left(\sum_i p_i \boldsymbol{\rho}_i\right). \tag{3.58}$$

But we can also apply eq.(3.57) to the ensemble of final states, concluding that the final state may alternatively be described by

$$\boldsymbol{\rho}' = \sum_i p_i \mathcal{E}(\boldsymbol{\rho}_i). \tag{3.59}$$

Equating the two expressions for $\boldsymbol{\rho}'$ we find that $\mathcal{E}$ must act linearly, at least on convex combinations of states.

Similar reasoning applies to quantum operations, if we regard the normalization of an operation $\mathcal{E}_a$ as indicating the probability of the corresponding measurement outcome. Suppose again that the initial state $\boldsymbol{\rho}_i$ is prepared with *a priori* probability p_i and subsequently measured. If the state is $\boldsymbol{\rho}_i$ then measurement outcome a occurs with conditional probability $p(a|i)$, and the post-measurement state is $\mathcal{E}_a(\boldsymbol{\rho}_i)/p(a|i)$; hence the state ensemble after the measurement is described by the density operator

$$\boldsymbol{\rho}' = \sum_i p(i|a) \frac{\mathcal{E}_a(\boldsymbol{\rho}_i)}{p(a|i)}, \tag{3.60}$$

where $p(i|a)$ is the *a posteriori* probability that state $\boldsymbol{\rho}_i$ was prepared, taking into account the information gained by doing the measurement. On the other hand, applying the operation $\mathcal{E}_a$ to the convex combination of the initial states $\{\boldsymbol{\rho}_i\}$ yields

$$\boldsymbol{\rho}' = \frac{\mathcal{E}_a\left(\sum_i p_i \boldsymbol{\rho}_i\right)}{p_a}. \tag{3.61}$$

Invoking Bayes' rule

$$p_i p(a|i) = p_a p(i|a) \tag{3.62}$$

we see that the operation $\mathcal{E}_a$ is required to be a linear map:

$$\mathcal{E}_a\left(\sum_i p_i \boldsymbol{\rho}_i\right) = \sum_i p_i \mathcal{E}_a(\boldsymbol{\rho}_i). \tag{3.63}$$

3.2.6 Complete positivity

A quantum channel is a linear map taking density operators to density operators. In particular, if its input is a nonnegative operator than so is its output. We therefore say that a channel is a *positive map*.

But a channel has a stronger property than mere positivity; it is *completely positive*. This means that the channel remains positive even when we consider it to be acting on just part of a larger system.

If a channel $\mathcal{E}$ maps linear operators on Hilbert space $\mathcal{H}_A$ to linear operators on Hilbert space $\mathcal{H}_{A'}$, we will usually express this more economically by saying $\mathcal{E}$ maps A to A'. We may extend the input Hilbert space to $\mathcal{H}_A \otimes \mathcal{H}_B$, and consider the extended channel $\mathcal{E} \otimes I$ mapping AB to $A'B$. We say that $\mathcal{E}$ is completely positive if any such extension of $\mathcal{E}$ is positive.

Clearly, quantum channels are completely positive, because if $\mathcal{E}$ has an operator-sum representation with Kraus operators $\{\boldsymbol{M}_a\}$, then $\mathcal{E} \otimes I$ has an operator-sum representation with Kraus operators $\{\boldsymbol{M}_a \otimes \boldsymbol{I}\}$.

Likewise, quantum operations, though not necessarily trace preserving, are also completely positive.

It is perfectly reasonable to demand that a channel be completely positive if it is to describe the time evolution of a quantum system — even though the channel acts on just part of the world, it should map an initial state of the whole world to a final state of the whole world. It is therefore important to note that not all positive maps are completely positive; complete positivity is a stronger condition.

A simple example is the transpose map T, mapping the d-dimensional system A to itself. In a particular basis $\{|i\rangle\}$, T acts as

$$T : |i\rangle\langle j| \mapsto |j\rangle\langle i| \tag{3.64}$$

and hence

$$T : \boldsymbol{\rho} \mapsto \boldsymbol{\rho}^T. \tag{3.65}$$

The map T is evidently positive because

$$\langle\psi|\boldsymbol{\rho}^T|\psi\rangle = \sum_{i,j} \psi_j^* \left(\boldsymbol{\rho}^T\right)_{ji} \psi_i = \sum_{i,j} \psi_i \left(\boldsymbol{\rho}\right)_{ij} \psi_j^* = \langle\psi^*|\boldsymbol{\rho}|\psi^*\rangle \tag{3.66}$$

for any vector $|\psi\rangle$; therefore $\boldsymbol{\rho}^T$ is nonnegative if $\boldsymbol{\rho}$ is.

But T is not completely positive. Consider the (unconventionally normalized) maximally entangled state on AB, where B is also d-dimensional:

$$|\tilde{\Phi}\rangle_{AB} = \sum_{i=0}^{d-1} |i\rangle_A \otimes |i\rangle_B \equiv \sum_i |i,i\rangle. \tag{3.67}$$

The extension of T acts on this state as

$$T \otimes I : |\tilde{\Phi}\rangle\langle\tilde{\Phi}| = \sum_{i,j} |i\rangle\langle j| \otimes |i\rangle\langle j| \mapsto \sum_{i,j} |j\rangle\langle i| \otimes |i\rangle\langle j| \equiv \sum_{i,j} |j,i\rangle\langle i,j|; \tag{3.68}$$

that is, it maps $|\tilde{\Phi}\rangle\langle\tilde{\Phi}|$ to the SWAP operator which interchanges the systems A and B:

$$\text{SWAP} : |\psi\rangle_A \otimes |\varphi\rangle_B = \sum_{i,j} \psi_i\varphi_j|i,j\rangle \mapsto \sum_{i,j} \varphi_j\psi_i|j,i\rangle = |\varphi\rangle_A \otimes |\psi\rangle_B \tag{3.69}$$

Since the square of SWAP is the identity, its eigenvalues are ± 1. States which are symmetric under interchange of A and B have eigenvalue 1, while antisymmetric states have eigenvalue -1. Thus SWAP has negative eigenvalues, which means that $T \otimes I$ is not positive and therefore T is not completely positive.

3.3 Channel-state duality and the dilation of a channel

We have now seen that a quantum channel acting on A, which arises from a unitary map on an extension of A, is a completely positive linear map of density operators to density operators. We have also argued that linearity and complete positivity are properties that should hold for any reasonable evolution law on quantum states. It is therefore satisfying to note that any trace-preserving completely positive linear map is a quantum channel — it has an operator sum representation and a unitary realization. When considering the (in general nonunitary) evolution of A, we are entitled to imagine that A is part of an extended system which evolves unitarily.

3.3.1 Channel-state duality

To prove this statement we will use a trick which is interesting in its own right and also has other applications. For the purpose of studying the properties of a map $\mathcal{E}$ taking A to A', we introduce an auxiliary system R with the same dimension as A, which we call the *reference system*. If $\mathcal{E}$ is completely positive, it maps a maximally entangled state on RA to a nonnegative operator on RA'. Conversely, we may associate with any nonnegative operator on RA' a corresponding CP map taking A to A'. This correspondence between maps and states, called the *Choi-Jamiolkowski isomorphism* or *channel-state duality*, is a very useful tool.

To see how it works, consider how $I \otimes \mathcal{E}$ acts on the maximally entangled state

$$|\tilde{\Phi}\rangle_{RA} = \sum_{i=0}^{d-1} |i\rangle_R \otimes |i\rangle_A. \tag{3.70}$$

where A and R both have dimension d. This vector has norm $\sqrt{d}$ instead of norm 1; we choose this unconventional normalization, highlighted by the tilde over Φ, to avoid annoying factors of d in the formulas that follow. If $\mathcal{E}_{A\to A'}$ is completely positive, then $I \otimes \mathcal{E}$ maps $|\tilde{\Phi}\rangle\langle\tilde{\Phi}|$ (up to normalization) to a density operator on RA', which like any density operator can be realized by an ensemble of pure states; hence

$$(I \otimes \mathcal{E}) \left(\left(|\tilde{\Phi}\rangle\langle\tilde{\Phi}| \right)_{RA} \right) = \sum_a \left(|\tilde{\Psi}_a\rangle\langle\tilde{\Psi}_a| \right)_{RA'}. \tag{3.71}$$

Here the normalization of $|\tilde{\Psi}_a\rangle$ may depend on a; in order to make the equation look less cluttered, we've absorbed the probability of each pure state occurring in the ensemble into that state's normalization.

Now we notice that

$$|\varphi\rangle_A = \sum_i \varphi_i |i\rangle_A = \sum_i \varphi_i \left({}_R\langle i|\tilde{\Phi}\rangle_{RA} \right) = {}_R\langle\varphi^*|\tilde{\Phi}\rangle_{RA}; \tag{3.72}$$

using the linearity of $\mathcal{E}$, eq.(3.71) then implies

$$\mathcal{E}\left(\left(|\varphi\rangle\langle\varphi|\right)_A\right) = \sum_a \left(\langle\varphi^*|\tilde{\Psi}_a\rangle\langle\tilde{\Psi}_a|\varphi^*\rangle\right)_{A'}. \tag{3.73}$$

(This scheme for extracting the action on $|\varphi\rangle_A$ using the dual vector ${}_R\langle\varphi^*|$ is called the *relative-state method*.) Given a vector $|\tilde{\Phi}\rangle_{RA'}$, where R is d dimensional, we may define an operator $\boldsymbol{M}_a$ mapping $\mathcal{H}_A$ to $\mathcal{H}_{A'}$ (where A is d dimensional) by

$$\boldsymbol{M}_a|\varphi\rangle_A = {}_R\langle\varphi^*|\tilde{\Psi}_a\rangle_{RA'}; \tag{3.74}$$

it is easy to check that $\boldsymbol{M}_a$ is linear. Thus eq.(3.73) provides an operator-sum representation of $\mathcal{E}$ acting on the pure state $\left(|\varphi\rangle\langle\varphi|\right)_A$ (and hence by linearity acting on any density operator):

$$\mathcal{E}(\boldsymbol{\rho}) = \sum_a \boldsymbol{M}_a \boldsymbol{\rho} \boldsymbol{M}_a^\dagger. \tag{3.75}$$

We have now established the desired isomorphism between states and CP maps: Eq.(3.71) tells us how to obtain a state on RA' from the channel $\mathcal{E}_{A\to A'}$, while eq.(3.74) and eq.(3.75) tells us how to recover the CP map from the state. Furthermore, the $\{\boldsymbol{M}_a\}$ must obey the completeness relation $\sum_a \boldsymbol{M}_a^\dagger \boldsymbol{M}_a = \boldsymbol{I}$ if $\mathcal{E}$ is trace-preserving.

Put succinctly, the argument went as follows. Because $\mathcal{E}_{A\to A'}$ is completely positive, $I \otimes \mathcal{E}$ takes a maximally entangled state on RA to a density operator on RA', up to normalization. This density operator can be expressed as an ensemble of pure states, and each of these pure states is associated with a Kraus operator in the operator-sum representation of $\mathcal{E}$.

From this viewpoint, we see that the freedom to choose the Kraus operators representing a channel in many different ways is really the same thing as the freedom to choose the ensemble of pure states representing a density operator in many different ways. According to the HJW theorem, two different ensemble realizations of the same density operator,

$$(I\otimes\mathcal{E})\left(\left(|\tilde{\Phi}\rangle\langle\tilde{\Phi}|\right)_{RA}\right) = \sum_a \left(|\tilde{\Psi}_a\rangle\langle\tilde{\Psi}_a|\right)_{RA'} = \sum_\mu \left(|\tilde{\gamma}_\mu\rangle\langle\tilde{\gamma}_\mu|\right)_{RA'}, \tag{3.76}$$

are related by a unitary change of basis,

$$|\tilde{\gamma}_\mu\rangle = \sum_a V_{\mu a}|\tilde{\Psi}_a\rangle. \tag{3.77}$$

Correspondingly, two different sets of Kraus operators $\{\boldsymbol{M}_a\}$ and $\{\boldsymbol{N}_\mu\}$ representing the same channel are related by

$$\boldsymbol{N}_\mu = \sum_a V_{\mu a}\boldsymbol{M}_a \tag{3.78}$$

where $V_{\mu a}$ is a unitary matrix.

Channel-state duality also makes clear how many Kraus operators are needed to describe a channel. A channel $\mathcal{E}$ mapping A to A', where A has dimension d and A' has dimension d', is equivalent to a density operator on RA', where R has dimension d, and the minimal number of Kraus operators needed to represent the channel is the same as the minimal number of pure states needed in an ensemble representation of the density operator. This is the density operator's rank (number of nonzero eigenvalues), which is no larger than dd'. Of course, there may be operator-sum representations of $\mathcal{E}$ which use many more than this minimal number of Kraus operators, just as an ensemble representation of a density operator might use many more than the minimal number of pure states.

The number of free parameters needed to specify a channel mapping A to A' is the number $(dd')^2$ needed to specify a density operator on RA', except that there are d^2 constraints because the map is trace preserving for each of d^2 linearly independent inputs. Therefore the number of real free parameters is

$$d^2\left(d'^2-1\right). \tag{3.79}$$

This is 12 parameters for a general channel taking qubits to qubits. In contrast, a unitary map on qubits has only 3 parameters, aside from the physically irrelevant overall phase.

3.3.2 Stinespring dilation

Once we have an operator-sum representation of the channel $\mathcal{E}_{A\to A'}$, it is easy to see how $\mathcal{E}$ can be realized by a unitary map acting on an extended system. We introduce an extra system E, the channel's *environment*, which has dimension equal to the number of Kraus operators and orthonormal basis $\{|a\rangle\}$. Then we define an inner-product preserving map (an *isometry*) which takes A to $A'E$ according to

$$\boldsymbol{U}_{A\to A'E}:|\psi\rangle\mapsto\sum_a \boldsymbol{M}_a|\psi\rangle\otimes|a\rangle. \tag{3.80}$$

The completeness relation satisfied by the $\{\boldsymbol{M}_a\}$ implies $\boldsymbol{U}^\dagger\boldsymbol{U}=\boldsymbol{I}_A$. Though $\boldsymbol{U}$ may not actually be unitary, it might as well be, because we can easily extend an isometry to a unitary transformation by expanding the input Hilbert space. This isometry, which yields $\mathcal{E}_{A\to A'}$ when we trace out the environment, is called the *Stinespring dilation* of the channel

Another way to think about the construction of the Stinespring dilation is that we have used E to construct a *purification* of the density operator

arising from channel-state duality:

$$|\bar{\Psi}\rangle_{RA'E} = \sum_a |\tilde{\Psi}_a\rangle_{RA'} \otimes |a\rangle_E. \tag{3.81}$$

Apart from a normalization factor of $\sqrt{d}$, this is the pure state of $RA'E$ that results when the dilation acts on the maximally entangled state $|\tilde{\Phi}\rangle_{RA}$; we may recover the dilation from $|\bar{\Psi}\rangle$ using

$$\boldsymbol{U}_{A\to A'E}|\psi\rangle_A = {}_R\langle\psi^*|\bar{\Psi}\rangle_{RA'E}. \tag{3.82}$$

This succinct way to characterize a channel using a pure state is sometimes quite convenient, and we'll make heavy use of it when studying quantum channels in Chapter 10.

3.3.3 Axioms revisited

In Chapter 2 we stated the axioms of quantum mechanics in a form appropriate for closed systems. With the theory of open systems now in hand, we can give an alternative formulation with revised notions of how states, measurements, and evolution are described.

States. A state is a *density operator*, a nonnegative Hermitian operator in Hilbert space with unit trace.

Measurement. A measurement is a *positive operator-valued measure (POVM)*, a partition of unity by nonnegative operators. When the measurement $\{\boldsymbol{E}_a\}$ is performed on the state $\boldsymbol{\rho}$, the outcome a occurs with probability $\mathrm{tr}\,(\boldsymbol{E}_a\boldsymbol{\rho})$.

Dynamics. Time evolution is described by a *trace-preserving completely positive map (TPCP map)*.

One could regard either the open-system or closed-system version as the fundamental formulation of the theory; it's really a matter of taste. We have already seen how the open-system axioms are obtained starting from the closed-system axioms. Alternatively, starting with the open-system axioms, pure states arise as the extremal points in the space of density operators, or from the observation that every density operator has a purification in an extended system. Similarly, orthogonal measurements and unitary evolution arise naturally because every POVM can be realized by an orthogonal measurement in an extended system, and every trace-preserving completely positive map has an isometric Stinespring dilation. The notion that an open system may always be regarded as part of a larger closed system is fondly known as the *church of the larger Hilbert space*.

3.4 Three quantum channels

The best way to familiarize ourselves with the concept of a quantum channel is to study a few examples. We will now consider three examples (all interesting and useful) of channels acting on a single qubit. If we wish we may imagine that the channel $\mathcal{E}$ describes the fate of quantum information that is transmitted with some loss of fidelity from a sender to a receiver. Or, if we prefer, we may imagine that the transmission is in time rather than space; that is, $\mathcal{E}$ describes the time evolution of a quantum system that interacts with its environment.

3.4.1 Depolarizing channel

The *depolarizing channel* is a model of a decohering qubit that has particularly nice symmetry properties. We can describe it by saying that, with probability $1-p$ the qubit remains intact, while with probability p an "error" occurs. The error can be of any one of three types, where each type of error is equally likely. If $\{|0\rangle, |1\rangle\}$ is an orthonormal basis for the qubit, the three types of errors can be characterized as:

1. Bit flip error: $\begin{smallmatrix}|0\rangle\mapsto|1\rangle\\|1\rangle\mapsto|0\rangle\end{smallmatrix}$ or $|\psi\rangle \mapsto \boldsymbol{\sigma}_1|\psi\rangle, \boldsymbol{\sigma}_1 = \left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$,
2. Phase flip error: $\begin{smallmatrix}|0\rangle\mapsto|0\rangle\\|1\rangle\mapsto-|1\rangle\end{smallmatrix}$ or $|\psi\rangle \mapsto \boldsymbol{\sigma}_3|\psi\rangle, \boldsymbol{\sigma}_3 = \left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)$,
3. Both: $\begin{smallmatrix}|0\rangle\mapsto+i|1\rangle\\|1\rangle\mapsto-i|0\rangle\end{smallmatrix}$ or $|\psi\rangle \mapsto \boldsymbol{\sigma}_2|\psi\rangle, \boldsymbol{\sigma}_2 = \left(\begin{smallmatrix}0&-i\\i&0\end{smallmatrix}\right)$.

If an error occurs, then $|\psi\rangle$ evolves to an ensemble of the three states $\boldsymbol{\sigma}_1|\psi\rangle, \boldsymbol{\sigma}_2|\psi\rangle, \boldsymbol{\sigma}_3|\psi\rangle$, all occurring with equal likelihood.

Unitary representation. The depolarizing channel mapping qubit A to A can be realized by an isometry mapping A to AE, where E is a four-dimensional environment, acting as

$$\begin{aligned}\boldsymbol{U}_{A\to AE} : |\psi\rangle_A \mapsto \sqrt{1-p}\,|\psi\rangle_A \otimes |0\rangle_E \\ + \sqrt{\frac{p}{3}}\left(\boldsymbol{\sigma}_1|\psi\rangle_A \otimes |1\rangle_E + \boldsymbol{\sigma}_2|\psi\rangle_A \otimes |2\rangle_E + \boldsymbol{\sigma}_3|\psi\rangle_A \otimes |3\rangle_E\right).\end{aligned} \tag{3.83}$$

The environment evolves to one of four mutually orthogonal states that "keep a record" of what transpired; if we could only measure the environment in the basis $\{|a\rangle_E, a = 0, 1, 2, 3\}$, we would know what kind of error had occurred (and we would be able to intervene and reverse the error).

Operator-sum representation. To obtain an operator-sum representation of the channel, we evaluate the partial trace over the environment in the $\{|a\rangle_E\}$ basis. Then

$$\boldsymbol{M}_a = {}_E\langle a|\boldsymbol{U}, \tag{3.84}$$

so that

$$\boldsymbol{M}_0 = \sqrt{1-p}\,\boldsymbol{I},\ \boldsymbol{M}_1 = \sqrt{\frac{p}{3}}\,\boldsymbol{\sigma}_1,\ \boldsymbol{M}_2 = \sqrt{\frac{p}{3}}\,\boldsymbol{\sigma}_2,\ \boldsymbol{M}_3 = \sqrt{\frac{p}{3}}\,\boldsymbol{\sigma}_3. \tag{3.85}$$

Using $\boldsymbol{\sigma}_i^2 = \boldsymbol{I}$, we can readily check the normalization condition

$$\sum_a \boldsymbol{M}_a^\dagger \boldsymbol{M}_a = \left((1-p) + 3\frac{p}{3}\right)\boldsymbol{I} = \boldsymbol{I}. \tag{3.86}$$

A general initial density matrix $\boldsymbol{\rho}$ of the qubit evolves as

$$\boldsymbol{\rho} \mapsto \boldsymbol{\rho}' = (1-p)\boldsymbol{\rho} + \frac{p}{3}\left(\boldsymbol{\sigma}_1\boldsymbol{\rho}\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_2\boldsymbol{\rho}\boldsymbol{\sigma}_2 + \boldsymbol{\sigma}_3\boldsymbol{\rho}\boldsymbol{\sigma}_3\right). \tag{3.87}$$

where we are summing over the four (in principle distinguishable) possible final states of the environment.

Relative-state representation. We can also characterize the channel by introducing a reference qubit R and describing how a maximally-entangled state of the two qubits RA evolves, when the channel acts only on A. There are four mutually orthogonal maximally entangled states, which may be denoted

$$\begin{aligned}
|\phi^+\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle),\\
|\phi^-\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle),\\
|\psi^+\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle),\\
|\psi^-\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).
\end{aligned} \tag{3.88}$$

If the initial state is $|\phi^+\rangle_{RA}$, then when the depolarizing channel acts on qubit A, the entangled state evolves as

$$|\phi^+\rangle\langle\phi^+| \mapsto (1-p)|\phi^+\rangle\langle\phi^+| + \frac{p}{3}\Bigg(|\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-| + |\phi^-\rangle\langle\phi^-|\Bigg). \tag{3.89}$$

The "worst possible" quantum channel has $p = 3/4$, for in that case the initial entangled state evolves as

$$|\phi^+\rangle\langle\phi^+| \mapsto \frac{1}{4}\Big(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + |\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|\Big) = \frac{1}{4}\boldsymbol{I}; \tag{3.90}$$

it becomes the maximally mixed density matrix on RA. By the relative-state method, then, we see that a pure state $|\psi\rangle$ of qubit A evolves as

$$(|\psi\rangle\langle\psi|)_A \mapsto \ {}_R\langle\psi^*|2\left(\frac{1}{4}\boldsymbol{I}_{RA}\right)|\psi^*\rangle_R = \frac{1}{2}\boldsymbol{I}_A, \tag{3.91}$$

where the factor of two has been inserted because here we have used the standard normalization of the entangled states, instead of the unconventional normalization used in our earlier discussion of the relative-state method. We see that, for $p = 3/4$, the qubit is mapped to the maximally mixed density operator on A, irrespective of the value of the initial state $|\psi\rangle_A$. It is as though the channel threw away the initial quantum state, and replaced it by completely random junk.

An alternative way to express the evolution of the maximally entangled state is

$$|\phi^+\rangle\langle\phi^+| \mapsto \left(1 - \frac{4}{3}p\right)|\phi^+\rangle\langle\phi^+| + \frac{4}{3}p\left(\frac{1}{4}\boldsymbol{I}_{RA}\right). \tag{3.92}$$

Thus instead of saying that an error occurs with probability p, with errors of three types all equally likely, we could instead say that an error occurs with probability $4/3p$, where the error completely "randomizes" the state (at least we can say that for $p \leq 3/4$). The existence of two natural ways to define an "error probability" for this channel can sometimes cause confusion.

One useful measure of how well the channel preserves the original quantum information is called the "entanglement fidelity" F_e. It quantifies how "close" the final density matrix is to the original maximally entangled state $|\phi^+\rangle$ after the action of $I \otimes \mathcal{E}$:

$$F_e = \langle\phi^+|\boldsymbol{\rho}'|\phi^+\rangle. \tag{3.93}$$

For the depolarizing channel, we have $F_e = 1 - p$, and we can interpret F_e as the probability that no error occurred.

Bloch-sphere representation. It is also instructive to see how the depolarizing channel acts on the Bloch sphere. An arbitrary density matrix for a single qubit can be written as

$$\boldsymbol{\rho}(\vec{P}) = \frac{1}{2}\left(\boldsymbol{I} + \vec{P}\cdot\vec{\boldsymbol{\sigma}}\right), \tag{3.94}$$

where $\vec{P}$ is the "spin polarization" of the qubit. The depolarizing channel maps this state to

$$\boldsymbol{\rho}' = \left(1 - \frac{4}{3}p\right)\boldsymbol{\rho} + \frac{4}{3}p\boldsymbol{I} = \boldsymbol{\rho}(\vec{P}') \tag{3.95}$$

where

$$\vec{P}' = \left(1 - \frac{4}{3}p\right)\vec{P} \tag{3.96}$$

Hence the Bloch sphere contracts uniformly under the action of the channel (for $p \leq 3/4$); the spin polarization shrinks by the factor $1 - \frac{4}{3}p$ (which is why we call it the depolarizing channel).

Reversibility? Why do we say that the channel is not invertible? Evidently we can reverse a uniform contraction of the sphere with a uniform inflation. But the trouble is that the inflation of the Bloch sphere is not a channel, because it is not positive. Inflation will take some values of $\vec{P}$ with $|\vec{P}| \leq 1$ to values with $|\vec{P}| > 1$, and so will take a density operator to an operator with a negative eigenvalue. Decoherence can shrink the ball, but no physical process can blow it up again! A channel running backwards in time is *not* a channel.

3.4.2 Dephasing channel

Our next example is the *dephasing channel*, also called the *phase-damping channel*. This case is particularly instructive, because it provides a revealing caricature of decoherence in realistic physical situations, with all inessential mathematical details stripped away.

Unitary representation. An isometric representation of the channel is

$$\begin{aligned} |0\rangle_A &\mapsto \sqrt{1-p}\,|0\rangle_A \otimes |0\rangle_E + \sqrt{p}\,|0\rangle_A \otimes |1\rangle_E, \\ |1\rangle_A &\mapsto \sqrt{1-p}\,|1\rangle_A \otimes |0\rangle_E + \sqrt{p}\,|1\rangle_A \otimes |2\rangle_E. \end{aligned} \tag{3.97}$$

In this case, unlike the depolarizing channel, qubit A does not make any transitions in the $\{|0\rangle, |1\rangle\}$ basis. Instead, the environment "scatters" off of the qubit occasionally (with probability p), being kicked into the state $|1\rangle_E$ if A is in the state $|0\rangle_A$ and into the state $|2\rangle_E$ if A is in the state $|1\rangle_A$. Furthermore, also unlike the depolarizing channel, the channel picks out a preferred basis for qubit A; the basis $\{|0\rangle, |1\rangle\}$ is the only basis in which bit flips never occur.

Kraus operators. Evaluating the partial trace over E in the $\{|0\rangle_E, |1\rangle_E, |2\rangle_E\}$ basis, we obtain the Kraus operators

$$\boldsymbol{M}_0 = \sqrt{1-p}\,\boldsymbol{I}, \quad \boldsymbol{M}_1 = \sqrt{p}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \boldsymbol{M}_2 = \sqrt{p}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.98}$$

It is easy to check that $\boldsymbol{M}_0^2 + \boldsymbol{M}_1^2 + \boldsymbol{M}_2^2 = \boldsymbol{I}$. In this case, three Kraus operators are not really needed; a representation with two Kraus operators is possible. Expressing

$$\boldsymbol{M}_1 = \frac{\sqrt{p}}{2}\left(\boldsymbol{I} + \boldsymbol{\sigma}_3\right), \quad \boldsymbol{M}_2 = \frac{\sqrt{p}}{2}\left(\boldsymbol{I} - \boldsymbol{\sigma}_3\right), \tag{3.99}$$

we find

$$\mathcal{E}(\boldsymbol{\rho}) = \sum_a \boldsymbol{M}_a \boldsymbol{\rho} \boldsymbol{M}_a = \left(1 - \frac{1}{2}p\right)\boldsymbol{\rho} + \frac{1}{2}p\,\boldsymbol{\sigma}_3 \boldsymbol{\rho} \boldsymbol{\sigma}_3, \tag{3.100}$$

so an alternative description of the channel is that $\boldsymbol{\sigma}_3$ is applied with probability $p/2$ and nothing happens with probability $(1 - p/2)$. An initial density matrix $\boldsymbol{\rho}$ evolves to

$$\mathcal{E}\begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} = \begin{pmatrix} \rho_{00} & (1-p)\rho_{01} \\ (1-p)\rho_{10} & \rho_{11} \end{pmatrix}; \tag{3.101}$$

the on-diagonal terms in $\boldsymbol{\rho}$ remain invariant while the off-diagonal terms decay.

Continuous dephasing. We may also consider dephasing that occurs continuously in time. Suppose that the probability of a scattering event per unit time is Γ, so that $p = \Gamma \Delta t \ll 1$ when a brief time interval Δt elapses. The evolution over a time $t = n\Delta t$ is governed by $\mathcal{E}^n$ ($\mathcal{E}$ repeated n times in succession), so that the off-diagonal terms in the density operator become suppressed by

$$(1-p)^n = (1 - \Gamma t/n)^n \to e^{-\Gamma t}, \tag{3.102}$$

taking the limit $n \to \infty$ with t fixed. Thus, if we prepare an initial pure state $\alpha|0\rangle + \beta|1\rangle$, then after a time $t \gg \Gamma^{-1}$, the density operator evolves as

$$\begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix} \mapsto \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}; \tag{3.103}$$

The state decoheres, in the preferred basis $\{|0\rangle, |1\rangle\}$.

Bloch-sphere representation. We can compute how the polarization of the density operator evolves using the representation of the channel eq.(3.100), finding

$$\boldsymbol{\rho}(\vec{P}) = \frac{1}{2}\left(\boldsymbol{I} + \vec{P}\cdot\vec{\boldsymbol{\sigma}}\right) \mapsto \boldsymbol{\rho}(\vec{P}') \tag{3.104}$$

where

$$P'_{1,2} = (1-p)P_{1,2}, \quad P'_3 = P_3; \tag{3.105}$$

the Bloch ball shrinks to a prolate spheroid aligned with the z axis. Under continuous dephasing, the ball deflates in the x-y plane, degenerating to the z axis in the limit of large Γt.

You might wonder whether there is a quantum channel which causes just one component of the polarization to decay, mapping the Bloch ball to an oblate spheroid which touches the Bloch sphere along its equator. In fact no such map can be completely positive (the *no-pancake theorem*). Up to a unitary change of basis, a pancake map shrinks the value of P_2 while preserving P_1 and P_3, which is equivalent to taking the transpose of the density operator with nonzero probability, and we have seen in §3.2.6 that the transpose map is not completely positive.

Interpretation. We might interpret the phase-damping channel as describing a heavy "classical" particle (e.g., an interstellar dust grain) interacting with a background gas of light particles (e.g., the $3K$ microwave photons). We can imagine that the dust is initially prepared in a superposition of position eigenstates $|\psi\rangle = \frac{1}{\sqrt{2}}(|x\rangle + |-x\rangle)$ (or more realistically a superposition of position-space wavepackets with little overlap). We might be able to monitor the behavior of the dust particle, but it is hopeless to keep track of the quantum state of all the photons that scatter from the particle; for our purposes, the quantum state of the particle is described by the density matrix $\boldsymbol{\rho}$ obtained by tracing over the photon degrees of freedom.

Our analysis of the phase damping channel indicates that if photons are scattered by the dust particle at a rate Γ, then the off-diagonal terms in $\boldsymbol{\rho}$ decay like $\exp(-\Gamma t)$, and so become completely negligible for $t \gg \Gamma^{-1}$. At that point, the coherence of the superposition of position eigenstates is completely lost – there is no chance that we can recombine the wavepackets and induce them to interfere. If we attempt to do a double-slit interference experiment with dust grains, we will not see any interference pattern if it takes a time $t \gg \Gamma^{-1}$ for the grain to travel from the source to the screen.

The dust grain is heavy. Because of its large inertia, its state of motion is little affected by the scattered photons. Thus, there are two disparate

time scales relevant to its dynamics. On the one hand, there is a damping time scale, the time for a significant amount of the particle's momentum to be transfered to the photons, which is a long time for such a heavy particle. On the other hand, there is the decoherence time scale. In this model, the time scale for decoherence is of order Γ, the time for a *single* photon to be scattered by the dust grain, which is far shorter than the damping time scale. For a macroscopic object, decoherence is *fast*.

As we have already noted, the phase-damping channel picks out a preferred basis for decoherence, which in our "interpretation" we have assumed to be the position-eigenstate basis. Physically, decoherence prefers the spatially localized states of the dust grain because the *interactions* of photons and grains are localized in space. Grains in distinguishable positions tend to scatter the photons of the environment into mutually orthogonal states.

Even if the separation between the "grains" were so small that it could not be resolved very well by the scattered photons, the decoherence process would still work in a similar way. Perhaps photons that scatter off grains at positions x and $-x$ are not mutually orthogonal, but instead have an overlap

$$\langle \gamma + | \gamma - \rangle = 1 - \varepsilon, \quad \varepsilon \ll 1. \tag{3.106}$$

The phase-damping channel would still describe this situation, but with p replaced by $p\varepsilon$ (if p is still the probability of a scattering event). Thus, the decoherence rate would become $\Gamma_{\rm dec} = \varepsilon \Gamma_{\rm scat}$, where $\Gamma_{\rm scat}$ is the scattering rate.

The intuition we distill from this simple model applies to a wide variety of physical situations. A coherent superposition of macroscopically distinguishable states of a "heavy" object decoheres very rapidly compared to its damping rate. The spatial locality of the interactions of the system with its environment gives rise to a preferred "local" basis for decoherence. The same principle applies to Schrödinger's unfortunate cat, delicately prepared in a coherent superposition of its dead state and its alive state, two states that are easily distinguished by spatially localized probes. The cat quickly interacts with its environment, which is "scattered" into one of two mutually orthogonal states perfectly correlated with the cat's state in the $\{|\text{dead}\rangle, |\text{alive}\rangle\}$ basis, thus transforming the cat into an incoherent mixture of those two basis states.

Visibility. On the other hand, for microscopic systems the time scale for decoherence need not be short compared to dynamical time scales. Consider for example a single atom, initially prepared in a uniform superposition of its ground state $|0\rangle$ and an excited state $|1\rangle$ with energy $\hbar\omega$ above the ground state energy. Neglecting decoherence, after time t the atom's

state will be

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{-i\omega t}|1\rangle\right). \tag{3.107}$$

If dephasing occurs in the $\{|0\rangle, |1\rangle\}$ basis with rate Γ, the off-diagonal terms in the density operator decay, yielding the density operator

$$\boldsymbol{\rho}(t) = \frac{1}{2}\begin{pmatrix} 1 & e^{i\omega t}e^{-\Gamma t} \\ e^{-i\omega t}e^{-\Gamma t} & 1 \end{pmatrix}. \tag{3.108}$$

If after time t we measure the atom in the basis

$$|\pm\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle \pm |1\rangle\right), \tag{3.109}$$

the probability of the + outcome is

$$\text{Prob}(+,t) = \langle +|\boldsymbol{\rho}(t)|+\rangle = \frac{1}{2}\left(1 + e^{-\Gamma t}\cos\omega t\right). \tag{3.110}$$

In principle this time dependence of the probability can be measured by varying the time t between the preparation and measurement, and by repeating the experiment many times for each t to estimate the probability with high statistical confidence. The decoherence rate Γ can be determined experimentally by fitting the declining *visibility* of the coherent oscillations of $\text{Prob}(+,t)$ to a decaying exponential function of t.

3.4.3 Amplitude-damping channel

The *amplitude-damping channel* is a schematic model of the decay of an excited state of a (two-level) atom due to spontaneous emission of a photon. By detecting the emitted photon ("observing the environment") we can perform a POVM that gives us information about the initial preparation of the atom.

Unitary representation. We denote the atomic ground state by $|0\rangle_A$ and the excited state of interest by $|1\rangle_A$. The "environment" is the electromagnetic field, assumed initially to be in its vacuum state $|0\rangle_E$. After we wait a while, there is a probability p that the excited state has decayed to the ground state and a photon has been emitted, so that the environment has made a transition from the state $|0\rangle_E$ ("no photon") to the state $|1\rangle_E$ ("one photon"). This evolution is described by a unitary transformation that acts on atom and environment according to

$$\begin{aligned} &|0\rangle_A \otimes |0\rangle_E \mapsto |0\rangle_A \otimes |0\rangle_E \\ &|1\rangle_A \otimes |0\rangle_E \mapsto \sqrt{1-p}\,|1\rangle_A \otimes |0\rangle_E + \sqrt{p}\,|0\rangle_A \otimes |1\rangle_E. \end{aligned} \tag{3.111}$$

(Of course, if the atom starts out in its ground state, and the environment in its vacuum state, then no transition occurs.)

Kraus operators. By evaluating the partial trace over the environment in the basis $\{|0\rangle_E, |1\rangle_E\}$, we find the Kraus operators

$$\boldsymbol{M}_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \quad \boldsymbol{M}_1 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}, \tag{3.112}$$

and we can check that

$$\boldsymbol{M}_0^\dagger \boldsymbol{M}_0 + \boldsymbol{M}_1^\dagger \boldsymbol{M}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1-p \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & p \end{pmatrix} = \boldsymbol{I}. \tag{3.113}$$

The operator $\boldsymbol{M}_1$ induces a "quantum jump," the decay from $|1\rangle_A$ to $|0\rangle_A$, and $\boldsymbol{M}_0$ describes how the state changes if no jump occurs. The density matrix evolves as

$$\begin{aligned} \boldsymbol{\rho} \mapsto \mathcal{E}(\boldsymbol{\rho}) &= \boldsymbol{M}_0 \boldsymbol{\rho} \boldsymbol{M}_0^\dagger + \boldsymbol{M}_1 \boldsymbol{\rho} \boldsymbol{M}_1^\dagger \\ &= \begin{pmatrix} \rho_{00} & \sqrt{1-p}\,\rho_{01} \\ \sqrt{1-p}\,\rho_{10} & (1-p)\,\rho_{11} \end{pmatrix} + \begin{pmatrix} p\rho_{11} & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \rho_{00} + p\rho_{11} & \sqrt{1-p}\,\rho_{01} \\ \sqrt{1-p}\,\rho_{10} & (1-p)\,\rho_{11} \end{pmatrix}. \end{aligned} \tag{3.114}$$

Time dependence. If Γ is the spontaneous decay rate per unit time, then the decay occurs with probability $p = \Gamma \Delta t \ll 1$ in a small time interval Δt. We find the density operator after time $t = n\Delta t$ by applying the channel n times in succession. The ρ_{11} matrix element then decays as

$$\rho_{11} \mapsto (1-p)^n \rho_{11}, \quad \text{where} \quad (1-p)^n = (1 - \Gamma t/n)^n \to e^{-\Gamma t}, \tag{3.115}$$

the expected exponential decay law, while the off-diagonal entries decay by the factor $(1-p)^{n/2} = e^{-\Gamma t/2}$; hence we find

$$\boldsymbol{\rho}(t) = \begin{pmatrix} \rho_{00} + \left(1 - e^{-\Gamma t}\right) \rho_{11} & e^{-\Gamma t/2} \rho_{01} \\ e^{-\Gamma t/2}\,\rho_{10} & e^{-\Gamma t} \rho_{11} \end{pmatrix} \tag{3.116}$$

It is customary to use "T_1" to denote the exponential decay time for the excited population, and to use "T_2" to denote the exponential decay time for the off-diagonal terms in the density operator. In some systems where dephasing is very rapid T_2 is much shorter than T_1, but we see that for the amplitude-damping channel these two times are related and comparable:

$$T_2 = 2\Gamma^{-1} = 2T_1. \tag{3.117}$$

By the time that $t \gg T_1$, the atom is in its ground state with high probability ($\rho_{00}(t) \approx 1$).

Watching the environment. So far we have described the evolution of the qubit under the assumption that the state of the environment is not observed. But now suppose we surround the atom with photon detectors, so we know whether a photon has been emitted or not. Rather than a channel, then, we consider a POVM performed on the atom.

Returning to the joint unitary dynamics of system and environment, we see that a coherent superposition of the atomic ground and excited states evolves as

$$\begin{aligned}(\alpha|0\rangle_A + \beta|1\rangle_A) \otimes |0\rangle_E \\ \mapsto (\alpha|0\rangle_A + \beta\sqrt{1-p}\ |1\rangle_A) \otimes |0\rangle_E + \beta\sqrt{p}\ |0\rangle_A \otimes |1\rangle_E;\end{aligned} \quad (3.118)$$

To describe the system evolving continuously in time, we may consider applying this unitary map $n \gg 1$ times in succession, but where photons emitted at different times are perfectly distinguishable and hence orthogonal. The resulting POVM has $n+1$ Kraus operators, associated with the vacuum state of the environment and n different possible single photon states:

$$\boldsymbol{M}_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{(1-p)^n} \end{pmatrix}, \quad \boldsymbol{M}_k = \begin{pmatrix} 0 & \sqrt{(1-p)^{k-1}p} \\ 0 & 0 \end{pmatrix}, \quad (3.119)$$

for $k = 1, 2, \ldots n$. Taking the continuous-time limit we find that if no spontaneous decay occurs for time t, the corresponding Kraus operator is

$$\boldsymbol{M}_0 = \begin{pmatrix} 1 & 0 \\ 0 & e^{-\Gamma t/2} \end{pmatrix}. \quad (3.120)$$

If we detect a photon (and so project out a single-photon state of the environment), then we have prepared the state $|0\rangle_A$ of the atom. Not only that, we have prepared a state in which we know with certainty that the initial atomic state was the excited state $|1\rangle_A$; if the atom had started out in the ground state than it could not have decayed and no photon could have been detected.

On the other hand, if we detect no photon, and our photon detector has perfect efficiency, then we have projected out the vacuum state of the environment, and so have prepared the atomic state

$$\boldsymbol{M}_0 \left(\alpha|0\rangle + \beta|1\rangle\right) = \alpha|0\rangle + e^{-\Gamma t/2}\beta|1\rangle, \quad (3.121)$$

up to a normalization factor. As time goes by, the *a posteriori* quantum state has larger and larger overlap with the ground state, because if it had started out in the excited state it should have decayed by now. In the limit $t \to \infty$ our POVM becomes an orthogonal measurement: either a photon is detected, in which case the initial state of the atom must have

been $|1\rangle$, or no photon is detected, in which case the initial state must have been $|0\rangle$. It's odd but true: we can project out the state $|0\rangle$ of the atom by *not* detecting anything.

3.5 Master equations for open quantum systems

3.5.1 Markovian evolution

Quantum channels provide a general description of the evolution of density operators, including the evolution of pure states to mixed states (decoherence). In the same sense, unitary transformations provide a general description of coherent quantum evolution. But in the case of coherent evolution, we often find it very convenient to characterize the dynamics of a quantum system with a *Hamiltonian*, which describes the evolution over an infinitesimal time interval. The dynamics is then encoded in a differential equation, the *Schrödinger equation*, and we may calculate the evolution over a finite time interval by integrating the equation, that is, by piecing together the evolution over many infinitesimal intervals. Likewise, it is often possible to describe the (not necessarily coherent) evolution of a density operator, at least to a good approximation, by a differential equation which is called the *master equation*.

It is not obvious that there should be a differential equation that describes the decoherence of an open system. Such a description is possible only if the evolution of the quantum system is *Markovian*, that is, *local* in time. For the evolution of the density operator $\boldsymbol{\rho}(t)$ to be governed by a (first-order) differential equation in t, $\boldsymbol{\rho}(t+dt)$ must be completely determined by $\boldsymbol{\rho}(t)$.

In the case of an open system A, we are to imagine that its evolution is actually unitary on the extended system AE, where E is the environment. But though the evolution of AE may be governed by a Schrödinger equation, that's not enough to ensure that the time evolution is Markovian for A by itself. The trouble is that information can flow from A to E and then return at a later time. In that case the density operator $\boldsymbol{\rho}_A(t+dt)$ is not fully determined by $\boldsymbol{\rho}_A(t)$; we need to know $\boldsymbol{\rho}_A$ at earlier times as well.

This quandary arises because information flow is a two-way street. An open system (whether classical or quantum) is *dissipative* because information and energy can flow from the system to the environment. But that means that information can also flow back from environment to system, resulting in non-Markovian *fluctuations* of the system. This inescapable connection underlies the fluctuation-dissipation theorem, a widely applicable tool of statistical physics.

For any open system these fluctuations are inevitable, and an exact

Markovian description of quantum dynamics is impossible. Nevertheless, a Markovian description can be a very good approximation if there is a clean separation between the typical correlation time of the fluctuations and the time scale of the evolution that we want to follow. Crudely speaking, we may denote by $(\Delta t)_{\rm env}$ the time it takes for the environment to "forget" information it acquired from the system — after time $(\Delta t)_{\rm env}$ we can regard that information as lost forever, and neglect the possibility that the information may return to influence the subsequent evolution of the system.

To describe the evolution we "coarse-grain" in time, perceiving the dynamics through a filter that screens out the high frequency components of the motion with $\omega \gg (\Delta t_{\rm coarse})^{-1}$. An approximately Markovian description should be possible for $(\Delta t)_{\rm env} \ll (\Delta t)_{\rm coarse}$; we may neglect the memory of the reservoir if we are unable to resolve its effects. This Markovian approximation is *useful* if the time scale of the dynamics that we want to observe is long compared to $(\Delta t)_{\rm coarse}$, for example if the *damping* time scale $(\Delta t)_{\rm damp}$ satisfies

$$(\Delta t)_{\rm damp} \gg (\Delta t)_{\rm coarse} \gg (\Delta t)_{\rm env}. \tag{3.122}$$

This is a good approximation in some physical settings, like an atom interacting with the radiation field, but more dubious in other cases, like an electron spin interacting with nuclear spins in a semiconductor.

We could attempt to derive the master equation starting with the Schrödinger equation for AE, treating the coupling between A and E in time-dependent perturbation theory, and carefully introducing a frequency cutoff, but we won't do that here. Instead let's take it for granted that the dynamics is Markovian, and use the theory of quantum channels to infer the form of the master equation.

3.5.2 The Liouvillian

For a closed quantum system, time evolution is governed by a self-adjoint Hamiltonian $\boldsymbol{H}$ according to

$$|\psi(t+dt)\rangle = (\boldsymbol{I} - idt\boldsymbol{H})\,|\psi(t)\rangle, \tag{3.123}$$

and correspondingly the density operator evolves as

$$\boldsymbol{\rho}(t+dt) = \boldsymbol{\rho}(t) - idt[\boldsymbol{H}, \boldsymbol{\rho}(t)]. \tag{3.124}$$

In the case of an open quantum system, Markovian evolution for the infinitesimal time interval dt may be expressed as

$$\boldsymbol{\rho}(t+dt) = \mathcal{E}_{dt}(\boldsymbol{\rho}(t)), \tag{3.125}$$

where $\mathcal{E}_{dt}$ is a quantum channel. By adopting this Markovian form, we take the view that, after each infinitesimal time increment in the joint evolution of the system and its environment, the state of the environment is discarded and replaced by a fresh state of the environment unentangled with the system. We already made this assumption implicitly when discussing continuous-time dephasing and spontaneous decay in §3.4.

Expanding $\mathcal{E}_{dt}$ to linear order,

$$\mathcal{E}_{dt} = \boldsymbol{I} + dt\mathcal{L} \tag{3.126}$$

we find

$$\dot{\boldsymbol{\rho}} = \mathcal{L}(\boldsymbol{\rho}), \tag{3.127}$$

where the linear map $\mathcal{L}$ generating time evolution is called the *Liouvillian* or *Lindbladian*. This evolution equation has the formal solution

$$\boldsymbol{\rho}(t) = \lim_{n\to\infty}\left(1 + \frac{\mathcal{L}t}{n}\right)^n (\boldsymbol{\rho}(0)) = e^{\mathcal{L}t}(\boldsymbol{\rho}(0)) \tag{3.128}$$

if $\mathcal{L}$ is time independent.

The channel has an operator-sum representation

$$\boldsymbol{\rho}(t+dt) = \mathcal{E}_{dt}(\boldsymbol{\rho}(t)) = \sum_a \boldsymbol{M}_a \boldsymbol{\rho}(t) \boldsymbol{M}_a^\dagger = \boldsymbol{\rho}(t) + O(dt), \tag{3.129}$$

where, if we retain only terms up to linear order in dt, we may assume without loss of generality that $\boldsymbol{M}_0 = \boldsymbol{I} + O(dt)$, and that $\boldsymbol{M}_a$ is of order $\sqrt{dt}$ for $a > 0$. Each of the Kraus operators $\boldsymbol{M}_{1,2,\ldots}$ describes a possible "quantum jump" that the system might undergo, which occurs during time interval dt with probability $O(dt)$, and $\boldsymbol{M}_0$ describes how the system evolves when no jump occurs. We may write

$$\begin{aligned} \boldsymbol{M}_0 &= \boldsymbol{I} + (-i\boldsymbol{H} + \boldsymbol{K})dt, \\ \boldsymbol{M}_a &= \sqrt{dt}\ \boldsymbol{L}_a, \end{aligned} \tag{3.130}$$

where $\boldsymbol{H}$ and $\boldsymbol{K}$ are both hermitian and $\boldsymbol{L}_a, \boldsymbol{H}$, and $\boldsymbol{K}$ are all zeroth order in dt. In fact, we can determine $\boldsymbol{K}$ by invoking the Kraus-operator completeness relation; keeping terms up to linear order in $O(dt)$, we find

$$\boldsymbol{I} = \sum_a \boldsymbol{M}_a^\dagger \boldsymbol{M}_a = \boldsymbol{I} + dt\left(2\boldsymbol{K} + \sum_{a>0} \boldsymbol{L}_a^\dagger \boldsymbol{L}_a\right) + \cdots, \tag{3.131}$$

or

$$\boldsymbol{K} = -\frac{1}{2}\sum_{a>0} \boldsymbol{L}_a^\dagger \boldsymbol{L}_a. \tag{3.132}$$

Substituting into eq. (3.129), we obtain the *Lindblad master equation*:

$$\dot{\boldsymbol{\rho}} = \mathcal{L}(\boldsymbol{\rho}) = -i[\boldsymbol{H}, \boldsymbol{\rho}] + \sum_{a>0} \left(\boldsymbol{L}_a \boldsymbol{\rho} \boldsymbol{L}_a^\dagger - \frac{1}{2} \boldsymbol{L}_a^\dagger \boldsymbol{L}_a \boldsymbol{\rho} - \frac{1}{2} \boldsymbol{\rho} \boldsymbol{L}_a^\dagger \boldsymbol{L}_a \right). \quad (3.133)$$

This is the general Markovian evolution law for quantum states in the Schrödinger picture, assuming time evolution is a trace-preserving completely positive linear map. The first term in $\mathcal{L}(\boldsymbol{\rho})$ is the familiar Hamiltonian term generating unitary evolution. The other terms describe the possible transitions that the system may undergo due to interactions with the environment. The operators $\boldsymbol{L}_a$ are called *Lindblad operators* or *quantum jump operators*. Each $\boldsymbol{L}_a \boldsymbol{\rho} \boldsymbol{L}_a^\dagger$ term induces one of the possible quantum jumps, while the terms $-1/2 \boldsymbol{L}_a^\dagger \boldsymbol{L}_a \boldsymbol{\rho} - 1/2 \boldsymbol{\rho} \boldsymbol{L}_a^\dagger \boldsymbol{L}_a$ are needed to normalize properly the case in which no jumps occur.

Alternatively, we can describe the evolution using the Heisenberg picture. Then instead of eq.(3.129), the density operator is time-independent, while an operator $\boldsymbol{A}$ evolves according to

$$\boldsymbol{A}(t + dt) = \mathcal{E}_{dt}^*(\boldsymbol{A}(t)) = \sum_a \boldsymbol{M}_a^\dagger \boldsymbol{A}(t) \boldsymbol{M}_a, \quad (3.134)$$

and hence

$$\dot{\boldsymbol{A}} = \mathcal{L}^*(\boldsymbol{A}) = i[\boldsymbol{H}, \boldsymbol{A}] + \sum_{a>0} \left(\boldsymbol{L}_a^\dagger \boldsymbol{A} \boldsymbol{L}_a - \frac{1}{2} \boldsymbol{L}_a^\dagger \boldsymbol{L}_a \boldsymbol{A} - \frac{1}{2} \boldsymbol{A} \boldsymbol{L}_a^\dagger \boldsymbol{L}_a \right). \quad (3.135)$$

Heisenberg-picture time evolution is unital rather than trace preserving; the identity operator $\boldsymbol{I}$ does not evolve.

As for any nonunitary quantum channel, we have the freedom to redefine the Kraus operators in the operator-sum representation of $\mathcal{E}_{dt}$, replacing $\{\boldsymbol{M}_a\}$ by operators $\{\boldsymbol{N}_\mu\}$ which differ by a unitary change of basis. In particular, invoking this freedom for the jump operators (while leaving $\boldsymbol{M}_0$ untouched), we may replace $\{\boldsymbol{L}_a\}$ by $\{\boldsymbol{L}'_\mu\}$ where

$$\boldsymbol{L}'_\mu = \sum_a V_{\mu a} \boldsymbol{L}_a \quad (3.136)$$

and $V_{\mu a}$ is a unitary matrix. We say that these two ways of choosing the jump operators are two different *unravelings* of the same Markovian dynamics.

The master equation describes what happens when the system interacts with an unobserved environment, but we may also consider what happens if the environment is continuously monitored. In that case each quantum jump is detected; we update the quantum state of the system whenever

a jump occurs, and an initial pure state remains pure at all later times. Specifically, a jump of type a occurs during the interval $(t, t+dt)$ with probability

$$\text{Prob}(a) = dt\langle\psi(t)|\boldsymbol{L}_a^\dagger \boldsymbol{L}_a|\psi(t)\rangle, \tag{3.137}$$

and when a type-a jump is detected the updated state is

$$|\psi(t+dt)\rangle = \frac{\boldsymbol{L}_a|\psi(t)\rangle}{\|\boldsymbol{L}_a|\psi(t)\rangle\|}, \tag{3.138}$$

while when no jump occurs the state evolves as

$$|\psi(t+dt)\rangle = \frac{\boldsymbol{M}_0|\psi(t)\rangle}{\|\boldsymbol{M}_0|\psi(t)\rangle\|}. \tag{3.139}$$

This *stochastic Schrödinger evolution* can be numerically simulated; each simulated *quantum trajectory* is different, but averaging over a sample of many such trajectories reproduces the evolution of the density operator as described by the master equation. Simulating the stochastic Schrödinger equation may have advantages over simulating the master equation, since it is less costly to follow the evolution of a d-dimensional state vector than a $d \times d$ density matrix.

3.5.3 Damped harmonic oscillator

As an example to illustrate the master equation, consider the case of a harmonic oscillator coupled to the electromagnetic field via

$$\boldsymbol{H}' = \sum_k g_k(\boldsymbol{a}\boldsymbol{b}_k^\dagger + \boldsymbol{a}^\dagger \boldsymbol{b}_k), \tag{3.140}$$

where $\boldsymbol{a}$ is the annihilation operator of the oscillator, $\boldsymbol{b}_k^\dagger$ creates a photon in mode k, and g_k is a coupling constant. Let's also suppose that the environment is at zero temperature; then the excitation level of the oscillator can cascade down by successive emission of photons, but no absorption of photons will occur. If each photon, once emitted, never interacts again with the oscillator, the evolution is Markovian, and there is only one Lindblad jump operator:

$$\boldsymbol{L} = \sqrt{\Gamma}\boldsymbol{a}. \tag{3.141}$$

Here Γ is the rate for the oscillator to decay from the first excited ($n = 1$) state to the ground ($n = 0$) state, which can be computed as $\Gamma = \sum_k \Gamma_k$, where Γ_k is the rate for emission into mode k. The rate for the decay from level n to $n-1$ is $n\Gamma$. (The nth level of excitation of the oscillator

may be interpreted as a state of n noninteracting particles; the rate is $n\Gamma$ because any one of the n particles can decay.)

The Schrödinger-picture master equation in the Lindblad form becomes

$$\dot{\boldsymbol{\rho}} = -i[\boldsymbol{H}_0, \boldsymbol{\rho}] + \Gamma\left(\boldsymbol{a}\boldsymbol{\rho}\boldsymbol{a}^\dagger - \frac{1}{2}\boldsymbol{a}^\dagger\boldsymbol{a}\boldsymbol{\rho} - \frac{1}{2}\boldsymbol{\rho}\boldsymbol{a}^\dagger\boldsymbol{a}\right), \tag{3.142}$$

where $\boldsymbol{H}_0 = \omega\boldsymbol{a}^\dagger\boldsymbol{a}$ is the Hamiltonian of the oscillator, while in the Heisenberg picture an operator $\boldsymbol{A}$ evolves according to

$$\dot{\boldsymbol{A}} = i[\boldsymbol{H}_0, \boldsymbol{A}] + \Gamma\left(\boldsymbol{a}^\dagger\boldsymbol{A}\boldsymbol{a} - \frac{1}{2}\boldsymbol{a}^\dagger\boldsymbol{a}\boldsymbol{A} - \frac{1}{2}\boldsymbol{A}\boldsymbol{a}^\dagger\boldsymbol{a}\right). \tag{3.143}$$

The jump term describes the *damping* of the oscillator due to photon emission.

To study the effect of the jumps, consider the Heisenberg-picture evolution of the annihilation operator $\boldsymbol{a}(t)$. We can solve the master equation by making an ansatz $\boldsymbol{a}(t) = f(t)\boldsymbol{a}$, where $f(t)$ is a function of time, and then checking the self-consistency of the ansatz. Plugging into eq.(3.143) we find

$$\begin{aligned}\dot{f}(t)\boldsymbol{a} &= i\omega f(t)[\boldsymbol{a}^\dagger\boldsymbol{a}, \boldsymbol{a}] + \Gamma f(t)\left(\boldsymbol{a}^\dagger\boldsymbol{a}\boldsymbol{a} - \frac{1}{2}\boldsymbol{a}^\dagger\boldsymbol{a}\boldsymbol{a} - \frac{1}{2}\boldsymbol{a}\boldsymbol{a}^\dagger\boldsymbol{a}\right) \\ &= \left(i\omega + \frac{\Gamma}{2}\right) f(t)[\boldsymbol{a}^\dagger, \boldsymbol{a}]\boldsymbol{a} \implies \dot{f}(t) = -\left(i\omega + \frac{\Gamma}{2}\right) f(t),\end{aligned} \tag{3.144}$$

which integrates to $f(t) = e^{-i\omega t - \Gamma t/2} f(0)$. We conclude that

$$\boldsymbol{a}(t) = e^{-i\omega t - \Gamma t/2}\boldsymbol{a}(0), \tag{3.145}$$

and therefore the occupation number of the oscillator $\boldsymbol{n} = \boldsymbol{a}^\dagger\boldsymbol{a}$ decays in the Heisenberg picture according to

$$\boldsymbol{n}(t) = e^{-\Gamma t}\boldsymbol{n}(0). \tag{3.146}$$

Thus Γ is indeed the damping rate of the oscillator. If we interpret the nth excitation state of the oscillator as a state of n noninteracting particles, each with a decay probability Γ per unit time, then eq. (3.146) is just the exponential law satisfied by the population of decaying particles.

More interesting is what the master equation tells us about decoherence. In our amplitude damping model, it is the annihilation operator $\boldsymbol{a}$ and its adjoint that appear in the coupling $\boldsymbol{H}'$ of oscillator to environment, so we can anticipate that the oscillator's state will decohere in the basis of $\boldsymbol{a}$ eigenstates. The *coherent state*

$$|\alpha\rangle = e^{-|\alpha|^2/2} e^{\alpha a^\dagger}|0\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle, \tag{3.147}$$

is the normalized eigenstate of $\boldsymbol{a}$ with complex eigenvalue α. The operator $\boldsymbol{a}$ is not Hermitian, and two coherent states with distinct eigenvalues α and β are not orthogonal; rather

$$\begin{aligned} |\langle\alpha|\beta\rangle|^2 &= e^{-|\alpha|^2} e^{-|\beta|^2} e^{2Re(\alpha^*\beta)} \\ &= \exp(-|\alpha-\beta|^2), \end{aligned} \tag{3.148}$$

so the overlap is very small when $|\alpha - \beta|$ is large.

The solution to the Schrödinger-picture master equation is worked out in Exercise 3.9, where we find that an initial coherent state remains coherent, but with a decaying amplitude; after time t the state $|\alpha\rangle$ evolves as

$$|\alpha\rangle \mapsto |\alpha e^{-\Gamma t/2}\rangle \tag{3.149}$$

(in the rotating frame where the Hamiltonian evolution of the oscillator has been transformed away). We may also consider what happens when the initial state is a superposition of coherent states (a "cat state")

$$|\psi\rangle = N_{\alpha,\beta}(|\alpha\rangle + |\beta\rangle), \tag{3.150}$$

(where $N_{\alpha,\beta}$ is a normalization factor), or

$$\boldsymbol{\rho} = N_{\alpha,\beta}^2 \left(|\alpha\rangle\langle\alpha| + |\alpha\rangle\langle\beta| + |\beta\rangle\langle\alpha| + |\beta\rangle\langle\beta|\right). \tag{3.151}$$

The off-diagonal terms in this density operator evolve as

$$|\alpha\rangle\langle\beta| \mapsto e^{i\phi(\alpha,\beta)} e^{-\Gamma t|\alpha-\beta|^2/2} |\alpha e^{-\Gamma t/2}\rangle\langle\beta e^{-\Gamma t/2}|, \tag{3.152}$$

where $e^{i\phi(\alpha,\beta)}$ is a phase factor. Thus the off-diagonal terms decay exponentially with time, at a rate

$$\Gamma_{\text{decohere}} = \frac{1}{2}\Gamma|\alpha-\beta|^2 \tag{3.153}$$

proportional to $|\alpha - \beta|^2$, the square of the separation of the two coherent states in phase space; the decoherence rate is much larger than the damping rate Γ for $|\alpha - \beta|^2 \gg 1$. This behavior is easy to interpret. The expectation value of the occupation number $\boldsymbol{n}$ in a coherent state is $\langle\boldsymbol{n}\rangle = \langle\alpha|\boldsymbol{a}^\dagger\boldsymbol{a}|\alpha\rangle = |\alpha|^2$. Therefore, if α, β have comparable moduli but significantly different phases (as for a superposition of minimum uncertainty wave packets centered at positions x and $-x$), the decoherence rate is comparable to $\Gamma\langle\boldsymbol{n}\rangle$, the rate for emission of a *single* photon. This rate is very large compared to the rate for a significant fraction of the oscillator energy to be dissipated, if n is large.

We can also consider an oscillator coupled to an environment with a nonzero temperature. Again, the decoherence rate is roughly the rate for a

single photon to be emitted or absorbed, but the rate may be much faster than at zero temperature. Because the photon modes with frequency comparable to the oscillator frequency ω have a thermal occupation number

$$n_\gamma \approx \frac{T}{\hbar\omega}, \tag{3.154}$$

(for $T \gg \hbar\omega$), the interaction rate is further enhanced by the factor n_γ. For an oscillator with energy $E = \hbar\omega n_{\rm osc}$, we have

$$\begin{aligned}\frac{\Gamma_{\rm decohere}}{\Gamma_{\rm damp}} &\sim n_{\rm osc} n_\gamma \sim \frac{E}{\hbar\omega}\frac{T}{\hbar\omega} \\ &\sim \frac{m\omega^2 x^2}{\hbar\omega}\frac{T}{\hbar\omega} \sim x^2\frac{mT}{\hbar^2} \sim \frac{x^2}{\lambda_T^2},\end{aligned} \tag{3.155}$$

where x is the amplitude of oscillation and λ_T is the thermal de Broglie wavelength of the oscillating object. For macroscopic objects, decoherence is really *fast*.

3.6 Non-Markovian noise

3.6.1 Gaussian phase noise

The master equation describes the evolution of a quantum system subject to Markovian noise, but in some experimental systems the Markovian approximation is not very accurate. In this section we will discuss some of the features of decoherence for a system subjected to non-Markovian noise.

As a simple example, consider a single qubit with energy eigenstates $|0\rangle$ and $|1\rangle$, where the energy splitting between the two states fluctuates. For example, the qubit could be a spin-$\frac{1}{2}$ particle in a magnetic field pointing along the z-axis, where the magnetic field is not perfectly controlled in the laboratory. The Hamiltonian for this system is

$$\boldsymbol{H} = -\frac{1}{2}\omega_{01}\boldsymbol{\sigma}_3 - \frac{1}{2}f(t)\boldsymbol{\sigma}_3 \ , \tag{3.156}$$

where $f(t)$ is the fluctuating component of the magnetic field. This is a model of *classical noise*, arising not because the system interacts with an unobserved environment, but rather because a term in the system's Hamiltonian fluctuates.

The function f is treated stochastically; that is, we consider an ensemble of possible functions $\{f\}$, each with an assigned probability weight $\mu(f)$. We imagine that the actual $f(t)$ in each run of the experiment is selected by sampling from this distribution, and predict the observed behavior of the system by averaging over the distribution $\mu(f)$. The model is

particularly simple because the unperturbed Hamiltonian $\boldsymbol{H}_0 = \frac{1}{2}\omega_{01}\boldsymbol{\sigma}_3$ commutes with the noise term $\boldsymbol{H}_f = \frac{1}{2}f(t)\boldsymbol{\sigma}_3$, and in fact we can transform $\boldsymbol{H}_0$ away by going to the interaction picture.

The fluctuations induce dephasing of the qubit in the energy eigenstate basis. To analyze the dephasing, we will make a further simplifying assumption, that the noise is *Gaussian*. Whether the noise is classical of quantum, this Gaussian approximation often applies in laboratory situations where the system is weakly coupled to many different fluctuating variables in the environment. We denote averaging over the distribution $\mu(f)$ by $[\cdot]$, and assume the distribution to be stationary with mean zero; that is $[f(t)] = 0$, and $[f(t)f(t')] = K(t-t')$ is a function only of the difference $t-t'$, which is called the *covariance* of the distribution. The Gaussian distribution can be characterized by its generating functional $Z[J]$, which can be expressed in terms of the covariance as

$$Z[J] \equiv \left[e^{\int dt J(t)f(t)}\right]_f = \exp\left(\frac{1}{2}\int dt dt' J(t)K(t-t')J(t')\right). \quad (3.157)$$

An initial density operator $\boldsymbol{\rho}(0)$ evolves in time T to

$$\boldsymbol{\rho}(T) = \left[\exp\left(i\int_o^T \frac{1}{2}f(t)\boldsymbol{\sigma}_3\right)\boldsymbol{\rho}(0)\exp\left(-i\int_o^T \frac{1}{2}f(t)\boldsymbol{\sigma}_3\right)\right]. \quad (3.158)$$

The energy eigenstates $|0\rangle\langle 0|$ or $|1\rangle\langle 1|$ are not affected, but using eq.(3.157) we see that the coefficients of the off-diagonal entries $|0\rangle\langle 1|$ and $|1\rangle\langle 0|$ decay by the factor

$$\begin{aligned}&\exp\left(-\frac{1}{2}\int_0^T dt\int_0^T dt' K(t-t')\right)\\ &= \exp\left(-\frac{1}{2}\int_0^T dt\int_0^T dt'\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}e^{-i\omega(t-t')}\tilde{K}(\omega)\right);\end{aligned} \quad (3.159)$$

here we have introduced the Fourier transform $\tilde{K}(\omega)$ of the covariance $K(t)$, which is said to be the *spectral density* or *power spectrum* of the noise. Doing the t and t' integrals we obtain

$$\exp\left(-\frac{1}{2}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\tilde{K}(\omega)W_T(\omega)\right) \quad (3.160)$$

where $W_T(\omega)$ is the smooth window function

$$W_T(\omega) = \left|\int_0^T dt\ e^{-i\omega t}\right|^2 = \frac{4}{\omega^2}\sin^2(\omega T/2), \quad (3.161)$$

which has most of its support on the interval $[0, 2\pi/T]$.

Assuming that $\tilde{K}(\omega = 0)$ is finite, we expect that for T sufficiently large, $\tilde{K}(\omega)$ can be regarded as approximately constant in the region where $W_T(\omega)$ is supported. Using $\int_{-\infty}^{\infty} dx \frac{\sin^2 x}{x^2} = \pi$, we then obtain the decay factor $e^{-\Gamma_2 T}$, where the dephasing rate Γ_2 is

$$\Gamma_2 = \tilde{K}(\omega = 0). \tag{3.162}$$

(Here we've assumed that $\tilde{K}(\omega)$ is continuous at $\omega = 0$ — otherwise we should average its limiting values as ω approaches zero from positive and negative values.) If the spectral density is flat ("white noise"), this formula for Γ_2 applies at any time T, but in general, the time scale for which dephasing can be described by a rate Γ_2 depends on the shape of the noise's spectral density. In effect, an experimentalist who measures the dephasing time $T_2 = \Gamma_2^{-1}$ of a qubit is probing the noise power at low frequency.

Crudely speaking, we expect $\tilde{K}(\omega)$ to be roughly constant in the interval $[0, \omega_c]$, where $\omega_c = 2\pi/\tau_c$, and τ_c is a characteristic "autocorrelation" or "memory" time of the noise. That is, τ_c is chosen so that the correlation function $K(t - t')$ is small for $|t - t'| \gg \tau_c$. Thus we see that one can speak of a "dephasing rate" Γ_2 (and a corresponding dephasing time $T_2 = \Gamma_2^{-1}$) if the evolution is sufficiently "coarse-grained" in time. For the purpose of describing evolution over a time period $T \gg \tau_c$, the window function $W_T(\omega)$ is mostly supported within the interval $[0, \omega_c]$ where $\tilde{K}(\omega)$ is approximately constant; therefore the non-Markovian noise model can be replaced by a corresponding effective Markovian model in which the memory of the fluctuations can be neglected, as in our analysis of dephasing in §3.4.2. But for $T \ll \tau_c$ such a description may not be applicable.

3.6.2 Spin echo

Strategies for mitigating the damaging effects of the noise become possible when the noise autocorrelation time τ_c is long compared to the time scale over which the experimentalist can manipulate the system. For example, when observing the dephasing of a spin evolving for time T, we may apply a fast pulse that flips the spin about the x-axis at time $T/2$. Then the effects of low-frequency phase noise during the second half of the evolution will tend to compensate for the effects of the phase noise during the first half. This trick is called the *spin echo* phenomenon.

If we use this trick, the damping factor applied to $|0\rangle\langle 1|$ is again given by

$$\exp\left(-\frac{1}{2}\int_{-\infty}^{\infty} \frac{dw}{2\pi} \tilde{K}(\omega) W_T(\omega)\right) \tag{3.163}$$

but with a modified window function

$$W_T(\omega) = \left| \int_0^T dt J(t) e^{i\omega t} \right|^2 , \tag{3.164}$$

where $J(t)$ is a modulating function that expresses the effect of the spin echo pulse sequence. For example, if we flip the spin at time $T/2$, then $J(t)$ is +1 in the interval $[0, T/2]$ and -1 in the interval $[T/2, T]$; therefore

$$\begin{aligned} W_T(w) &= \frac{1}{\omega^2} \left| 1 - 2e^{i\omega T/2} + e^{i\omega T} \right|^2 \\ &= \frac{1}{\omega^2} \left| \frac{1 - e^{i\omega T/2}}{1 + e^{i\omega T/2}} \left(1 - e^{i\omega T}\right) \right|^2 \\ &= \tan^2(\omega T/4) \cdot \frac{4}{\omega^2} \sin^2(\omega t/2). \end{aligned} \tag{3.165}$$

In effect, the spin echo modifies $\tilde{K}(\omega)$ by the multiplicative factor $\tan^2(\omega T/4)$, which suppresses the low frequency noise.

The suppression can be improved further by using more pulses. In practice, pulses have bounded strength and nonzero duration, which places limitations on the effectiveness of this strategy.

3.6.3 Qubits as noise spectrometers

Now let's consider a different model of classical noise, in which the fluctuating term does not commute with the unperturbed Hamiltonian:

$$H = -\frac{1}{2}\omega_{01}\boldsymbol{\sigma}_3 + f(t)\boldsymbol{\sigma}_1. \tag{3.166}$$

In this model the fluctuating field can induce transitions among the energy eigenstates, at a rate that can be computed using lowest-order interaction-picture perturbation theory if the noise is weak. The probability that a qubit prepared in the state $|1\rangle$ at time 0 is observed in the state $|0\rangle$ at time T, averaged over the fluctuating classical field, is

$$\begin{aligned} \text{Prob}(1 \to 0) &= \left[\left| -i \int_0^T dt \; f(t) e^{-i\omega_{01} t} \langle 0|\boldsymbol{\sigma}_1|1\rangle \right|^2 \right] \\ &= \int_0^T dt \int_0^T dt' \; e^{-i\omega_{01}(t-t')} \left[f(t) f(t') \right] \\ &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \tilde{K}(\omega) W_T(\omega - \omega_{01}) \; . \end{aligned} \tag{3.167}$$

This expression is similar to the formula eq.(3.160) for the off-diagonal term in the density operator obtained in the dephasing model, except that

now the center of the window function has been shifted to the frequency ω_{01} of the transition.

As before, if we consider the observation time T to be large compared to the autocorrelation time τ_c of the noise, then the support of the window function is narrow, and $\tilde{K}(\omega)$ is approximately constant in the window. Thus, after a suitable coarse-graining of the time evolution, we may identify a rate for the decay of the qubit

$$\Gamma_\downarrow = \tilde{K}(\omega = \omega_{01}). \tag{3.168}$$

Similarly, for the transition from ground state to excited state, we find

$$\Gamma_\uparrow = \tilde{K}(\omega = -\omega_{01}). \tag{3.169}$$

Thus negative frequency noise transfers energy from the noise reservoir to the qubit, exciting the qubit, while positive frequency noise transfers energy from qubit to the noise reservoir, returning the excited qubit to the ground state. (Dephasing of a qubit, in contrast, involves a negligible exchange of energy and therefore is controlled by low frequency noise.) We conclude that an experimentalist capable of varying the energy splitting ω_{01} and measuring the qubit's transition rate can determine how the noise power depends on the frequency.

For the case we have considered in which the noise source is classical, $f(t)$ and $f(t')$ are real commuting variables; therefore $K(t)$ is an even function of t and correspondingly $\tilde{K}(\omega)$ is an even function of ω. Classical noise is spectrally symmetric, and the rates for excitation and decay are equal.

On the other hand, noise driven by a quantized thermal "bath" can be spectrally asymmetric. When the qubit comes to thermal equilibrium with the bath, up and down transitions occur at equal rates. If p_0 denotes the probability that the qubit is in the ground state $|0\rangle$ and p_1 denotes the probability that the qubit is in the excited state $|1\rangle$, then in equilibrium

$$p_0\Gamma_\uparrow = p_1\Gamma_\downarrow \Rightarrow \frac{\tilde{K}(-\omega_{01})}{\tilde{K}(\omega_{01})} = \frac{p_1}{p_0} = e^{-\beta\omega_{01}}; \tag{3.170}$$

the ratio of noise strengths at positive and negative frequencies is given (for a thermal bath) by a Boltzmann factor; this property of the noise is called the Kubo-Martin-Schwinger (KMS) condition. The noise becomes classical in the high-temperature limit $\beta\omega_{01} \ll 1$, and is in the deeply quantum regime for $\beta\omega_{01} \gg 1$.

3.6.4 Spin-boson model at nonzero temperature

To turn our model of classical dephasing noise into a quantum model, we replace the stochastic classical field $f(t)$ by an operator acting on a

quantized bath. The noise will still be Gaussian if the bath is a system of harmonic oscillators, uncoupled to one another and each coupled linearly to the dephasing qubit. The Hamiltonian for the system A and bath B is

$$\boldsymbol{H}_A + \boldsymbol{H}_B + \boldsymbol{H}_{AB} = -\frac{1}{2}\omega_{01}\boldsymbol{\sigma}_3 + \sum_k \omega_k \boldsymbol{a}_k^\dagger \boldsymbol{a}_k - \frac{1}{2}\boldsymbol{\sigma}_3 \left(\sum_k g_k \boldsymbol{a}_k + g_k^* \boldsymbol{a}_k^\dagger \right), \tag{3.171}$$

which is called the *spin-boson model*, as it describes a single spin-$\frac{1}{2}$ particle coupled to many bosonic variables. This is a model of dephasing because the coupling of the spin to the bath is diagonal in the spin's energy eigenstate basis. (Otherwise the physics of the model would be harder to analyze.) Despite its simplicity, the spin-boson model provides a reasonably realistic description of dephasing for a qubit weakly coupled to many degrees of freedom in the environment.

If there are many oscillators, the sum over k can be approximated by a frequency integral:

$$\sum_k |g_k|^2 \approx \int_0^\infty d\omega J(\omega), \tag{3.172}$$

where $J(\omega)$ is said to be the *spectral function* of the oscillator bath. Let's assume that the bath is in thermal equilibrium at temperature β^{-1}. In principle, the coupling to the system could tweak the equilibrium distribution of the bath, but we assume that this effect is negligible, because the bath is much bigger than the system. The fluctuations of the bath are Gaussian, and the average over the ensemble of classical functions in our previous analysis can be replaced by the thermal expectation value:

$$[\boldsymbol{f}(t)\boldsymbol{f}(0)] \mapsto \langle \boldsymbol{f}(t)\boldsymbol{f}(0)\rangle_\beta \equiv \mathrm{tr}\left(e^{-\beta \boldsymbol{H}_B} \boldsymbol{f}(t)\boldsymbol{f}(0) \right), \tag{3.173}$$

where now $\boldsymbol{f}(t)$ denotes the operator

$$\boldsymbol{f}(t) = e^{it\boldsymbol{H}_B} \boldsymbol{f}(0) e^{-it\boldsymbol{H}_B} = \sum_k \left(g_k \boldsymbol{a}_k e^{-i\omega_k t} + g_k^* \boldsymbol{a}_k^\dagger e^{i\omega_k t} \right). \tag{3.174}$$

We see that

$$K_\beta(t) \equiv \langle \boldsymbol{f}(t)\boldsymbol{f}(0)\rangle_\beta = \sum_k |g_k|^2 \langle e^{-i\omega_k t} \boldsymbol{a}_k \boldsymbol{a}_k^\dagger + e^{i\omega_k t} \boldsymbol{a}_k^\dagger \boldsymbol{a}_k \rangle_\beta. \tag{3.175}$$

From the Planck distribution, we find

$$\begin{aligned} \langle \boldsymbol{a}_k^\dagger \boldsymbol{a}_k \rangle_\beta &= \frac{1}{e^{\beta\omega_k} - 1} = \frac{1}{2}\coth(\beta\omega_k/2) - \frac{1}{2}, \\ \langle \boldsymbol{a}_k \boldsymbol{a}_k^\dagger \rangle_\beta &= \langle \boldsymbol{a}_k^\dagger \boldsymbol{a}_k + 1 \rangle_\beta = \frac{1}{2}\coth(\beta\omega_k/2) + \frac{1}{2}, \end{aligned} \tag{3.176}$$

and by Fourier transforming we obtain the spectral density of the noise

$$\begin{aligned}\tilde{K}_\beta(\omega) &\equiv \int_{-\infty}^{\infty} dt\ e^{i\omega t} K_\beta(t) \\ &= \sum_k |g_k|^2 \left(2\pi\delta(\omega-\omega_k)\langle \boldsymbol{a}_k \boldsymbol{a}_k^\dagger\rangle_\beta + 2\pi\delta(\omega+\omega_k)\langle \boldsymbol{a}_k^\dagger \boldsymbol{a}_k\rangle_\beta\right),\end{aligned} \tag{3.177}$$

which may be written as

$$\begin{aligned}\tilde{K}_\beta(\omega) &= \pi J(\omega)\left(\coth(\beta\omega/2)+1\right), \quad \omega>0, \\ \tilde{K}_\beta(\omega) &= \pi J(\omega)\left(\coth(\beta\omega/2)-1\right), \quad \omega<0.\end{aligned} \tag{3.178}$$

Thus, as we anticipated, the noise power spectrum exhibits the spectral asymmetry required by the KMS condition — the spectral density $\tilde{K}_\beta(-\omega)$ of the noise at negative frequency is supressed relative to the spectral density $\tilde{K}_\beta(\omega)$ at positive frequency by the Boltzmann factor $e^{-\beta\omega}$.

Since the window function $W_T(\omega)$ is an even function of ω, only the even part of $\tilde{K}_\beta(\omega)$ contributes to the attenuation of $|0\rangle\langle 1|$; the attenuation factor

$$\exp\left(-\frac{1}{2}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\tilde{K}_\beta(\omega)W_T(\omega)\right), \tag{3.179}$$

therefore becomes

$$\exp\left(-\int_0^\infty d\omega J(\omega)\frac{2\sin^2(\omega T/2)}{\omega^2}\coth(\beta\omega/2)\right). \tag{3.180}$$

A dephasing *rate* can be identified if the spectral function $J(\omega)$ behaves suitably at low frequency; the attenuation factor is $e^{-\Gamma_2 T}$ in the limit $T\to\infty$ where

$$\Gamma_2 = \lim_{\omega\to 0}\tilde{K}_\beta(\omega) = \lim_{\omega\to 0} 2\pi J(\omega)/(\beta\omega), \tag{3.181}$$

assuming that this limit exists. The noise is said to be *Ohmic* if $J(\omega)\approx A\omega$ is linear in ω at low frequency, and in that case the dephasing rate becomes $\Gamma_2 = 2\pi A\beta^{-1}$ in the limit of long time T.

3.7 Summary

POVM. If we restrict our attention to a subspace of a larger Hilbert space, then an orthogonal (Von Neumann) measurement performed on the larger space cannot in general be described as an orthogonal measurement

on the subspace. Rather, it is a *generalized measurement* or *POVM* — the outcome a occurs with a probability

$$\mathrm{Prob}(a) = \mathrm{tr}\,(\boldsymbol{E}_a \boldsymbol{\rho}) \ , \tag{3.182}$$

where $\boldsymbol{\rho}$ is the density matrix of the subsystem, each $\boldsymbol{E}_a$ is a positive hermitian operator, and the $\boldsymbol{E}_a$'s satisfy

$$\sum_a \boldsymbol{E}_a = \boldsymbol{I} \ . \tag{3.183}$$

A POVM in $\mathcal{H}_A$ can be realized as a unitary transformation on the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$, followed by an orthogonal measurement in $\mathcal{H}_B$.

Quantum channel. Unitary evolution on $\mathcal{H}_A \otimes \mathcal{H}_B$ will not in general appear to be unitary if we restrict our attention to $\mathcal{H}_A$ alone. Rather, evolution in $\mathcal{H}_A$ will be described by a *quantum channel*, (which can be inverted by another channel only if unitary). A general channel $\mathcal{E}$ has an operator-sum representation:

$$\mathcal{E} : \boldsymbol{\rho} \to \mathcal{E}(\boldsymbol{\rho}) = \sum_a \boldsymbol{M}_a \boldsymbol{\rho} \boldsymbol{M}_a^\dagger \ , \tag{3.184}$$

where

$$\sum_a \boldsymbol{M}_a^\dagger \boldsymbol{M}_a = \boldsymbol{I}. \tag{3.185}$$

In fact, any reasonable (linear, trace preserving, and completely positive) mapping of density operators to density operators has such an operator-sum representation.

Decoherence. Decoherence — the decay of quantum information due to the interaction of a system with its environment — can be described by a quantum channel. If the environment frequently "scatters" off the system, and the state of the environment is not monitored, then off-diagonal terms in the density operator of the system decay rapidly in a preferred basis (typically a spatially localized basis selected by the nature of the coupling of the system to the environment). The time scale for decoherence is set by the scattering rate, which may be much larger than the damping rate for the system.

Master Equation. When the relevant dynamical time scale of an open quantum system is long compared to the time for the environment to "forget" quantum information, the evolution of the system is effectively local in time (the Markovian approximation). Much as general unitary evolution is generated by a Hamiltonian, a general Markovian superoperator is generated by a *Liouvillian* $\mathcal{L}$ as described by the *master equation*:

$$\dot{\boldsymbol{\rho}} \equiv \mathcal{L}(\boldsymbol{\rho}) = -i[\boldsymbol{H}, \boldsymbol{\rho}] + \sum_a \left(\boldsymbol{L}_a \boldsymbol{\rho} \boldsymbol{L}_a^\dagger - \frac{1}{2} \boldsymbol{L}_a^\dagger \boldsymbol{L}_a \boldsymbol{\rho} - \frac{1}{2} \boldsymbol{\rho} \boldsymbol{L}_a^\dagger \boldsymbol{L}_a \right). \tag{3.186}$$

Here each *Lindblad operator* (or *quantum jump operator*) $\mathcal{L}_a$ describes a "quantum jump" that could in principle be detected if we monitored the environment faithfully. By solving the master equation, we can compute the decoherence rate of an open system.

Non-Markovian noise. Non-Markovian noise can be characterized by its power spectrum, and the effects of the noise on dephasing over a long time period are determined by the behavior of the power spectrum at low frequency. Quantum noise in thermal equilibrium at temperature β^{-1} has a spectral asymmetry — the noise at negative frequency $(-\omega)$ is suppressed compared to the noise at positive frequency ω by a Boltzmann factor $e^{-\beta\omega}$ (the *KMS condition*).

Further important ideas are developed in the Exercises.

3.8 Exercises

3.1 Which state did Alice make?

Consider a game in which Alice prepares one of two possible states: either ρ_1 with *a priori* probability p_1, or ρ_2 with *a priori* probability $p_2 = 1 - p_1$. Bob is to perform a measurement and on the basis of the outcome to guess which state Alice prepared. If Bob's guess is right, he wins; if he guesses wrong, Alice wins.

In this exercise you will find Bob's best strategy, and determine his optimal probability of error.

Let's suppose (for now) that Bob performs a POVM with two possible outcomes, corresponding to the two nonnegative Hermitian operators $\boldsymbol{E}_1$ and $\boldsymbol{E}_2 = \boldsymbol{I} - \boldsymbol{E}_1$. If Bob's outcome is $\boldsymbol{E}_1$, he guesses that Alice's state was ρ_1, and if it is $\boldsymbol{E}_2$, he guesses ρ_2. Then the probability that Bob guesses wrong is

$$p_{\text{error}} = p_1 \text{ tr } (\rho_1 \boldsymbol{E}_2) + p_2 \text{ tr } (\rho_2 \boldsymbol{E}_1) \ . \tag{3.187}$$

a) Show that

$$p_{\text{error}} = p_1 + \sum_i \lambda_i \langle i|\boldsymbol{E}_1|i\rangle \ , \tag{3.188}$$

where $\{|i\rangle\}$ denotes the orthonormal basis of eigenstates of the Hermitian operator $p_2\boldsymbol{\rho}_2 - p_1\boldsymbol{\rho}_1$, and the λ_i's are the corresponding eigenvalues.

b) Bob's best strategy is to perform the two-outcome POVM that minimizes this error probability. Find the nonnegative operator $\boldsymbol{E}_1$ that minimizes p_{error}, and show that error probability

when Bob performs this optimal two-outcome POVM is

$$(p_{\text{error}})_{\text{optimal}} = p_1 + \sum_{\text{neg}} \lambda_i \ . \tag{3.189}$$

where $\sum_{\text{neg}}$ denotes the sum over all of the *negative* eigenvalues of $p_2\boldsymbol{\rho}_2 - p_1\boldsymbol{\rho}_1$.

c) It is convenient to express this optimal error probability in terms of the L^1 norm of the operator $p_2\boldsymbol{\rho}_2 - p_1\boldsymbol{\rho}_1$,

$$\|p_2\boldsymbol{\rho}_2 - p_1\boldsymbol{\rho}_1\|_1 = \text{tr } |p_2\boldsymbol{\rho}_2 - p_1\boldsymbol{\rho}_1| = \sum_{\text{pos}} \lambda_i - \sum_{\text{neg}} \lambda_i \ , \tag{3.190}$$

the difference between the sum of positive eigenvalues and the sum of negative eigenvalues. Use the property tr $(p_2\boldsymbol{\rho}_2 - p_1\boldsymbol{\rho}_1) = p_2 - p_1$ to show that

$$(p_{\text{error}})_{\text{optimal}} = \frac{1}{2} - \frac{1}{2}\|p_2\boldsymbol{\rho}_2 - p_1\boldsymbol{\rho}_1\|_1 \ . \tag{3.191}$$

Check whether the answer makes sense in the case where $\boldsymbol{\rho}_1 = \boldsymbol{\rho}_2$ and in the case where $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$ have support on orthogonal subspaces.

d) Now suppose that Alice decides at random (with $p_1 = p_2 = 1/2$) to prepare one of two pure states $|\psi_1\rangle, |\psi_2\rangle$ of a single qubit, with

$$|\langle\psi_1|\psi_2\rangle| = \sin(2\alpha) \ , \quad 0 \le \alpha \le \pi/4 \ . \tag{3.192}$$

With a suitable choice of basis, the two states can be expressed as

$$|\psi_1\rangle = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \ , \qquad |\psi_2\rangle = \begin{pmatrix} \sin\alpha \\ \cos\alpha \end{pmatrix} \ . \tag{3.193}$$

Find Bob's optimal two-outcome measurement, and compute the optimal error probability.

e) Bob wonders whether he can find a better strategy if his POVM $\{\boldsymbol{E}_i\}$ has more than two possible outcomes. Let $p(a|i)$ denote the probability that state a was prepared, given that the measurement outcome was i; it can be computed using the relations

$$\begin{aligned} p_i p(1|i) &= p_1 p(i|1) = p_1 \text{ tr } \boldsymbol{\rho}_1 \boldsymbol{E}_i \ , \\ p_i p(2|i) &= p_2 p(i|2) = p_2 \text{ tr } \boldsymbol{\rho}_2 \boldsymbol{E}_i \ ; \end{aligned} \tag{3.194}$$

here $p(i|a)$ denotes the probability that Bob finds measurement outcome i if Alice prepared the state $\boldsymbol{\rho}_a$, and p_i denotes the

probability that Bob finds measurement outcome i, averaged over Alice's choice of state. For each outcome i, Bob will make his decision according to which of the two quantities

$$p(1|i) \ , \qquad p(2|i) \tag{3.195}$$

is the larger; the probability that he makes a mistake is the smaller of these two quantities. This probability of error, given that Bob obtains outcome i, can be written as

$$p_{\text{error}}(i) = \min\left(p(1|i), p(2|i)\right) = \frac{1}{2} - \frac{1}{2}\left|p(2|i) - p(1|i)\right| \ . \tag{3.196}$$

Show that the probability of error, averaged over the measurement outcomes, is

$$p_{\text{error}} = \sum_i p_i \ p_{\text{error}}(i) = \frac{1}{2} - \frac{1}{2}\sum_i \left|\text{tr}\left(p_2\boldsymbol{\rho}_2 - p_1\boldsymbol{\rho}_1\right)\boldsymbol{E}_i\right| \ . \tag{3.197}$$

f) By expanding in terms of the basis of eigenstates of $p_2\boldsymbol{\rho}_2 - p_1\boldsymbol{\rho}_1$, show that

$$p_{\text{error}} \geq \frac{1}{2} - \frac{1}{2}\|p_2\boldsymbol{\rho}_2 - p_1\boldsymbol{\rho}_1\|_1 \ . \tag{3.198}$$

(**Hint**: Use the completeness property $\sum_i \boldsymbol{E}_i = \boldsymbol{I}$.) Since we have already shown that this bound can be saturated with a two-outcome POVM, the POVM with many outcomes is no better.

3.2 Eavesdropping and disturbance

Alice wants to send a message to Bob. Alice is equipped to prepare either one of the two states $|u\rangle$ or $|v\rangle$. These two states, in a suitable basis, can be expressed as

$$|u\rangle = \begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix} \ , \quad |v\rangle = \begin{pmatrix}\sin\alpha\\ \cos\alpha\end{pmatrix} \ , \tag{3.199}$$

where $0 < \alpha < \pi/4$. Suppose that Alice decides at random to send either $|u\rangle$ or $|v\rangle$ to Bob, and Bob is to make a measurement to determine what she sent. Since the two states are not orthogonal, Bob cannot distinguish the states perfectly.

a) Bob realizes that he can't expect to be able to identify Alice's qubit every time, so he settles for a procedure that is successful only some of the time. He performs a POVM with three possible outcomes: $\neg u$, $\neg v$, or DON'T KNOW. If he obtains the

result $\neg u$, he is certain that $|v\rangle$ was sent, and if he obtains $\neg v$, he is certain that $|u\rangle$ was sent. If the result is DON'T KNOW, then his measurement is inconclusive. This POVM is defined by the operators

$$\boldsymbol{E}_{\neg u} = A(\boldsymbol{I} - |u\rangle\langle u|) \ , \quad \boldsymbol{E}_{\neg v} = A(\boldsymbol{I} - |v\rangle\langle v|) \ ,$$
$$\boldsymbol{E}_{\mathrm{DK}} = (1-2A)\boldsymbol{I} + A\left(|u\rangle\langle u| + |v\rangle\langle v|\right) \ , \quad (3.200)$$

where A is a positive real number. How should Bob choose A to minimize the probability of the outcome DK, and what is this minimal DK probability (assuming that Alice chooses from $\{|u\rangle, |v\rangle\}$ equiprobably)? **Hint:** If A is too large, $\boldsymbol{E}_{\mathrm{DK}}$ will have negative eigenvalues, and Eq.(3.200) will not be a POVM.

b) Eve also wants to know what Alice is sending to Bob. Hoping that Alice and Bob won't notice, she intercepts each qubit that Alice sends, by performing an orthogonal measurement that projects onto the basis $\left\{\binom{1}{0}, \binom{0}{1}\right\}$. If she obtains the outcome $\binom{1}{0}$, she sends the state $|u\rangle$ on to Bob, and if she obtains the outcome $\binom{0}{1}$, she sends $|v\rangle$ on to Bob. Therefore each time Bob's POVM has a conclusive outcome, Eve knows with certainty what that outcome is. But Eve's tampering causes detectable errors; sometimes Bob obtains a "conclusive" outcome that actually differs from what Alice sent. What is the probability of such an error, when Bob's outcome is conclusive?

3.3 Minimal disturbance

Consider a game in which Alice decides at random (equiprobably) whether to prepare one of two possible pure states of a single qubit, either

$$|\psi\rangle = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \ , \qquad \text{or} \qquad |\tilde{\psi}\rangle = \begin{pmatrix} \sin\alpha \\ \cos\alpha \end{pmatrix} \ , \quad (3.201)$$

and sends the state to Bob. By performing an orthogonal measurement in the basis $\{|0\rangle, |1\rangle\}$, Bob can identify the state with minimal error probability

$$(p_{\mathrm{error}})_{\mathrm{optimal}} = \sin^2\alpha = \frac{1}{2}(1 - \sin\theta) \ , \quad (3.202)$$

where we have defined θ by

$$\langle\psi|\tilde{\psi}\rangle \equiv \cos\theta = \sin(2\alpha) \ . \quad (3.203)$$

But now let's suppose that Eve wants to *eavesdrop* on the state as it travels from Alice to Bob. Like Bob, she wishes to extract optimal information that distinguishes $|\psi\rangle$ from $|\tilde{\psi}\rangle$, and she also wants to minimize the disturbance introduced by her eavesdropping, so that Alice and Bob are not likely to notice that anything is amiss.

Eve realizes that the optimal POVM can be achieved by measurement operators

$$\boldsymbol{M}_0 = |\phi_0\rangle\langle 0| \ , \qquad \boldsymbol{M}_1 = |\phi_1\rangle\langle 1| \ , \tag{3.204}$$

where the vectors $|\phi_0\rangle$, and $|\phi_1\rangle$ are arbitrary. If Eve performs this measurement, then Bob receives the state

$$\boldsymbol{\rho}' = \cos^2\alpha|\phi_0\rangle\langle\phi_0| + \sin^2\alpha|\phi_1\rangle\langle\phi_1| \ , \tag{3.205}$$

if Alice sent $|\psi\rangle$, and the state

$$\tilde{\boldsymbol{\rho}}' = \sin^2\alpha|\phi_0\rangle\langle\phi_0| + \cos^2\alpha|\phi_1\rangle\langle\phi_1| \ , \tag{3.206}$$

if Alice sent $|\tilde{\psi}\rangle$.

Eve wants the average fidelity of the state received by Bob to be as large as possible. The quantity that she wants to minimize, which we will call the "disturbance" D, measures how close this average fidelity is to one:

$$D = 1 - \frac{1}{2}(F + \tilde{F}) \ , \tag{3.207}$$

where

$$F = \langle\psi|\boldsymbol{\rho}'|\psi\rangle \ , \qquad \tilde{F} = \langle\tilde{\psi}|\tilde{\boldsymbol{\rho}}'|\tilde{\psi}\rangle \ . \tag{3.208}$$

The purpose of this exercise is to examine how effectively Eve can reduce the disturbance by choosing her measurement operators properly.

a) Show that $F + \tilde{F}$ can be expressed as

$$F + \tilde{F} = \langle\phi_0|\boldsymbol{A}|\phi_0\rangle + \langle\phi_1|\boldsymbol{B}|\phi_1\rangle \ , \tag{3.209}$$

where

$$\begin{aligned} \boldsymbol{A} &= \begin{pmatrix} 1 - 2\cos^2\alpha\sin^2\alpha & \cos\alpha\sin\alpha \\ \cos\alpha\sin\alpha & 2\cos^2\alpha\sin^2\alpha \end{pmatrix} \ , \\ \boldsymbol{B} &= \begin{pmatrix} 2\cos^2\alpha\sin^2\alpha & \cos\alpha\sin\alpha \\ \cos\alpha\sin\alpha & 1 - 2\cos^2\alpha\sin^2\alpha \end{pmatrix} \ . \end{aligned} \tag{3.210}$$

b) Show that if $|\phi_0\rangle$ and $|\phi_1\rangle$ are chosen optimally, the minimal disturbance that can be attained is

$$D_{\min}(\cos^2\theta) = \frac{1}{2}(1 - \sqrt{1 - \cos^2\theta + \cos^4\theta}) \ . \tag{3.211}$$

[**Hint**: We can choose $|\phi_0\rangle$ and $|\phi_1\rangle$ to maximize the two terms in eq. (3.209) independently. The maximal value is the maximal eigenvalue of $\boldsymbol{A}$, which since the eigenvalues sum to 1, can be expressed as $\lambda_{\max} = \frac{1}{2}\left(1 + \sqrt{1 - 4 \cdot \det \boldsymbol{A}}\right)$.] Of course, Eve could reduce the disturbance further were she willing to settle for a less than optimal probability of guessing Alice's state correctly.

c) Sketch a plot of the function $D_{\min}(\cos^2\theta)$. Interpret its value for $\cos\theta = 1$ and $\cos\theta = 0$. For what value of θ is $D_{\min}$ largest? Find $D_{\min}$ and $(p_{\text{error}})_{\text{optimal}}$ for this value of θ.

3.4 The price of quantum state encryption

Alice and Bob are working on a top secret project. I can't tell you exactly what the project is, but I will reveal that Alice and Bob are connected by a perfect quantum channel, and that Alice uses the channel to send quantum states to Bob. Alice and Bob are worried that an eavesdropper (Eve) might intercept some of Alice's transmissions. By measuring the intercepted quantum state, Eve could learn something about what Alice is sending, and perhaps make an inference about the nature of the project.

To protect against eavesdropping, Alice and Bob decide to *encrypt* the quantum states that Alice sends. They share a *secret key*, a string of random bits about which the eavesdropper knows nothing. By consuming $2n$ bits of secret key, Alice can encrypt, and Bob can decrypt, an arbitrary n-qubit state $\boldsymbol{\rho}$. For every possible state $\boldsymbol{\rho}$, the encrypted state looks exactly the same to Eve, so she cannot find out anything about $\boldsymbol{\rho}$.

Here is how the encryption procedure works: We may express the $2n$ bit string x as $x = x_0x_1x_2\cdots x_{n-1}$, where $x_i \in \{0, 1, 2, 3\}$, and denote a tensor product of n Pauli operators as

$$\boldsymbol{\sigma}(x) = \boldsymbol{\sigma}_{x_0} \otimes \boldsymbol{\sigma}_{x_2} \otimes \cdots \otimes \boldsymbol{\sigma}_{x_{n-1}} \tag{3.212}$$

(where $\boldsymbol{\sigma}_0 = \boldsymbol{I}$). Note that $\boldsymbol{\sigma}(x)^2 = \boldsymbol{I}^{\otimes n}$, the identity operator acting on n qubits. To encrypt, Alice consults her random string to determine x (which is chosen uniformly at random), and applies $\boldsymbol{\sigma}(x)$ to the state, obtaining $\boldsymbol{\sigma}(x)\boldsymbol{\rho}\boldsymbol{\sigma}(x)$. To decrypt, Bob, consults the same string and applies $\boldsymbol{\sigma}(x)$ to recover $\boldsymbol{\rho}$.

a) Since Eve does not know the secret key, to her the encrypted state is indistinguishable from

$$\mathcal{E}(\boldsymbol{\rho}) = \frac{1}{2^{2n}} \sum_x \boldsymbol{\sigma}(x) \boldsymbol{\rho} \boldsymbol{\sigma}(x) \ . \tag{3.213}$$

Show that, for any n-qubit state $\boldsymbol{\rho}$

$$\mathcal{E}(\boldsymbol{\rho}) = \frac{1}{2^n} \boldsymbol{I}^{\otimes n} \ . \tag{3.214}$$

Since $\mathcal{E}(\boldsymbol{\rho})$ is independent of $\boldsymbol{\rho}$, no information about $\boldsymbol{\rho}$ is accessible to Eve.

b) Alice wonders if it is possible to encrypt the state using a shorter key. Alice and Bob could use their shared randomness to sample an arbitrary probability distribution. That is, they could agree on a set of N unitary matrices $\{\boldsymbol{U}_a, a = 1, 2, 3, \ldots, N\}$, and Alice could encrypt by applying $\boldsymbol{U}_a$ with probability p_a. Then Bob could decrypt by applying $\boldsymbol{U}_a^{-1}$. To Eve, the encrypted state would then appear to be

$$\mathcal{E}'(\boldsymbol{\rho}) = \sum_a p_a \boldsymbol{U}_a \boldsymbol{\rho} \boldsymbol{U}_a^{-1} \ . \tag{3.215}$$

Show that, if $\mathcal{E}'(\boldsymbol{\rho}) = \boldsymbol{I}^{\otimes n}$, then $p_a \leq 2^{-2n}$ for each a.

Hint: Note that $\mathcal{E}$ has an operator sum representation with Kraus operators $\{\boldsymbol{\sigma}(x)/2^n\}$ and that $\mathcal{E}'$ has an operator sum representation with Kraus operators $\{\sqrt{p_a}\ \boldsymbol{U}_a\}$. Furthermore $\mathcal{E} = \mathcal{E}'$. Therefore, there exists an $M \times M$ unitary matrix V_{ax} (where $M = \max(N, 2^{2n})$) such that $\sqrt{p}_a \boldsymbol{U}_a = \sum_x V_{ax} \boldsymbol{\sigma}(x)/2^n$. Now express $p_a \mathrm{tr}\left(\boldsymbol{U}_a \boldsymbol{U}_a^\dagger\right)$ in terms of V.

Remark: The result shows that encryption requires $N \geq 2^{2n}$, and that at least $2n$ bits of key are required to specify $\boldsymbol{U}_a$. Thus the encryption scheme in which $\boldsymbol{\sigma}(x)$ is applied is the most efficient possible scheme. (For encryption to be effective, it is enough for $\mathcal{E}(\boldsymbol{\rho})$ to be independent of $\boldsymbol{\rho}$; it is not necessary that $\mathcal{E}(\boldsymbol{\rho}) = I^{\otimes n}/2^n$. But the same result applies under the weaker assumption that $\mathcal{E}(\boldsymbol{\rho})$ is independent of $\boldsymbol{\rho}$.)

3.5 Unital maps and majorization

Recall that the action of a trace-preserving completely positive (TPCP) map $\mathcal{E}$ can be expressed as

$$\mathcal{E}(\boldsymbol{\rho}) = \sum_a \boldsymbol{M}_a \boldsymbol{\rho} \boldsymbol{M}_a^\dagger \ , \tag{3.216}$$

where

$$\sum_a \boldsymbol{M}_a^\dagger \boldsymbol{M}_a = \boldsymbol{I} \ . \tag{3.217}$$

A TPCP map is said to be *unital* if $\mathcal{E}(\boldsymbol{I}) = \boldsymbol{I}$, or equivalently if

$$\sum_a \boldsymbol{M}_a \boldsymbol{M}_a^\dagger = \boldsymbol{I} \ . \tag{3.218}$$

If $\boldsymbol{A}$ is a nonnegative Hermitian operator with unit trace ($\mathrm{tr}\ \boldsymbol{A} = 1$), let $\lambda(\boldsymbol{A})$ denote the vector of eigenvalues of $\boldsymbol{A}$, which can be regarded as a probability vector. If $\boldsymbol{A}$ and $\boldsymbol{B}$ are nonnegative Hermitian operators with unit trace, we say that $\boldsymbol{A} \prec \boldsymbol{B}$ ("$\boldsymbol{A}$ is majorized by $\boldsymbol{B}$") if $\lambda(\boldsymbol{A}) \prec \lambda(\boldsymbol{B})$. (Recall that for two probability vectors p and q, we say that $p \prec q$ if there is a doubly stochastic matrix D such that $p = Dq$.)

Show that if $\boldsymbol{\rho}$ is a density operator and $\mathcal{E}$ is a unital map, then

$$\mathcal{E}(\boldsymbol{\rho}) \prec \boldsymbol{\rho} \ . \tag{3.219}$$

Hint: Express $\boldsymbol{\rho} = \boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{U}^\dagger$ where $\boldsymbol{\Delta}$ is diagonal and $\boldsymbol{U}$ is unitary, and express $\boldsymbol{\rho}' \equiv \mathcal{E}(\boldsymbol{\rho}) = \boldsymbol{V}\boldsymbol{\Delta}'\boldsymbol{V}^\dagger$, where $\boldsymbol{\Delta}'$ is diagonal and $\boldsymbol{V}$ is unitary. Then try to show that the diagonal entries of $\boldsymbol{\Delta}'$ can be expressed as a doubly stochastic matrix acting on the diagonal entries of $\boldsymbol{\Delta}$.

Remark: A unital map is the natural quantum generalization of a doubly stochastic map (a doubly stochastic map can be regarded as the special case of a unital map that preserves the basis in which $\boldsymbol{\rho}$ is diagonal). The result of the exercise shows that a unital map takes an input density operator to an output density operator that is no less random than the input.

3.6 What transformations are possible for bipartite pure states?

Alice and Bob share a bipartite pure state $|\Psi\rangle$. Using a 2-LOCC protocol, they wish to transform it to another bipartite pure state $|\Phi\rangle$. Furthermore, the protocol must be *deterministic* — the state $|\Phi\rangle$ is obtained with probability one irrespective of the outcomes of the measurements that Alice and Bob perform.

Suppose that these initial and final states have Schmidt decompositions

$$|\Psi\rangle = \sum_i \sqrt{(p_\Psi)_i}\ |\alpha_i\rangle \otimes |\beta_i\rangle \ , \quad |\Phi\rangle = \sum_i \sqrt{(p_\Phi)_i}\ |\alpha_i'\rangle \otimes |\beta_i'\rangle \ . \tag{3.220}$$

Show that if the deterministic transformation $|\Psi\rangle \mapsto |\Phi\rangle$ is possible, then $p_\Psi \prec p_\Phi$.

Hints: Using the Lo-Popescu Theorem from Exercise 2.9, we can reduce the 2-LOCC to an equivalent 1-LOCC. That is, if the deterministic transformation is possible, then there is a generalized measurement that can be applied by Alice, and an operation depending on Alice's measurement outcome that can be applied by Bob, such that for each possible measurement outcome Alice's measurement followed by Bob's operation maps $|\Psi\rangle$ to $|\Phi\rangle$. Recall that a generalized measurement is defined by a set of operators $\{\boldsymbol{M}_a\}$ such that $\sum_a \boldsymbol{M}_a^\dagger \boldsymbol{M}_a = \boldsymbol{I}$, and that the action of the measurement on a pure state $|\psi\rangle$ if outcome a occurs is

$$|\psi\rangle \mapsto \frac{\boldsymbol{M}_a|\psi\rangle}{\sqrt{\langle\psi|\boldsymbol{M}_a^\dagger \boldsymbol{M}_a|\psi\rangle}} \,. \tag{3.221}$$

Think about how the 1-LOCC protocol transforms Alice's density operator. You might want to use the *polar decomposition*: a matrix $\boldsymbol{A}$ can be expressed as $\sqrt{\boldsymbol{A}\boldsymbol{A}^\dagger}\,\boldsymbol{U}$, where $\boldsymbol{U}$ is unitary.

Remark: The converse is also true. Thus majorization provides the necessary and sufficient condition for the deterministic transformation of one bipartite pure state to another (*Nielsen's Theorem*). In this respect, majorization defines a partial order on bipartite pure states such that we may say that $|\Psi\rangle$ is no less entangled than $|\Phi\rangle$ if $p_\Psi \prec p_\Phi$.

3.7 Fidelity and overlap

The *overlap* of two probability distributions $\{p_i\}$ and $\{\tilde{p}_i\}$ is defined as

$$\text{Overlap}(\{p_i\}, \{\tilde{p}_i\}) \equiv \sum_i \sqrt{p_i \cdot \tilde{p}_i} \,. \tag{3.222}$$

Suppose that we try to distinguish the two states $\boldsymbol{\rho}$ and $\tilde{\boldsymbol{\rho}}$ by performing the POVM $\{\boldsymbol{E}_i\}$. Then the two corresponding probability distributions have the overlap

$$\text{Overlap}(\boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}; \{\boldsymbol{E}_i\}) \equiv \sum_i \sqrt{\text{tr } \boldsymbol{\rho}\boldsymbol{E}_i} \cdot \sqrt{\text{tr } \tilde{\boldsymbol{\rho}}\boldsymbol{E}_i} \,. \tag{3.223}$$

It turns out that the minimal overlap that can be achieved by any POVM is related to the fidelity $F(\boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}) = \left\| \tilde{\boldsymbol{\rho}}^{\frac{1}{2}} \boldsymbol{\rho}^{\frac{1}{2}} \right\|_1^2$:

$$\min_{\{\boldsymbol{E}_i\}} \left[\text{Overlap}(\boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}; \{\boldsymbol{E}_i\})\right] = \sqrt{F(\boldsymbol{\rho}, \tilde{\boldsymbol{\rho}})} \,. \tag{3.224}$$

In this exercise, you will show that the square root of the fidelity is a lower bound on the overlap, but not that the bound can be saturated.

b) The space of linear operators acting on a Hilbert space is itself a Hilbert space, where the inner product $(\boldsymbol{A}, \boldsymbol{B})$ of two operators $\boldsymbol{A}$ and $\boldsymbol{B}$ is

$$(\boldsymbol{A}, \boldsymbol{B}) \equiv \mathrm{tr}\left(\boldsymbol{A}^\dagger \boldsymbol{B}\right) \ . \tag{3.225}$$

For this inner product, the Schwarz inequality becomes

$$|\mathrm{tr}\ \boldsymbol{A}^\dagger \boldsymbol{B}| \leq \left(\mathrm{tr}\ \boldsymbol{A}^\dagger \boldsymbol{A}\right)^{1/2} \left(\mathrm{tr}\ \boldsymbol{B}^\dagger \boldsymbol{B}\right)^{1/2} \ , \tag{3.226}$$

Choosing $\boldsymbol{A} = \boldsymbol{\rho}^{\frac{1}{2}} \boldsymbol{E}_i^{\frac{1}{2}}$ and $\boldsymbol{B} = \boldsymbol{U}\tilde{\boldsymbol{\rho}}^{\frac{1}{2}} \boldsymbol{E}_i^{\frac{1}{2}}$ (for an arbitrary unitary $\boldsymbol{U}$), use this form of the Schwarz inequality to show that

$$\mathrm{Overlap}(\boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}; \{\boldsymbol{E}_i\}) \geq |\mathrm{tr}\ \boldsymbol{\rho}^{\frac{1}{2}} \boldsymbol{U} \tilde{\boldsymbol{\rho}}^{\frac{1}{2}}| \ . \tag{3.227}$$

c) Now use the polar decomposition

$$\boldsymbol{A} = \boldsymbol{V}\sqrt{\boldsymbol{A}^\dagger \boldsymbol{A}} \tag{3.228}$$

(where $\boldsymbol{V}$ is unitary) to write

$$\tilde{\boldsymbol{\rho}}^{\frac{1}{2}} \boldsymbol{\rho}^{\frac{1}{2}} = \boldsymbol{V}\sqrt{\boldsymbol{\rho}^{\frac{1}{2}} \tilde{\boldsymbol{\rho}} \boldsymbol{\rho}^{\frac{1}{2}}} \ , \tag{3.229}$$

and by choosing the unitary $\boldsymbol{U}$ in eq. (3.227) to be $\boldsymbol{U} = \boldsymbol{V}^{-1}$, show that

$$\mathrm{Overlap}(\boldsymbol{\rho}, \tilde{\boldsymbol{\rho}}; \{\boldsymbol{E}_i\}) \geq \sqrt{F(\boldsymbol{\rho}, \tilde{\boldsymbol{\rho}})} \ . \tag{3.230}$$

d) We can obtain an explicit formula for the fidelity in the case of two states of a single qubit. Using the Bloch parametrization

$$\boldsymbol{\rho}(\vec{P}) = \frac{1}{2}\left(I + \vec{\boldsymbol{\sigma}} \cdot \vec{P}\right) \ , \tag{3.231}$$

show that the fidelity of two single-qubit states with polarization vectors $\vec{P}$ and $\vec{Q}$ is

$$F(\vec{P}, \vec{Q}) = \frac{1}{2}\left(1 + \vec{P} \cdot \vec{Q} + \sqrt{(1 - \vec{P}^2)(1 - \vec{Q}^2)}\ \right) \ . \tag{3.232}$$

Hint: First note that the eigenvalues of a 2×2 matrix can be expressed in terms of the trace and determinant of the matrix. Then evaluate the determinant and trace of $\left(\boldsymbol{\rho}^{\frac{1}{2}} \tilde{\boldsymbol{\rho}} \boldsymbol{\rho}^{\frac{1}{2}}\right)$, and calculate the fidelity using the corresponding expression for the eigenvalues.

3.8 Semicausal and semilocal maps in the Heisenberg picture

In the Schrödinger picture, a completely positive (CP) map $\mathcal{E}$ leaves observables fixed and takes an input density operator to an output density operator, $\mathcal{E} : \boldsymbol{\rho}_{in} \mapsto \boldsymbol{\rho}_{out} = \mathcal{E}(\boldsymbol{\rho}_{in})$. In the Heisenberg picture, the dual map $\mathcal{E}^*$ leaves density operators fixed and takes an input observable to an output observable, $\mathcal{E}^* : \boldsymbol{a}_{in} \mapsto \boldsymbol{a}_{out} = \mathcal{E}^*(\boldsymbol{a}_{in})$. If $\mathcal{E}$ has the operator sum representation $\mathcal{E}(\boldsymbol{\rho}) = \sum_\mu \boldsymbol{M}_a \boldsymbol{\rho} \boldsymbol{M}_a^\dagger$, then its dual has operator sum representation

$$\mathcal{E}^*(\boldsymbol{a}) = \sum_a \boldsymbol{M}_a^\dagger \boldsymbol{a} \boldsymbol{M}_a \ . \tag{3.233}$$

a) If $\mathcal{E}$ is a TPCP map, show that its dual $\mathcal{E}^*$ can be represented as

$$\mathcal{E}^*(a) = {}_C\langle 0|\boldsymbol{U}_{AC}^\dagger \, (\boldsymbol{a}_A \otimes \boldsymbol{I}_C) \, \boldsymbol{U}_{AC}|0\rangle_C \ , \tag{3.234}$$

where $\boldsymbol{U}_{AC}$ is a unitary transformation on AC, $\boldsymbol{a}_A$ is an observable on A, $\boldsymbol{I}_C$ is the identity on C, and $|0\rangle_C$ is a fixed pure state in $\mathcal{H}_C$. (You may use the corresponding property of the TPCP map $\mathcal{E}$.)

b) Consider a CP map $\mathcal{E}$ acting on a bipartite quantum system AB. We way that $\mathcal{E}$ is *semicausal* if the map does not convey any information from B to A. That is, suppose that Alice and Bob share an initial state $\boldsymbol{\rho}_{AB}$. Then if Bob performs an operation on B before the map $\mathcal{E}$ acts, and Alice makes a measurement on A after the map $\mathcal{E}$ acts, Alice's measurement collects no information about the operation that Bob performed. Show that if $\mathcal{E}$ is semicausal, then there is an operation $\tilde{\mathcal{E}}$ on A such that

$$\mathcal{E}^*(\boldsymbol{a}_A \otimes \boldsymbol{I}_B) = \tilde{\mathcal{E}}^*(\boldsymbol{a}_A) \otimes \boldsymbol{I}_B \ . \tag{3.235}$$

c) We say that $\mathcal{E}$ is *semilocal* if it can be performed by means of local operations and one-way *quantum communication* from A to B. That is, there is a message system C that can be passed from Alice to Bob. We may assume that the initial state of ABC is a product $\boldsymbol{\rho}_{AB} \otimes \boldsymbol{\rho}_C$ — the state of the message is uncorrelated with the joint state held by Alice and Bob. To apply $\mathcal{E}$ to $\boldsymbol{\rho}_{AB}$, Alice applies an operation to AC, and sends C to Bob. Then Bob applies an operation to BC, and discards C. Show that if $\mathcal{E}$ is semilocal, then there are CP maps $\mathcal{G}_{AC}$ from A to AC and $\mathcal{F}_{BC}$ from BC to B such that

$$\mathcal{E}^* = (\mathcal{G}_{AC}^* \otimes I_B) \circ (I_A \otimes \mathcal{F}_{BC}^*) \ ; \tag{3.236}$$

here $\circ$ denotes composition of maps, with the map on the right acting first.

d) Using the Heisenberg-picture characterizations of semicausal and semilocal maps found in (*b*) and (*c*), show that a semilocal map is semicausal, and express $\tilde{\mathcal{E}}$ in terms of $\mathcal{F}$ and $\mathcal{G}$.

Remark. The result (*d*) is intuitively obvious — communication from Alice to Bob cannot convey a signal from Bob to Alice. What is less obvious is that the converse is also true: every semicausal map is semilocal.

3.9 Damped harmonic oscillator at zero temperature

Let's suppose the oscillations of a quantum harmonic oscillator with circular frequency ω are damped because the oscillator can emit photons with energy $\hbar\omega$. When a photon is emitted, the oscillator makes a transition from the energy eigenstate with energy $E_n = n\hbar\omega$ to the energy eigenstate with enenrgy $E_{n-1} = (n-1)\hbar\omega$, and the photon carries away the lost energy. The probability that a photon is emitted in an infinitesimal time interval dt is Γdt; we say that Γ is the emission rate. Therefore, the coupled evolution of the oscillator and the electromagnetic field for time interval dt can be described as:

$$\begin{aligned} &|\Psi(0)\rangle = |\psi\rangle \otimes |0\rangle \mapsto \\ &|\Psi(dt)\rangle = \sqrt{\Gamma dt}\ \boldsymbol{a}|\psi\rangle \otimes |1\rangle + \left(\boldsymbol{I} - \frac{1}{2}\Gamma dt\ \boldsymbol{a}^\dagger \boldsymbol{a}\right)|\psi\rangle \otimes |0\rangle. \end{aligned} \tag{3.237}$$

Here $|\psi\rangle$ is the initial normalized state vector of the oscillator and $\{|0\rangle, |1\rangle\}$ are orthonormal states of the electromagnetic field; $|0\rangle$ denotes the state in which no photon has been emitted and $|1\rangle$ denotes the state containing one photon. The operator $\boldsymbol{a}$ reduces the excitation level of the oscillator by one unit, and the $\boldsymbol{a}^\dagger\boldsymbol{a}$ factor in the second term is needed to ensure that the evolution is unitary.

a) Check unitarity by verifying that $\langle\Psi(dt)|\Psi(dt)\rangle = 1$, to linear order in the small quantity dt.

Because the states $\{|0\rangle, |1\rangle\}$ of the electromagnetic field are orthogonal, the quantum state of the oscillator may decohere. Summing over these basis states, we see that the initial pure state $|\psi\rangle\langle\psi|$ of the oscillator evolves in time dt as

$$\begin{aligned} &|\psi\rangle\langle\psi| \mapsto \langle 0|\Psi(dt)\rangle\langle\Psi(dt)|0\rangle + \langle 1|\Psi(dt)\rangle\langle\Psi(dt)|1\rangle \\ &= \Gamma dt\ \boldsymbol{a}|\psi\rangle\langle\psi|\boldsymbol{a}^\dagger + \left(\boldsymbol{I} - \frac{1}{2}\Gamma dt\ \boldsymbol{a}^\dagger\boldsymbol{a}\right)|\psi\rangle\langle\psi|\left(\boldsymbol{I} - \frac{1}{2}\Gamma dt\ \boldsymbol{a}^\dagger\boldsymbol{a}\right); \end{aligned}$$

More generally, the initial (not necessarily pure) density operator $\boldsymbol{\rho}$ of the oscillator evolves as

$$\boldsymbol{\rho} \mapsto \Gamma dt\ \boldsymbol{a}\boldsymbol{\rho}\boldsymbol{a}^\dagger + \left(\boldsymbol{I} - \frac{1}{2}\Gamma dt\ \boldsymbol{a}^\dagger\boldsymbol{a}\right)\boldsymbol{\rho}\left(\boldsymbol{I} - \frac{1}{2}\Gamma dt\ \boldsymbol{a}^\dagger\boldsymbol{a}\right). \quad (3.238)$$

Now suppose that the initial state of the oscillator is a coherent state

$$|\alpha\rangle = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle, \quad (3.239)$$

where α is a complex number. For this problem, we will ignore the usual dynamics of the oscillator that causes α to rotate uniformly in time: $\alpha \mapsto \alpha e^{-i\omega t}$; equivalently, we will assume that the dynamics is described in a "rotating frame" such that the rotation of α is transformed away. We will only be interested in how the states of the oscillator are affected by the damping described by eq.(3.238).

b) Show that, to linear order in dt,

$$\left(\boldsymbol{I} - \frac{1}{2}\Gamma dt\ \boldsymbol{a}^\dagger\boldsymbol{a}\right)|\alpha\rangle \approx e^{-\Gamma dt|\alpha|^2/2}|\alpha\ e^{-\Gamma dt/2}\rangle. \quad (3.240)$$

Note that there are two things to check in eq.(3.240): that the value of α decays with time, and that the normalization of the state decays with time.

c) Verify that, also to linear order in dt,

$$\Gamma dt\ \boldsymbol{a}|\alpha\rangle\langle\alpha|\boldsymbol{a}^\dagger \approx \Gamma dt|\alpha|^2\ |\alpha\ e^{-\Gamma dt/2}\rangle\langle\alpha\ e^{-\Gamma dt/2}|, \quad (3.241)$$

and thus show that, to linear order in dt, $|\alpha\rangle\langle\alpha|$ evolves as

$$|\alpha\rangle\langle\alpha| \mapsto |\alpha\ e^{-\Gamma dt/2}\rangle\langle\alpha\ e^{-\Gamma dt/2}|. \quad (3.242)$$

By considering many consecutive small time increments, argue that, in a finite time t, the initial coherent state evolves as

$$|\alpha\rangle \mapsto |\alpha\ e^{-\Gamma t/2}\rangle. \quad (3.243)$$

Thus, the state remains a (pure) coherent state at all times, with the value of α decaying exponentially with time. Since the energy stored in the oscillator is proportional to α^2, which decays like $e^{-\Gamma t}$, we say that Γ is the *damping rate* of the oscillator.

Now consider what happens if the initial state of the oscillator is a superposition of two coherent states:

$$|\psi\rangle = N_{\alpha,\beta}\left(|\alpha\rangle + |\beta\rangle\right). \tag{3.244}$$

Here $N_{\alpha,\beta}$ is a real nonnegative normalization constant (note that, though the states $|\alpha\rangle$ and $\beta\rangle$ are both normalized, they are not orthogonal).

d) Evaluate $\langle\beta|\alpha\rangle$, and determine $N_{\alpha,\beta}$.

For example we might choose $\alpha = \xi_0/\sqrt{2}$ and $\beta = -\xi_0/\sqrt{2}$, so that the two superposed coherent states are minimum uncertainty wavepackets (with width $\Delta\xi = 1/\sqrt{2}$) centered at dimensionless positions $\pm\xi_0$. If $|\alpha - \beta| \gg 1$, then the two wavepackets are well separated compared to their width, and we might say that oscillator state $|\psi\rangle$ is "in two places at once." How quickly will such a superposition of two separated wavepackets decohere?

The initial density operator of the oscillator is

$$\boldsymbol{\rho} = N_{\alpha,\beta}^2\big(|\alpha\rangle\langle\alpha| + |\alpha\rangle\langle\beta| + |\beta\rangle\langle\alpha| + |\beta\rangle\langle\beta|\big). \tag{3.245}$$

We already know from part (*c*) how the "diagonal" terms $|\alpha\rangle\langle\alpha|$ and $|\beta\rangle\langle\beta|$ evolve, but what about the "off-diagonal" terms $|\alpha\rangle\langle\beta|$ and $|\beta\rangle\langle\alpha|$?

e) Using arguments similar to those used in parts (*b*) and (*c*), show that in time t, the operator $|\alpha\rangle\langle\beta|$ evolves as

$$|\alpha\rangle\langle\beta| \mapsto e^{i\phi(\alpha,\beta)}e^{-\Gamma t|\alpha-\beta|^2/2}|\alpha e^{-\Gamma t/2}\rangle\langle\beta e^{-\Gamma t/2}|, \tag{3.246}$$

and find the phase factor $e^{i\phi(\alpha,\beta)}$. Thus the off-diagonal terms decay exponentially with time, at a rate

$$\Gamma_{\text{decohere}} = \frac{1}{2}\Gamma|\alpha - \beta|^2 \tag{3.247}$$

proportional to the distance squared $|\alpha - \beta|^2$.

f) Consider an oscillator with mass $m = 1\ g$, circular frequency $\omega = 1\ s^{-1}$ and (*very good*) quality factor $Q \equiv \omega/\Gamma = 10^9$. Thus the damping time is very long: over 30 years. A superposition of minimum uncertainty wavepackets is prepared, centered at positions $x = \pm 1\ cm$. Estimate the decoherence rate. (Wow! For macroscopic objects, decoherence is really *fast*!)

3.10 One-qubit decoherence

The matrices $\boldsymbol{I}, \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3$, where $\boldsymbol{\sigma}_{1,2,3}$ are the Pauli matrices and bfI is the identity matrix, are a basis for the four-dimensional space of 2×2 matrices. Let us denote $\boldsymbol{I}$ as $\boldsymbol{\sigma}_0$.

a) Let $\mathcal{E}$ be a quantum operation (a completely positive map) acting on the density operator $\boldsymbol{\rho}$ of a single qubit. Show that we may express $\mathcal{E}(\boldsymbol{\rho})$ as

$$\mathcal{E}(\boldsymbol{\rho}) = \sum_{\mu,\nu=0}^{3} \mathcal{E}_{\mu\nu}\ \boldsymbol{\sigma}_\mu \boldsymbol{\rho} \boldsymbol{\sigma}_\nu\ , \tag{3.248}$$

where the $\mathcal{E}_{\mu\nu}$'s are complex numbers satisfying $\mathcal{E}_{\mu\nu} = \mathcal{E}^*_{\nu\mu}$. **Hint**: The operation $\mathcal{E}$ has an operator-sum representation with operation elements $\{\boldsymbol{M}_a\}$. Each $\boldsymbol{M}_a$ can be expanded in the basis $\{\boldsymbol{\sigma}_\mu, \mu = 0, 1, 2, 3\}$.

b) Find four independent conditions that must be satisfied by the $\mathcal{E}_{\mu\nu}$'s in order that the operation $\mathcal{E}$ be trace-preserving (a *channel*).

c) A Hermitian 2×2 operator can be expressed as

$$\boldsymbol{\rho}(P) = \frac{1}{2} \sum_{\mu=0}^{3} P_\mu \boldsymbol{\sigma}_\mu\ , \tag{3.249}$$

where P_0, P_1, P_2, P_3 are real numbers. Show that a linear map that takes Hermitian operators to Hermitian operators acts as

$$\mathcal{E}(\boldsymbol{\rho}(P)) = \boldsymbol{\rho}(P')\ , \tag{3.250}$$

where $P' = MP$ and M is a real matrix. What is the (real) dimension of the space of such linear maps?

d) Suppose that tr $\boldsymbol{\rho} = 1$ and that $\mathcal{E}$ is trace preserving, so that $P_0 = P'_0 = 1$. Show that

$$\vec{P}' = M\vec{P} + \vec{v}\ , \tag{3.251}$$

where $\vec{P}$ and $\vec{P}'$ are real three-component polarization vectors, M is a real matrix, and $\vec{v}$ is a real three-component vector. What is the (real) dimension of the space of such trace-preserving maps?

e) Express $\vec{v}$ in terms of the $\mathcal{E}_{0k}$'s. **Hint**: Use the result of (*b*).

f) On a Hilbert space of dimension d, the space of linear maps from Hermitian operators to Hermitian operators has real dimension d^4. What is the dimension of the space of trace-preserving maps? **Hint**: Count the number of independent conditions that must be imposed to ensure that the map is trace preserving.

3.11 Orthogonal or not?

Consider a generalized measurement (POVM) on an d-dimensional Hilbert space. There are d possible outcomes for the measurement corresponding to the d nonnegative operators $\boldsymbol{E}_a$, $a = 0, 1, 2, \ldots, d-1$, where $\sum_{a=0}^{d-1} \boldsymbol{E}_a = \boldsymbol{I}$. Suppose that each $\boldsymbol{E}_a$ is one-dimensional (has one nonzero eigenvalue). Is this POVM necessarily an orthogonal measurement? Explain your answer.

3.12 Heterodyne measurement of an oscillator

The coherent states $\{|\alpha\rangle,\ \alpha \in \mathbf{C}\}$ are an overcomplete basis for a one-dimensional harmonic oscillator, satisfying

$$\langle\beta|\alpha\rangle = \exp\left(-\frac{1}{2}|\beta|^2 + \beta^*\alpha - \frac{1}{2}|\alpha|^2\right) \tag{3.252}$$

a) Show that

$$\int d^2\alpha\ \boldsymbol{E}_\alpha = \boldsymbol{I}\ , \tag{3.253}$$

where

$$\boldsymbol{E}_\alpha = \frac{1}{\pi}|\alpha\rangle\langle\alpha|\ . \tag{3.254}$$

Hint: Evaluate matrix elements of both sides of the equation between coherent states.

b) Since the $\boldsymbol{E}_\alpha$'s provide a partition of unity, they define a POVM (an "ideal heterodyne measurement" of the oscillator). Suppose that a coherent state $|\beta\rangle$ is prepared, and that an ideal heterodyne measurement is performed, so that the coherent state $|\alpha\rangle$ is obtained with probability distribution $P(\alpha)\ d^2\alpha = \langle\beta|\boldsymbol{E}_\alpha|\beta\rangle\ d^2\alpha$. With what fidelity does the measurement outcome $|\alpha\rangle$ approximate the initial coherent state $|\beta\rangle$, averaged over the possible outcomes?

3.13 Master equation for the depolarizing channel

a) Consider a depolarizing qubit that is subjected to "Pauli errors" at a rate $\tilde{\Gamma}$, where $\boldsymbol{\sigma}_1$, $\boldsymbol{\sigma}_2$, and $\boldsymbol{\sigma}_3$ errors are all equally likely.

The depolarization can be described by a master equation with Lindblad operators $\sqrt{\tilde{\Gamma}/3}\ \sigma_1$, $\sqrt{\tilde{\Gamma}/3}\ \sigma_2$, and $\sqrt{\tilde{\Gamma}/3}\ \sigma_3$. Show that this master equation has the form

$$\dot{\boldsymbol{\rho}} = -i[\boldsymbol{H}, \boldsymbol{\rho}] - \Gamma \left(\boldsymbol{\rho} - \frac{1}{2}\boldsymbol{I} \right) \ . \tag{3.255}$$

How is Γ related to $\tilde{\Gamma}$?

b) Up to an irrelevant term proportional to the identity, the most general 2×2 Hermitian matrix is

$$\boldsymbol{H} = \frac{\omega}{2}\ \boldsymbol{n} \cdot \boldsymbol{\sigma} \ , \tag{3.256}$$

where $\boldsymbol{n}$ is a unit vector. Use this form of $\boldsymbol{H}$ and the Bloch parametrization

$$\boldsymbol{\rho} = \frac{1}{2}(\boldsymbol{I} + \vec{P} \cdot \vec{\boldsymbol{\sigma}}) \ , \tag{3.257}$$

to show that the master equation eq. (3.255) can be rewritten as

$$\dot{\vec{P}} = \omega(\boldsymbol{n} \times \vec{P}) - \Gamma \vec{P} \ . \tag{3.258}$$

Thus the polarization precesses uniformly with circular frequency ω about the $\boldsymbol{n}$-axis as it contracts with lifetime Γ^{-1}.

c) Alice and Bob play a game in which Alice decides to "turn on" one of the two Hamiltonians

$$\boldsymbol{H} = \frac{\omega}{2}\ \sigma_3 \ , \qquad \boldsymbol{H'} = 0 \ , \tag{3.259}$$

and Bob is to guess which Hamiltonian Alice chose. Bob has a supply of qubits, and he can observe whether the qubits "precess" in order to distinguish $\boldsymbol{H}$ from $\boldsymbol{H'}$. However, his qubits are also subject to depolarization at the rate Γ as in eq. (3.255). Suppose that Bob prepares his qubits at time 0 with polarization $\vec{P}_0 = (1, 0, 0)$; after time t elapses, (1) find the polarization $\vec{P}(t)$ if the Hamiltonian is $\boldsymbol{H}$ and (2) find the polarization $\vec{P}'(t)$ if the Hamiltonian is $\boldsymbol{H'}$.

d) What is Bob's optimal measurement for distinguishing the polarizations $\vec{P}(t)$ and $\vec{P}'(t)$ (assuming that Alice is as likely to choose $\boldsymbol{H}$ as $\boldsymbol{H'}$? What is his optimal probability of error $(p_{\rm e})_{\rm opt}\,(t)$?

e) The probability of error is smallest if Bob waits for a time $t_{\rm best}$ before measuring. Find $t_{\rm best}$ as a function of Γ and ω. Does your answer make sense in the limits $\Gamma \gg \omega$ and $\Gamma \ll \omega$?