A straightforward introduction to continuous quantum measurement

KURT JACOBS*†‡ and DANIEL A. STECK§

†Department of Physics, University of Massachusetts at Boston,
Boston, MA 02124, USA
‡Quantum Sciences and Technologies Group, Hearne Institute for Theoretical Physics,
Department of Physics & Astronomy, Louisiana State University, Baton Rouge,
LA 70803-4001, USA
§Department of Physics and Oregon Center for Optics, 1274 University of Oregon,
Eugene, OR 97403-1274, USA

We present a pedagogical treatment of the formalism of continuous quantum measurement. Our aim is to show the reader how the equations describing such measurements are derived and manipulated in a direct manner. We also give elementary background material for those new to measurement theory, and describe further various aspects of continuous measurements that should be helpful to those wanting to use such measurements in applications. Specifically, we use the simple and direct approach of generalized measurements to derive the stochastic master equation describing the continuous measurements of observables, give a tutorial on stochastic calculus, treat multiple observers and inefficient detection, examine a general form of the measurement master equation, and show how the master equation leads to information gain and disturbance. To conclude, we give a detailed treatment of imaging the resonance fluorescence from a single atom as a concrete example of how a continuous position measurement arises in a physical system.

1. Introduction

When measurement is first introduced to students of quantum mechanics, it is invariably treated by ignoring any consideration of the time the measurement takes: the measurement just 'happens', for all intents and purposes, instantaneously. This treatment is good for a first introduction, but is not sufficient to describe two important situations. The first is when some aspect of a system is continually monitored. This happens, for example, when one illuminates an object and continually detects the reflected light in order to track the object's motion. In this case, information is obtained about the object at a finite rate, and one needs to understand what happens to the object *while* the measurement takes place. It is the subject of *continuous quantum measurement* that describes such a measurement. The second situation arises because nothing really happens instantaneously. Even rapid, 'single shot' measurements take some time. If this time is not short compared to the dynamics of the measured system, then it is once again important to understand both the dynamics of the flow of information to the observer and the effect of the measurement on the system.

Continuous measurement has become increasingly important in the last decade, due mainly to the growing interest in the application of *feedback control* in quantum systems [1–11].

---

*Corresponding author. Email: kurt.jacobs@umb.edu or kjacobs@cs.umb.edu

In feedback control a system is continuously measured, and this information is used while the measurement proceeds (that is, in real time) to modify the system Hamiltonian so as to obtain some desired behaviour. Thus, continuous measurement theory is essential for describing feedback control. The increasing interest in continuous measurement is also due to its applications in metrology [12–16], quantum information [17–19], quantum computing [20–22], and its importance in understanding the quantum to classical transition [23–29].

While the importance of continuous measurement grows, to date there is really only one introduction to the subject that could be described as both easily accessible and extensive, that being the one by Brun in the *American Journal of Physics* [30] (some other pedagogical treatments can be found in [31–33]). While the analysis in Brun's work is suitably direct, it treats explicitly only measurements on two-state systems, and due to their simplicity the derivations used there do not easily extend to measurements of more general observables. Since many applications involve measurements of observables in infinite-dimensional systems (such as the position of a particle), we felt that an introductory article that derived the equations for such measurements in the simplest and most direct fashion would fill an important gap in the literature. This is what we do here. Don't be put off by the length of this article—a reading of only a fraction of the article is sufficient to understand how to derive the basic equation that describes continuous measurement, the mathematics required to manipulate it (the so-called Itô calculus), and how it can be solved. This is achieved in sections 4, 5 and 6. If the reader is not familiar with the density operator, then this preliminary material is explained in section 2 and generalized quantum measurements (POVMs) are explained in section 3.

The rest of the article gives some more information about continuous measurements. In section 7 we show how to treat multiple, simultaneous observers and inefficient detectors, both of which involve simple and quite straightforward generalizations of the basic equation. In section 8 we discuss the most general form that the continuous-measurement equation can take. In section 9 we present explicit calculations to explain the meaning of the various terms in the measurement equation. Since our goal in the first part of this article was to derive a continuous measurement equation in the shortest and most direct manner, this did not involve a concrete physical example. In the second-to-last (and longest) section, we provide such an example, showing in considerable detail how a continuous measurement arises when the position of an atom is monitored by detecting the photons it emits. The final section concludes with some pointers for further reading.

2. Describing an observer's state of knowledge of a quantum system

2.1 *The density operator*

Before getting on with measurements, we will briefly review the density operator, since it is so central to our discussion. The density operator represents the state of a quantum system in a more general way than the state vector, and equivalently represents an observer's *state of knowledge* of a system.

When a quantum state can be represented by a state vector $|\psi\rangle$, the density operator is defined as the product

$$\rho := |\psi\rangle\langle\psi|. \tag{1}$$

In this case, it is obvious that the information content of the density operator is equivalent to that of the state vector (except for the overall phase, which is not of physical significance).

The state vector can represent states of *coherent* superposition. The power of the density operator lies in the fact that it can represent *incoherent* superpositions as well. For example, let $|\psi_\alpha\rangle$ be a set of states (without any particular restrictions). Then the density operator

$$\rho = \sum_\alpha p_\alpha |\psi_\alpha\rangle\langle\psi_\alpha| \tag{2}$$

models the fact that we don't know *which* of the states $|\psi_\alpha\rangle$ the system is in, but we know that it is in the state $|\psi_\alpha\rangle$ with probability p_α. Another way to say it is this: the state vector $|\psi\rangle$ represents a certain *intrinsic uncertainty* with respect to quantum observables; the density operator can represent uncertainty *beyond* the minimum required by quantum mechanics. Equivalently, the density operator can represent an *ensemble* of identical systems in possibly different states.

A state of the form (1) is said to be a *pure state*. One that cannot be written in this form is said to be *mixed*, and can be written in the form (2).

Differentiating the density operator and employing the Schrödinger equation $i\hbar\partial_t|\psi\rangle = H|\psi\rangle$, we can write down the equation of motion for the density operator:

$$\partial_t\rho = -\frac{i}{\hbar}[H,\rho]. \tag{3}$$

This is referred to as the *Schrödinger–von Neumann equation*. Of course, the use of the density operator allows us to write down more general evolution equations than those implied by state-vector dynamics.

2.2 *Expectation values*

We can compute expectation values with respect to the density operator via the trace operation. The trace of an

operator A is simply the sum over the diagonal matrix elements with respect to any complete, orthonormal set of states $|\beta\rangle$:

$$\text{Tr}[A] := \sum_{\beta} \langle\beta|A|\beta\rangle. \tag{4}$$

An important property of the trace is that the trace of a product is invariant under cyclic permutations of the product. For example, for three operators,

$$\text{Tr}[ABC] = \text{Tr}[BCA] = \text{Tr}[CAB]. \tag{5}$$

This amounts to simply an interchange in the order of summations. For example, for two operators, working in the position representation, we can use the fact that $\int dx \langle x|x\rangle$ is the identity operator to see that

$$\begin{aligned}
\text{Tr}[AB] &= \int dx \langle x|AB|x\rangle \\
&= \int dx \int dx' \langle x|A|x'\rangle\langle x'|B|x\rangle \\
&= \int dx' \int dx \langle x'|B|x\rangle\langle x|A|x'\rangle \\
&= \int dx' \langle x'|BA|x'\rangle \\
&= \text{Tr}[BA]. \tag{6}
\end{aligned}$$

Note that this argument assumes sufficiently 'nice' operators (it fails, for example, for $\text{Tr}[xp]$). More general permutations (e.g. of the form (5)) are obtained by replacements of the form $B \to BC$. Using this property, we can write the expectation value with respect to a pure state as

$$\langle A \rangle = \langle\psi|A|\psi\rangle = \text{Tr}[A\rho]. \tag{7}$$

This argument extends to the more general form (2) of the density operator.

2.3 *The density matrix*

The physical content of the density operator is more apparent when we compute the elements $\rho_{\alpha\alpha'}$ of the *density matrix* with respect to a complete, orthonormal basis. The density matrix elements are given by

$$\rho_{\alpha\alpha'} := \langle\alpha|\rho|\alpha'\rangle. \tag{8}$$

To analyse these matrix elements, we will assume the simple form $\rho = |\psi\rangle\langle\psi|$ of the density operator, though the arguments generalize easily to arbitrary density operators.

The diagonal elements $\rho_{\alpha\alpha}$ are referred to as *populations*, and give the probability of being in the state $|\alpha\rangle$:

$$\rho_{\alpha\alpha} = \langle\alpha|\rho|\alpha\rangle = |\langle\alpha|\psi\rangle|^2. \tag{9}$$

The off-diagonal elements $\rho_{\alpha\alpha'}$ (with $\alpha \neq \alpha'$) are referred to as *coherences*, since they give information about the relative phase of different components of the superposition. For example, if we write the state vector as a superposition with explicit phases,

$$|\psi\rangle = \sum_{\alpha} c_\alpha|\alpha\rangle = \sum_{\alpha} |c_\alpha| \exp(i\phi_\alpha)|\alpha\rangle, \tag{10}$$

then the coherences are

$$\rho_{\alpha\alpha'} = |c_\alpha c_{\alpha'}| \exp[i(\phi_\alpha - \phi_{\alpha'})]. \tag{11}$$

Notice that for a density operator not corresponding to a pure state, the coherences in general will be the sum of complex numbers corresponding to different states in the incoherent sum. The phases will not in general line up, so that while $|\rho_{\alpha\alpha'}|^2 = \rho_{\alpha\alpha}\rho_{\alpha'\alpha'}$ for a pure state, we expect $|\rho_{\alpha\alpha'}|^2 < \rho_{\alpha\alpha}\rho_{\alpha'\alpha'}$ ($\alpha \neq \alpha'$) for a generic mixed state.

2.4 *Purity*

The difference between pure and mixed states can be formalized in another way. Notice that the diagonal elements of the density matrix form a probability distribution. Proper normalization thus requires

$$\text{Tr}[\rho] = \sum_{\alpha} \rho_{\alpha\alpha} = 1. \tag{12}$$

We can do the same computation for ρ^2 and we will define the *purity* to be $\text{Tr}[\rho^2]$. For a pure state, the purity is simple to calculate:

$$\text{Tr}[\rho^2] = \text{Tr}[|\psi\rangle\langle\psi|\psi\rangle\langle\psi|] = \text{Tr}[\rho] = 1. \tag{13}$$

But for mixed states, $\text{Tr}[\rho^2] < 1$. For example, for the density operator in (2),

$$\text{Tr}[\rho^2] = \sum_{\alpha} p_\alpha^2, \tag{14}$$

if we assume the states $|\psi_\alpha\rangle$ to be orthonormal. For equal probability of being in N such states, $\text{Tr}[\rho^2] = 1/N$. Intuitively, then, we can see that $\text{Tr}[\rho^2]$ drops to zero as the state becomes more mixed—that is, as it becomes an incoherent superposition of more and more orthogonal states.

3. Weak measurements and POVMs

In undergraduate courses the only kind of measurement that is usually discussed is one in which the system is projected onto one of the possible eigenstates of a given observable. If we write these eigenstates as $\{|n\rangle : n = 1, \ldots, n_{\max}\}$, and the state of the system is $|\psi\rangle = \sum_n c_n|n\rangle$, the

probability that the system is projected onto $|n\rangle$ is $|c_n|^2$. In fact, these kinds of measurements, which are often referred to as *von Neumann* measurements, represent only a special class of all the possible measurements that can be made on quantum systems. However, all measurements can be derived from von Neumann measurements.

One reason that we need to consider a larger class of measurements is so we can describe measurements that extract only partial information about an observable. A von Neumann measurement provides complete information—after the measurement is performed we know exactly what the value of the observable is, since the system is projected into an eigenstate. Naturally, however, there exist many measurements which, while reducing on average our uncertainty regarding the observable of interest, do not remove it completely.

First, it is worth noting that a von Neumann measurement can be described by using a set of projection operators $\{P_n = |n\rangle\langle n|\}$. Each of these operators describes what happens on one of the possible outcomes of the measurement: if the initial state of the system is $\rho = |\psi\rangle\langle\psi|$, then the nth possible outcome of the final state is given by

$$\rho_{\text{f}} = |n\rangle\langle n| = \frac{P_n \rho P_n}{\text{Tr}[P_n \rho P_n]}, \tag{15}$$

and this result is obtained with probability

$$P(n) = \text{Tr}[P_n \rho P_n] = c_n, \tag{16}$$

where c_n defines the superposition of the initial state $|\psi\rangle$ given above. It turns out that *every* possible measurement may be described in a similar fashion by generalizing the set of operators. Suppose we pick a set of m_{max} operators Ω_m, the only restriction being that $\sum_{m=1}^{m_{\text{max}}} \Omega_m^\dagger \Omega_m = I$, where I is the identity operator. Then it is in principle possible to design a measurement that has N possible outcomes,

$$\rho_{\text{f}} = \frac{\Omega_m \rho \Omega_m^\dagger}{\text{Tr}[\Omega_m \rho \Omega_m^\dagger]}, \tag{17}$$

with

$$P(m) = \text{Tr}[\Omega_m \rho \Omega_m^\dagger] \tag{18}$$

giving the probability of obtaining the mth outcome.

Every one of these more general measurements may be implemented by performing a unitary interaction between the system and an auxiliary system, and then performing a von Neumann measurement on the auxiliary system. Thus all possible measurements may be derived from the basic postulates of unitary evolution and von Neumann measurement [34,35].

These 'generalized' measurements are often referred to as POVMs, where the acronym stands for 'positive operator-valued measure'. The reason for this is somewhat technical, but we explain it here because the terminology is so common. Note that the probability for obtaining a result in the range [*a,b*] is

$$P(m \in [a,b]) = \sum_{m=a}^{b} \text{Tr}[\Omega_m \rho \Omega_m^\dagger] = \text{Tr}\left[\sum_{m=a}^{b} \Omega_m^\dagger \Omega_m \rho\right]. \tag{19}$$

The positive operator $M = \sum_{m=a}^{b} \Omega_m^\dagger \Omega_m$ thus determines the probability that m lies in the subset [*a,b*] of its range. In this way the formalism associates a positive operator with *every* subset of the range of m, and is therefore a *positive operator-valued measure*.

Let us now put this into practice to describe a measurement that provides partial information about an observable. In this case, instead of our measurement operators Ω_m being projectors onto a single eigenstate, we choose them to be a weighted sum of projectors onto the eigenstates $|n\rangle$, each one peaked about a different value of the observable. Let us assume now, for the sake of simplicity, that the eigenvalues n of the observable N take on all the integer values. In this case we can choose

$$\Omega_m = \frac{1}{\mathcal{N}} \sum_n \exp\left[-k(n-m)^2/4\right]|n\rangle\langle n|, \tag{20}$$

where $\mathcal{N}$ is a normalization constant chosen so that $\sum_{m=-\infty}^{\infty} \Omega_m^\dagger \Omega_m = I$. We have now constructed a measurement that provides partial information about the observable N. This is illustrated clearly by examining the case where we start with no information about the system. In this case the density matrix is completely mixed, so that $\rho \propto I$. After making the measurement and obtaining the result m, the state of the system is

$$\rho_{\text{f}} = \frac{\Omega_m \rho \Omega_m^\dagger}{\text{Tr}[\Omega_m \rho \Omega_m^\dagger]} = \frac{1}{\mathcal{N}} \sum_n \exp\left[-k(n-m)^2/2\right]|n\rangle\langle n|. \tag{21}$$

The final state is thus peaked about the eigenvalue m, but has a width given by $1/k^{1/2}$. The larger k, the less our final uncertainty regarding the value of the observable. Measurements for which k is large are often referred to as *strong* measurements, and conversely those for which k is small are *weak* measurements [36]. These are the kinds of measurements that we will need in order to derive a continuous measurement in the next section.

4. A continuous measurement of an observable

A continuous measurement is one in which information is continually extracted from a system. Another way to say

this is that when one is making such a measurement, the amount of information obtained goes to zero as the duration of the measurement goes to zero. To construct a measurement like this, we can divide time into a sequence of intervals of length Δt, and consider a weak measurement in each interval. To obtain a continuous measurement, we make the strength of each measurement proportional to the time interval, and then take the limit in which the time intervals become infinitesimally short.

In what follows, we will denote the observable we are measuring by X (i.e. X is a Hermitian operator), and we will assume that it has a continuous spectrum of eigenvalues x. We will write the eigenstates as $|x\rangle$, so that $\langle x|x'\rangle = \delta(x - x')$. However, the equation that we will derive will be valid for measurements of *any* Hermitian operator.

We now divide time into intervals of length Δt. In each time interval, we will make a measurement described by the operators

$$A(\alpha) = \left(\frac{4k\Delta t}{\pi}\right)^{1/4} \int_{-\infty}^{\infty} \exp\left[-2k\Delta t(x - \alpha)^2\right]|x\rangle\langle x|\mathrm{d}x. \quad (22)$$

Each operator $A(\alpha)$ is a Gaussian-weighted sum of projectors onto the eigenstates of X. Here α is a continuous index, so that there is a continuum of measurement results labelled by α.

The first thing we need to know is the probability density $P(\alpha)$ of the measurement result α when Δt is small. To work this out we first calculate the mean value of α. If the initial state is $|\psi\rangle = \int\psi(x)|x\rangle\mathrm{d}x$ then $P(\alpha) = \mathrm{Tr}[A(\alpha)^\dagger A(\alpha)|\psi\rangle\langle\psi|]$, and we have

$$\begin{aligned}\langle\alpha\rangle &= \int_{-\infty}^{\infty} \alpha P(\alpha)\mathrm{d}\alpha \\ &= \int_{-\infty}^{\infty} \alpha\mathrm{Tr}[A(\alpha)^\dagger A(\alpha)|\psi\rangle\langle\psi|]\mathrm{d}\alpha \\ &= \sqrt{\frac{4k\Delta t}{\pi}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \alpha|\psi(x)|^2\exp\left[-4k\Delta t(x - \alpha)^2\right]\mathrm{d}x\,\mathrm{d}\alpha \\ &= \int_{-\infty}^{\infty} x|\psi(x)|^2\mathrm{d}x = \langle X\rangle.\end{aligned} \quad (23)$$

To obtain $P(\alpha)$ we now write

$$\begin{aligned}P(\alpha) &= \mathrm{Tr}[A(\alpha)^\dagger A(\alpha)|\psi\rangle\langle\psi|] \\ &= \left(\frac{4k\Delta t}{\pi}\right)^{1/2} \int_{-\infty}^{\infty} |\psi(x)|^2 \exp\left[-4k\Delta t(x - \alpha)^2\right]\mathrm{d}x.\end{aligned} \quad (24)$$

If Δt is sufficiently small then the Gaussian is much broader than $\psi(x)$. This means we can approximate $|\psi(x)|^2$ by a delta function, which must be centred at the expected

position $\langle X\rangle$ so that $\langle\alpha\rangle = \langle X\rangle$ as calculated above. We therefore have

$$\begin{aligned}P(\alpha) &\approx \left(\frac{4k\Delta t}{\pi}\right)^{1/2} \int_{-\infty}^{\infty} \delta(x - \langle X\rangle) \exp\left[-4k\Delta t(x - \alpha)^2\right]\mathrm{d}x \\ &= \left(\frac{4k\Delta t}{\pi}\right)^{1/2} \exp\left[-4k\Delta t(\alpha - \langle X\rangle)^2\right].\end{aligned} \quad (25)$$

We can also write α as the stochastic quantity

$$\alpha_\mathrm{s} = \langle X\rangle + \frac{\Delta W}{(8k)^{1/2}\Delta t}, \quad (26)$$

where ΔW is a zero-mean, Gaussian random variable with variance Δt. This alternate representation as a stochastic variable will be useful later. Since it will be clear from context, we will use α interchangeably with α_s in referring to the measurement results, although technically we should distinguish between the index α and the stochastic variable α_s.

A *continuous* measurement results if we make a sequence of these measurements and take the limit as $\Delta t \to 0$ (or equivalently, as $\Delta t \to \mathrm{d}t$). As this limit is taken, more and more measurements are made in any finite time interval, but each is increasingly weak. By choosing the variance of the measurement result to scale as Δt, we have ensured that we obtain a sensible continuum limit. A stochastic equation of motion results due to the random nature of the measurements (a *stochastic* variable is one that fluctuates randomly over time). We can derive this equation of motion for the system under this continuous measurement by calculating the change induced in the quantum state by the single weak measurement in the time step Δt, to first order in Δt. We will thus compute the evolution when a measurement, represented by the operator $A(\alpha)$, is performed in each time step. This procedure gives

$$\begin{aligned}|\psi(t + \Delta t)\rangle &\propto A(\alpha)|\psi(t)\rangle \\ &\propto \exp\left[-2k\Delta t(\alpha - X)^2\right]|\psi(t)\rangle \\ &\propto \exp\left(-2k\Delta t X^2 + X[4k\langle X\rangle\Delta t + (2k)^{1/2}\Delta W]\right)|\psi(t)\rangle.\end{aligned} \quad (27)$$

We now expand the exponential to first order in Δt, which gives

$$\begin{aligned}|\psi(t + \Delta t)\rangle \propto \{1 &- 2k\Delta t X^2 \\ &+ X[4k\langle X\rangle\Delta t + (2k)^{1/2}\Delta W + kX(\Delta W)^2]\}|\psi(t)\rangle.\end{aligned} \quad (28)$$

Note that we have included the *second-order* term in ΔW in the power series expansion for the exponential. We need to include this term because it turns out that in the limit in

which $\Delta t \to 0$, $(\Delta W)^2 \to (\mathrm{d}W)^2 = \mathrm{d}t$. Because of this, the $(\Delta W)^2$ term contributes to the final differential equation. The reason for this will be explained in the next section, but for now we ask the reader to indulge us and accept that it is true.

To take the limit as $\Delta t \to 0$, we set $\Delta t = \mathrm{d}t$, $\Delta W = \mathrm{d}W$ and $(\Delta W)^2 = \mathrm{d}t$, and the result is

$$|\psi(t + \mathrm{d}t)\rangle \propto \{1 - [kX^2 - 4kX\langle X\rangle]\mathrm{d}t + (2k)^{1/2}X\,\mathrm{d}W\}|\psi(t)\rangle. \tag{29}$$

This equation does not preserve the norm $\langle\psi|\psi\rangle$ of the wave function, because before we derived it we threw away the normalization. We can easily obtain an equation that *does* preserve the norm simply by normalizing $|\psi(t + \mathrm{d}t)\rangle$ and expanding the result to first order in $\mathrm{d}t$ (again, keeping terms to order $\mathrm{d}W^2$). Writing $|\psi(t + \mathrm{d}t)\rangle = |\psi(t)\rangle + \mathrm{d}|\psi\rangle$, the resulting stochastic differential equation is given by

$$\mathrm{d}|\psi\rangle = \{-k(X - \langle X\rangle)^2\,\mathrm{d}t + (2k)^{1/2}(X - \langle X\rangle)\mathrm{d}W\}|\psi(t)\rangle. \tag{30}$$

This is the equation we have been seeking—it describes the evolution of the state of a system in a time interval $\mathrm{d}t$ given that the observer obtains the measurement result

$$\mathrm{d}y = \langle X\rangle\mathrm{d}t + \frac{\mathrm{d}W}{(8k)^{1/2}} \tag{31}$$

in that time interval. The measurement result gives the expected value $\langle X\rangle$ plus a random component due to the width of $P(\alpha)$, and we write this as a differential since it corresponds to the information gained in the time interval $\mathrm{d}t$. As the observer integrates $\mathrm{d}y(t)$ the quantum state progressively collapses, and this integration is equivalent to solving (30) for the quantum-state evolution.

The *stochastic Schrödinger equation* (SSE) in equation (30) is usually described as giving the evolution *conditioned* upon the stream of measurement results. The state $|\psi\rangle$ evolves randomly, and $|\psi(t)\rangle$ is called the *quantum trajectory* [33]. The set of measurement results $\mathrm{d}y(t)$ is called the *measurement record*. We can also write this SSE in terms of the density operator ρ instead of $|\psi\rangle$. Remembering that we must keep all terms proportional to $\mathrm{d}W^2$, and defining $\rho(t + \mathrm{d}t) \equiv \rho(t) + \mathrm{d}\rho$, we have

$$\begin{aligned}
\mathrm{d}\rho &= (\mathrm{d}|\psi\rangle)\langle\psi| + |\psi\rangle(\mathrm{d}\langle\psi|) + (\mathrm{d}|\psi\rangle)(\mathrm{d}\langle\psi|)\\
&= -k[X[X,\rho]]\,\mathrm{d}t\\
&\quad + (2k)^{1/2}(X\rho + \rho X - 2\langle X\rangle\rho)\mathrm{d}W.
\end{aligned} \tag{32}$$

This is referred to as a *stochastic master equation* (SME), which also defines a quantum trajectory $\rho(t)$. This SME was

first derived by Belavkin [1]. Note that in general, the SME also includes a term describing Hamiltonian evolution as in equation (3).

The density operator at time t gives the observer's state of knowledge of the system, given that she has obtained the measurement record $y(t)$ up until time t. Since the observer has access to $\mathrm{d}y$ but not to $\mathrm{d}W$, to calculate $\rho(t)$ she must calculate $\mathrm{d}W$ at each time step from the measurement record in that time step along with the expectation value of X at the previous time:

$$\mathrm{d}W = (8k)^{1/2}(\mathrm{d}y - \langle X\rangle\mathrm{d}t). \tag{33}$$

By substituting this expression in the SME (equation (32)), we can write the evolution of the system directly in terms of the measurement record, which is the natural thing to do from the point of the view of the observer. This is

$$\begin{aligned}
\mathrm{d}\rho &= -k[X[X,\rho]]\mathrm{d}t\\
&\quad + 4k(X\rho + \rho X - 2\langle X\rangle\rho)(\mathrm{d}y - \langle X\rangle\mathrm{d}t).
\end{aligned} \tag{34}$$

In section 6 we will explain how to solve the SME analytically in a special case, but it is often necessary to solve it numerically. The simplest method of doing this is to take small time steps Δt, and use a random number generator to select a new ΔW in each time step. One then uses Δt and ΔW in each time step to calculate $\Delta\rho$ and adds this to the current state ρ. In this way we generate a specific trajectory for the system. Each possible sequence of $\mathrm{d}W$'s generates a different trajectory, and the probability that a given trajectory occurs is the probability that the random number generator gives the corresponding sequence of $\mathrm{d}W$'s. A given sequence of $\mathrm{d}W$'s is often referred to as a 'realization' of the noise, and we will refer to the process of generating a sequence of $\mathrm{d}W$'s as 'picking a noise realization'. Further details regarding the numerical methods for solving stochastic equations are given in [37].

If the observer makes the continuous measurement, but throws away the information regarding the measurement results, the observer must average over the different possible results. Since ρ and $\mathrm{d}W$ are statistically independent, $\langle\!\langle\rho\mathrm{d}W\rangle\!\rangle = 0$, where the double brackets denote this average (as we show in section 5.2.3). The result is thus given by setting to zero all terms proportional to $\rho\mathrm{d}W$ in equation (32),

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -k[X[X,\rho]], \tag{35}$$

where the density operator here represents the state averaged over all possible measurement results. We note that the method we have used above to derive the stochastic Schrödinger equation is an extension of a method initially developed by Caves and Milburn to derive the (non-stochastic) master equation (35) [38].

5. An introduction to stochastic calculus

Now that we have encountered a noise process in the quantum evolution, we will explore in more detail the formalism for handling this. It turns out that adding a white-noise stochastic process changes the basic structure of the calculus for treating the evolution equations. There is more than one formulation to treat stochastic processes, but the one referred to as *Itô calculus* is used in almost all treatments of noisy quantum systems, and so this is the one we describe here. The main alternative formalism may be found in [37,39].

5.1 *Usage*

First, let us review the usual calculus in a slightly different way. A differential equation

$$\frac{dy}{dt} = \alpha \tag{36}$$

can be instead written in terms of differentials as

$$dy = \alpha \, dt. \tag{37}$$

The basic rule in the familiar *deterministic* calculus is that $(dt)^2 = 0$. To see what we mean by this, we can try calculating the differential dz for the variable $z = \exp y$ in terms of the differential for dy as follows:

$$dz = \exp(y + dy) - \exp y = z[\exp(\alpha \, dt) - 1]. \tag{38}$$

Expanding the exponential and applying the rule $(dt)^2 = 0$, we find

$$dz = z\alpha \, dt. \tag{39}$$

This is, of course, the same result as that obtained by using the chain rule to calculate dz/dy and multiplying through by dy. The point here is that calculus breaks up functions and considers their values within short intervals Δt. In the infinitesimal limit, the quadratic and higher order terms in Δt end up being too small to contribute.

In Itô calculus, we have an additional differential element dW, representing white noise. The basic rule of Itô calculus is that $dW^2 = dt$, while $dt^2 = dt \, dW = 0$. We will justify this later, but to use this calculus, we simply note that we 'count' the increment dW as if it were equivalent to $dt^{1/2}$ in deciding what orders to keep in series expansions of functions of dt and dW. As an example, consider the stochastic differential equation

$$dy = \alpha \, dt + \beta \, dW. \tag{40}$$

We obtain the corresponding differential equation for $z = \exp y$ by expanding to *second* order in dy:

$$dz = \exp y[\exp(dy) - 1] = z\left(dy + \frac{(dy)^2}{2}\right). \tag{41}$$

Only the dW component contributes to the quadratic term; the result is

$$dz = z\left(\alpha + \frac{\beta^2}{2}\right)dt + z\beta \, dW. \tag{42}$$

The extra β^2 term is crucial in understanding many phenomena that arise in continuous-measurement processes.

5.2 *Justification*

5.2.1. Wiener process. To see why all this works, let us first define the *Wiener process* $W(t)$ as an 'ideal' random walk with arbitrarily small, independent steps taken arbitrarily often. (The Wiener process is thus scale-free and in fact fractal.) Being a symmetric random walk, $W(t)$ is a normally distributed random variable with zero mean, and we choose the variance of $W(t)$ to be t (i.e. the width of the distribution is $t^{1/2}$, as is characteristic of a diffusive process). We can thus write the probability density for $W(t)$ as

$$P(W, t) = \frac{1}{(2\pi t)^{1/2}} \exp\left(-W^2/2t\right). \tag{43}$$

In view of the central-limit theorem, *any* simple random walk gives rise to a Wiener process in the continuous limit, independent of the one-step probability distribution (so long as the one-step variance is finite).

Intuitively, $W(t)$ is a continuous but everywhere non-differentiable function. Naturally, the first thing we will want to do is to develop the analogue of the derivative for the Wiener process. We can start by defining the Wiener increment

$$\Delta W(t) := W(t + \Delta t) - W(t) \tag{44}$$

corresponding to a time increment Δt. Again, ΔW is a normally distributed random variable with zero mean and variance Δt. Note again that this implies that the root-mean-square amplitude of ΔW scales as $(\Delta t)^{1/2}$. We can understand this intuitively since the *variances* add for successive steps in a random walk. Mathematically, we can write the variance as

$$\langle\!\langle (\Delta W)^2 \rangle\!\rangle = \Delta t, \tag{45}$$

where the double angle brackets $\langle\!\langle \ \rangle\!\rangle$ denote an ensemble average over all possible realizations of the Wiener process. This relation suggests the above notion that second-order terms in ΔW contribute at the same level as first-order terms in Δt. In the infinitesimal limit of $\Delta t \to 0$, we write $\Delta t \to dt$ and $\Delta W \to dW$.

5.2.2. Itô rule. We now want to show that the Wiener differential dW satisfies the Itô rule $dW^2 = dt$. Note that we

want this to hold *without* the ensemble average, which is surprising since dW is a stochastic quantity, while dt obviously is not. To do this, consider the probability density function for $(\Delta W)^2$, which we can obtain by a simple transformation of the Gaussian probability density for ΔW (which is equation (43) with $t \to \Delta t$ and $W \to \Delta W$):

$$P\left[(\Delta W)^2\right] = \frac{\exp\left[-(\Delta W)^2/2\Delta t\right]}{[2\pi\Delta t(\Delta W)^2]^{1/2}}. \qquad (46)$$

In particular, the mean and variance of this distribution for $(\Delta W)^2$ are

$$\langle\langle(\Delta W)^2\rangle\rangle = \Delta t \qquad (47)$$

and

$$\mathrm{Var}[(\Delta W)^2] = 2(\Delta t)^2, \qquad (48)$$

respectively. To examine the continuum limit, we will sum the Wiener increments over N intervals of duration $\Delta t_N = t/N$ between 0 and t. The corresponding Wiener increments are

$$\Delta W_n := W[(n+1)\Delta t_N] - W(n\Delta t_N). \qquad (49)$$

Now consider the sum of the squared increments

$$\sum_{n=0}^{N-1}(\Delta W_n)^2, \qquad (50)$$

which corresponds to a random walk of N steps, where a single step has average value t/N and variance $2t^2/N^2$. According to the central limit theorem, for large N the sum (50) is a Gaussian random variable with mean t and variance $2t^2/N$. In the limit $N \to \infty$, the variance of the sum vanishes, and the sum becomes t with certainty. Symbolically, we can write

$$\int_0^t [dW(t')]^2 := \lim_{N\to\infty} \sum_{n=0}^{N-1}(\Delta W_n)^2 = t = \int_0^t dt'. \qquad (51)$$

For this to hold over any interval $(0,t)$, we must make the formal identification $dt = dW^2$. This means that even though dW is a random variable, dW^2 is not, since it has no variance when integrated over any finite interval.

5.2.3. Ensemble averages. Finally, we need to justify a relation useful for averaging over noise realizations, namely that

$$\langle\langle y\, dW\rangle\rangle = 0 \qquad (52)$$

for a solution $y(t)$ of equation (40). This makes it particularly easy to compute averages of functions of $y(t)$

over all possible realizations of a Wiener process, since we can simply set $dW = 0$, even when it is multiplied by y. We can see this as follows. Clearly, $\langle\langle dW\rangle\rangle = 0$. Also, equation (40) is the continuum limit of the discrete relation

$$y(t + \Delta t) = y(t) + \alpha\Delta t + \beta\Delta W(t). \qquad (53)$$

Thus, $y(t)$ depends on $\Delta W(t - \Delta t)$, but is independent of $W(t)$, which gives the desired result, equation (52). More detailed discussions of Wiener processes and Itô calculus may be found in [39,40].

6. Solution of a continuous measurement

The stochastic equation (32) that describes the dynamics of a system subjected to a continuous measurement is nonlinear in ρ, which makes it difficult to solve. However, it turns out that this equation can be recast in an effectively equivalent but linear form. We now derive this linear form, and then show how to use it to obtain a complete solution to the SME. To do this, we first return to the unnormalized stochastic Schrödinger equation (29). Writing this in terms of the measurement record dy from equation (31), we have

$$|\tilde{\psi}(t + dt)\rangle = \{1 - kX^2\, dt + 4kX\, dy\}|\tilde{\psi}(t)\rangle, \qquad (54)$$

where the tilde denotes that the state is not normalized (hence the equality here). Note that the nonlinearity in this equation is entirely due to the fact that dy depends upon $\langle X\rangle$ (and $\langle X\rangle$ depends upon ρ). So what would happen if we simply replaced dy in this equation with $dW/(8k)^{1/2}$? This would mean that we would be choosing the measurement record incorrectly in each time step dt. But the ranges of both dy and dW are the full real line, so replacing dy by $dW/(8k)^{1/2}$ still corresponds to a *possible* realization of dy. However, we would then be using the wrong *probability density* for dy because dy and $dW/(8k)^{1/2}$ have different means. Thus, if we were to use $dW/(8k)^{1/2}$ in place of dy we would obtain all the correct trajectories, but with the wrong probabilities.

Now recall from section 3 that when we apply a measurement operator to a quantum state, we must explicitly renormalize it. If we do not renormalize, the new norm contains information about the prior state: it represents the *prior* probability that the particular measurement outcome actually occurred. Because the operations that result in each succeeding time interval dt are independent, and probabilities for independent events multiply, this statement remains true after any number of time steps. That is, after n time steps, the norm of the state records the probability that the sequence of measurements led to that state. To put it yet another way, it records the probability that that particular trajectory occurred. This is

extremely useful, because it means that we do not have to choose the trajectories with the correct probabilities—we can recover these at the end merely by examining the final norm!

To derive the linear form of the SSE we use the observations above. We start with the *normalized* form given by equation (30), and write it in terms of dy, which gives

$$|\psi(t+dt)\rangle = \{1 - k(X - \langle X \rangle)^2 \, dt \\ + 4k(X - \langle X \rangle)(dy - \langle X \rangle dt)\}|\psi(t)\rangle. \quad (55)$$

We then replace dy by $dW/(8k)^{1/2}$ (that is, we remove the mean from dy at each time step). In addition, we multiply the state by the square root of the actual probability for getting that state (the probability for dy) and divide by the square root of the probability for dW. To first order in dt, the factor we multiply by is therefore

$$\left(\frac{P(dW)}{P(dy)}\right)^{1/2} = 1 + (2k)^{1/2}\langle X \rangle dW - k\langle X \rangle^2 \, dt. \quad (56)$$

The resulting stochastic equation is linear, being

$$|\tilde{\psi}(t+dt)\rangle = \{1 - kX^2 \, dt + (2k)^{1/2} X \, dW\}|\tilde{\psi}(t)\rangle. \quad (57)$$

The linear stochastic master equation equivalent to this linear SSE is

$$d\tilde{\rho} = -k[X,[X,\tilde{\rho}]]dt + (2k)^{1/2}(X\tilde{\rho} + \tilde{\rho}X)dW. \quad (58)$$

Because of the way we have constructed this equation, the actual probability at time t for getting a particular trajectory is the product of (1) the norm of the state at time t and (2) the probability that the trajectory is generated by the linear equation (the latter factor being the probability for picking the noise realization that generates the trajectory). This may sound complicated, but it is actually quite simple in practice, as we will now show. Further information regarding linear SSEs may be found in the accessible and detailed discussion given by Wiseman in [32].

We now solve the linear SME to obtain a complete solution to a quantum measurement in the special case in which the Hamiltonian commutes with the measured observable X. A technique that allows a solution to be obtained in some more general cases may be found in [41]. To solve equation (58), we include a Hamiltonian of the form $H = f(X)$, and write the equation as an exponential to first order in dt. The result is

$$\tilde{\rho}(t+dt) = \exp\{[-iH/\hbar - 2kX^2]dt + (2k)^{1/2}X\,dW\}\tilde{\rho}(t) \\ \times \exp\{[iH/\hbar - 2kX^2]dt + (2k)^{1/2}X\,dW\}, \quad (59)$$

which follows by expanding the exponentials (again to first order in dt and second order in dW) to see that this expression is equivalent to equation (58). What we have written is the generalization of the usual unitary time-evolution operator under standard Schrödinger-equation evolution. The evolution for a finite time t is easily obtained now by repeatedly multiplying on both sides by these exponentials. We can then combine all the exponentials on each side in a single exponential, since all the operators commute. The result is

$$\tilde{\rho}(t;W) = \exp\{[-iH/\hbar - 2kX^2]t + (2k)^{1/2}XW\}\tilde{\rho}(0) \\ \times \exp\{[iH/\hbar - 2kX^2]t + (2k)^{1/2}XW\}, \quad (60)$$

where the final states $\tilde{\rho}(t;W)$ are parameterized by W, with

$$W = \int_0^t dW(t'). \quad (61)$$

The probability density for W, being the sum of the Gaussian random variables dW, is Gaussian. In particular, as in equation (43), at time t the probability density is

$$\tilde{P}(W,t) = \frac{1}{(2\pi t)^{1/2}} \exp[-W^2/(2t)]. \quad (62)$$

That is, at time t, W is a Gaussian random variable with mean zero and variance t.

As we discussed above, however, the probability for obtaining $\rho(t)$ is not the probability with which it is generated by picking a noise realization. To calculate the 'true' probability for $\rho(t)$ we must multiply the density $P(W,t)$ by the norm of $\tilde{\rho}(t)$. Thus, the actual probability for getting a final state $\rho(t)$ (that is, a specific value of W at time t) is

$$P(W,t) = \frac{1}{(2\pi t)^{1/2}} \exp[-W^2/(2t)] \\ \times \text{Tr}[\exp\{[-4kX^2]t + (8k)^{1/2}XW\}\rho(0)]. \quad (63)$$

At this point, X can just as well be *any* Hermitian operator. Let us now assume that $X = J_z$ for some quantum number j of the angular momentum. In this case X has $2j+1$ eigenvectors $|m\rangle$, with eigenvalues $m = -j, -j+1, \ldots, j$. If we have no information about the system at the start of the measurement, so that the initial state is $\rho(0) = I/(2j+1)$, then the solution is quite simple. In particular, $\rho(t)$ is diagonal in the J_z eigenbasis, and

$$\langle m|\rho(t)|m\rangle = \frac{\exp[-4kt(m-Y)^2]}{\mathcal{N}}, \quad (64)$$

where $\mathcal{N}$ is the normalization and $Y := W/((8k)^{1/2} t)$. The true probability density for Y is

$$P(Y, t) = \frac{1}{2j+1} \sum_{n=-j}^{j} \left(\frac{4kt}{\pi} \right)^{1/2} \exp\left[-4kt(Y-n)^2 \right]. \quad (65)$$

We therefore see that after a sufficiently long time, the density for Y is sharply peaked about the $2j+1$ eigenvalues of J_z. This density is plotted in figure 1 for three values of t. At long times, Y becomes very close to one of these eigenvalues. Further, we see from the solution for $\rho(t)$ that when Y is close to an eigenvalue m, then the state of the system is sharply peaked about the eigenstate $|m\rangle$. Thus, we see that after a sufficiently long time, the system is projected into one of the eigenstates of J_z.

The random variable Y has a physical meaning. Since we replaced the measurement record dy by $dW/(8k)^{1/2}$ to obtain the linear equation, when we transform from the raw probability density $\tilde{P}$ to the true density P this transforms the driving noise process dW back into $(8k)^{1/2}$ $dy = (8k)^{1/2} \langle X(t) \rangle dt + dW$, being a scaled version of the measurement record. Thus, $Y(t)$, as we have defined it, is actually the output record up until time t, divided by t. That is,

$$Y = \frac{1}{t} \int_0^t \langle J_z(t) \rangle dt + \frac{1}{(8k)^{1/2} t} \int_0^t dW. \quad (66)$$

Thus, Y is the measurement result. When making the measurement the observer integrates up the measurement record, and then divides the result by the final time. The result is Y, and the closer Y is to one of the eigenvalues, and the longer the time of the measurement, the more certain the observer is that the system has been collapsed onto the eigenstate with that eigenvalue. Note that as the measurement progresses, the second, explicitly stochastic term
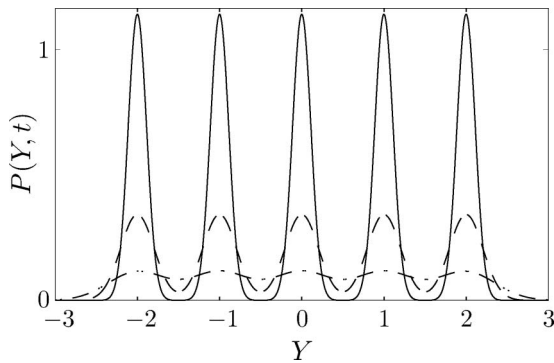

Figure 1. Here we show the probability density for the result of a measurement of the z component of angular momentum for $j = 2$, and with measurement strength k. This density is shown for three different measurement times: dot-dashed line: $t = 1/k$; dashed line: $t = 3/k$; solid line: $t = 10/k$.

converges to zero, while the expectation value in the first term evolves to the measured eigenvalue.

7. Multiple observers and inefficient detection

It is not difficult to extend the above analysis to describe what happens when more than one observer is monitoring the system. Consider two observers Alice and Bob, who measure the same system. Alice monitors X with strength k, and Bob monitors Y with strength κ. From Alice's point of view, since she has no access to Bob's measurement results, she must average over them. Thus, as far as Alice is concerned, Bob's measurement simply induces the dynamics $d\rho_1 = -\kappa[Y,[Y,\rho_1]]$ where ρ_1 is her state of knowledge. The full dynamics of her state of knowledge, including her measurement, evolves according to

$$d\rho_1 = -k[X[X,\rho_1]]dt - \kappa[Y[Y,\rho_1]]dt$$
$$+ (2k)^{1/2}(X\rho_1 + \rho_1 X - 2\langle X \rangle_1 \rho_1)dW_1, \quad (67)$$

where $\langle X \rangle_1 := \mathrm{Tr}[X\rho_1]$, and her measurement record is $dr_1 = \langle X \rangle_1 dt + dW_1/(8k)^{1/2}$. Similarly, the equation of motion for Bob's state of knowledge is

$$d\rho_2 = -\kappa[Y[Y,\rho_2]]dt - k[X[X,\rho_2]]dt$$
$$+ (2\kappa)^{1/2}(Y\rho_2 + \rho_2 Y - 2\langle Y \rangle_2 \rho_2)dW_2, \quad (68)$$

and his measurement record is $dr_2 = \langle Y \rangle_2 dt + dW_2/(8\kappa)^{1/2}$.

We can also consider the state of knowledge of a single observer, Charlie, who has access to both measurement records dr_1 and dr_2. The equation for Charlie's state of knowledge, ρ, is obtained simply by applying both measurements simultaneously, giving

$$d\rho = -k[X[X,\rho]]dt + (2k)^{1/2}(X\rho + \rho X - 2\langle X \rangle \rho)dV_1$$
$$- \kappa[Y[Y,\rho]]dt + (2\kappa)^{1/2}(Y\rho + \rho Y - 2\langle Y \rangle \rho)dV_2, \quad (69)$$

where $\langle X \rangle := \mathrm{Tr}[X\rho]$. Note that dV_1 and dV_2 are independent noise sources. In terms of Charlie's state of knowledge the two measurement records are

$$dr_1 = \langle X \rangle dt + \frac{dV_1}{(8k)^{1/2}},$$
$$dr_2 = \langle Y \rangle dt + \frac{dV_2}{(8\kappa)^{1/2}}. \quad (70)$$

In general Charlie's state of knowledge $\rho(t) \neq \rho_1(t) \neq \rho_2(t)$, but Charlie's measurement records are the same as Alice's and Bob's. Equating Charlie's expressions for the measurement records with Alice's and Bob's, we obtain the

relationship between Charlie's noise sources and those of Alice and Bob:

$$dV_1 = (8k)^{1/2}(\langle X \rangle_1 - \langle X \rangle)dt + dW_1,$$
$$dV_2 = (8\kappa)^{1/2}(\langle Y \rangle_2 - \langle Y \rangle)dt + dW_2. \tag{71}$$

We note that in quantum optics, each measurement is often referred to as a separate 'output channel' for information, and so multiple simultaneous measurements are referred to as multiple output channels. Multiple observers were first treated explicitly by Barchielli, who gives a rigorous and mathematically sophisticated treatment in [42]. A similarly detailed and considerably more accessible treatment is given in [43].

We turn now to *inefficient* measurements, which can be treated in the same way as multiple observers. An inefficient measurement is one in which the observer is not able to pick up all the measurement signal. The need to consider inefficient measurements arose originally in quantum optics, where photon counters will only detect some fraction of the photons incident upon them. This fraction, usually denoted by η, is referred to as the *efficiency of the detector* [44]. A continuous measurement in which the detector is inefficient can be described by treating the single measurement as two measurements, where the strengths of each of them sum to the strength of the single measurement. Thus we rewrite the equation for a measurement of X at strength k as

$$d\rho = -k_1[X,[X,\rho]]dt + (2k_1)^{1/2}(X\rho + \rho X - 2\langle X \rangle \rho)dV_1$$
$$- k_2[X,[X,\rho]]dt + (2k_2)^{1/2}(X\rho + \rho X - 2\langle X \rangle \rho)dV_2, \tag{72}$$

where $k_1 + k_2 = k$. We now give the observer access to only the measurement with strength k_1. From our discussion above, the equation for the observer's state of knowledge, ρ_1, is

$$d\rho_1 = -(k_1 + k_2)[X,[X,\rho_1]]dt$$
$$+ (2k_1)^{1/2}(X\rho_1 + \rho_1 X - 2\langle X \rangle_1 \rho_1)dW_1$$
$$= -k[X,[X,\rho_1]]dt$$
$$+ (2\eta k)^{1/2}(X\rho_1 + \rho_1 X - 2\langle X \rangle_1 \rho_1)dW_1, \tag{73}$$

where, as before, the measurement record is

$$dr_1 = \langle X \rangle_1 dt + \frac{dW_1}{(8k_1)^{1/2}} = \langle X \rangle_1 dt + \frac{dW_1}{(8\eta k)^{1/2}} \tag{74}$$

and

$$\eta = \frac{k_1}{k_1 + k_2} = \frac{k_1}{k} \tag{75}$$

is the efficiency of the detector.

8. General form of the stochastic master equation

Before looking at a physical example of a continuous measurement process, it is interesting to ask, what is the most general form of the measurement master equation when the measurements involve Gaussian noise? In this section we present a simplified version of an argument by Adler [45] that allows one to derive a form that is close to the fully general one and sufficient for most purposes. We also describe briefly the extension that gives the fully general form, the details of which have been worked out by Wiseman and Diosi [46].

Under unitary (unconditioned) evolution, the Schrödinger equation tells us that in a short time interval dt, the state vector undergoes the transformation

$$|\psi\rangle \rightarrow |\psi\rangle + d|\psi\rangle = \left(1 - i\frac{H}{\hbar}dt\right)|\psi\rangle, \tag{76}$$

where H is the Hamiltonian. The same transformation applied to the density operator gives the Schrödinger–von Neumann equation of equation (3):

$$\rho + d\rho = \left(1 - i\frac{H}{\hbar}dt\right)\rho\left(1 + i\frac{H}{\hbar}dt\right) = \rho - \frac{i}{\hbar}[H,\rho]dt. \tag{77}$$

To be physical, any transformation of the density operator must be *completely positive*. That is, the transformation must preserve the fact that the density operator has only nonnegative eigenvalues. This property guarantees that the density operator can generate only sensible (nonnegative) probabilities. (To be more precise, *complete* positivity means that the transformation for a system's density operator must preserve the positivity of the density operator—the fact that the density operator has no negative eigenvalues—of any larger system containing the system [34].) It turns out that the most general form of a completely positive transformation is

$$\rho \rightarrow \sum_n A_n \rho A_n^\dagger, \tag{78}$$

where the A_n are arbitrary operators. The Hamiltonian evolution above corresponds to a single infinitesimal transformation operator $A = 1 - iH dt/\hbar$.

Now let us examine the transformation for a more general, *stochastic* operator of the form

$$A = 1 - i\frac{H}{\hbar}dt + b\,dt + c\,dW, \tag{79}$$

where b and c are operators. We will use this operator to 'derive' a Markovian master equation, then indicate how it can be made more general. We may assume here that b is Hermitian, since we can absorb any antihermitian part into

the Hamiltonian. Putting this into the transformation (78), we find

$$\mathrm{d}\rho = -\frac{\mathrm{i}}{\hbar}[H, \rho]\mathrm{d}t + [b, \rho]_+\mathrm{d}t$$
$$+ c\rho c^\dagger\, \mathrm{d}t + (c\rho + \rho c^\dagger)\mathrm{d}W, \qquad (80)$$

where $[A,B]_+ := AB + BA$ is the anticommutator. We can then take an average over all possible Wiener processes, which again we denote by the double angle brackets $\langle\!\langle\ \rangle\!\rangle$. From equation (52), $\langle\!\langle\rho\,\mathrm{d}W\rangle\!\rangle = 0$ in Itô calculus, so

$$\mathrm{d}\langle\!\langle\rho\rangle\!\rangle = -\frac{\mathrm{i}}{\hbar}[H, \langle\!\langle\rho\rangle\!\rangle]\mathrm{d}t + [b, \langle\!\langle\rho\rangle\!\rangle]_+\,\mathrm{d}t + c\langle\!\langle\rho\rangle\!\rangle c^\dagger\,\mathrm{d}t. \quad (81)$$

Since the operator $\langle\!\langle\rho\rangle\!\rangle$ is an average over valid density operators, it is also a valid density operator and must therefore satisfy $\mathrm{Tr}[\langle\!\langle\rho\rangle\!\rangle] = 1$. Hence we must have $\mathrm{dTr}[\langle\!\langle\rho\rangle\!\rangle] = \mathrm{Tr}[\mathrm{d}\langle\!\langle\rho\rangle\!\rangle] = 0$. Using the cyclic property of the trace, this gives

$$\mathrm{Tr}\big[\langle\!\langle\rho\rangle\!\rangle\big(2b + c^\dagger c\big)\big] = 0. \qquad (82)$$

This holds for an arbitrary density operator only if

$$b = -\frac{c^\dagger c}{2}. \qquad (83)$$

Thus we obtain the *Lindblad form* [47] of the master equation (averaged over all possible noise realizations):

$$\mathrm{d}\langle\!\langle\rho\rangle\!\rangle = -\frac{\mathrm{i}}{\hbar}[H, \langle\!\langle\rho\rangle\!\rangle]\mathrm{d}t + \mathcal{D}[c]\langle\!\langle\rho\rangle\!\rangle\mathrm{d}t. \qquad (84)$$

Here, we have defined the Lindblad superoperator

$$\mathcal{D}[c]\rho := c\rho c^\dagger - \frac{1}{2}\big(c^\dagger c\rho + \rho c^\dagger c\big), \qquad (85)$$

where 'superoperator' refers to the fact that $\mathcal{D}[c]$ operates on ρ from both sides. This is the most general (Markovian) form of the unconditioned master equation for a single dissipation process.

The full transformation from equation (80) then becomes

$$\mathrm{d}\rho = -\frac{\mathrm{i}}{\hbar}[H, \rho]\mathrm{d}t + \mathcal{D}[c]\rho\,\mathrm{d}t + (c\rho + \rho c^\dagger)\mathrm{d}W. \quad (86)$$

This is precisely the linear master equation, for which we already considered the special case of $c = (2k)^{1/2}X$ for the measurement parts in equation (58). Again, this form of the master equation does not in general preserve the trace of the density operator, since the condition $\mathrm{Tr}[\mathrm{d}\rho] = 0$ implies

$$\mathrm{Tr}\big[\rho\big(c + c^\dagger\big)\mathrm{d}W\big] = 0. \qquad (87)$$

We could interpret this relation as a constraint on c [45], but we will instead keep c an arbitrary operator and explicitly renormalize ρ at each time step by adding a term proportional to the left-hand side of (87). The result is the nonlinear form

$$\mathrm{d}\rho = -\frac{\mathrm{i}}{\hbar}[H, \rho]\mathrm{d}t + \mathcal{D}[c]\rho\,\mathrm{d}t + \mathcal{H}[c]\rho\,\mathrm{d}W, \qquad (88)$$

where the measurement superoperator is

$$\mathcal{H}[c]\rho := c\rho + \rho c^\dagger - \langle c + c^\dagger\rangle\rho. \qquad (89)$$

When c is Hermitian, the measurement terms again give precisely the stochastic master equation (32).

More generally, we may have any number of measurements, sometimes referred to as *output channels*, happening simultaneously. The result is

$$\mathrm{d}\rho = -\frac{\mathrm{i}}{\hbar}[H, \rho]\mathrm{d}t + \sum_n (\mathcal{D}[c_n]\rho\,\mathrm{d}t + \mathcal{H}[c_n]\rho\,\mathrm{d}W_n). \qquad (90)$$

This is the same as equation (88), but this time summed (integrated) over multiple possible measurement operators c_n, each with a separate Wiener noise process independent of all the others.

In view of the arguments of section 7, when the measurements are inefficient, we have

$$\mathrm{d}\rho = -\frac{\mathrm{i}}{\hbar}[H, \rho]\mathrm{d}t + \sum_n (\mathcal{D}[c_n]\rho\,\mathrm{d}t, + \eta_n^{1/2}\mathcal{H}[c_n]\rho\,\mathrm{d}W), \qquad (91)$$

where η_n is the efficiency of the nth detection channel. The corresponding measurement record for the nth process can be written

$$\mathrm{d}r(t) = \frac{\langle c_n + c_n^\dagger\rangle}{2}\,\mathrm{d}t + \frac{\mathrm{d}W_n}{(4\eta_n)^{1/2}}. \qquad (92)$$

Again, for a single, position-measurement channel of the form $c = (2k)^{1/2}X$, we recover equations (31) and (74) if we identify $\mathrm{d}r_n/(2k)^{1/2}$ as a rescaled measurement record.

The SME in equation (91) is sufficiently general for most purposes when one is concerned with measurements resulting in Wiener noise, but is not quite the most general form for an SME driven by such noise. The most general form is worked out in [46], and includes the fact that the noise sources may also be complex and mutually correlated.

9. Interpretation of the master equation

Though we now have the general form of the master equation (91), the interpretation of each of the measurement terms is not entirely obvious. In particular, the $\mathcal{H}[c]\rho$ terms (i.e. the noise terms) represent the information gain

due to the measurement process, while the $\mathcal{D}[c]\rho$ terms represent the disturbance to, or the *backaction* on, the state of the system due to the measurement. Of course, as we see from the dependence on the efficiency η, the backaction occurs independently of whether the observer uses or discards the measurement information (corresponding to $\eta = 1$ or 0, respectively).

To examine the roles of these terms further, we will now consider the equations of motion for the moments (expectation values of powers of X and P) of the canonical variables. In particular, we will specialize to the case of a single measurement channel,

$$\mathrm{d}\rho = -\frac{\mathrm{i}}{\hbar}[H, \rho]\mathrm{d}t + \mathcal{D}[c]\rho\,\mathrm{d}t + \eta^{1/2}\mathcal{H}[c]\rho\,\mathrm{d}W. \quad (93)$$

For an arbitrary operator A, we can use the master equation and $\mathrm{d}\langle A\rangle = \mathrm{Tr}[A\,\mathrm{d}\rho]$ to obtain the following equation of motion for the expectation value $\langle A\rangle$:

$$\begin{aligned}
\mathrm{d}\langle A\rangle = &-\frac{\mathrm{i}}{\hbar}\langle[A, H]\rangle\mathrm{d}t \\
&+ \left\langle c^\dagger A c - \frac{1}{2}\left(c^\dagger c A + A c^\dagger c\right)\right\rangle\mathrm{d}t \\
&+ \eta^{1/2}\langle c^\dagger A + A c - \langle A\rangle\langle c + c^\dagger\rangle\rangle\mathrm{d}W. \quad (94)
\end{aligned}$$

Now we will consider the effects of measurements on the relevant expectation values in two example cases: a position measurement, corresponding to an observable, and an antihermitian operator, corresponding to an energy damping process. As we will see, the interpretation differs slightly in the two cases. For concreteness and simplicity, we will assume the system is a harmonic oscillator of the form

$$H = \frac{P^2}{2m} + \frac{1}{2}m\omega_0^2 X^2, \quad (95)$$

and consider the lowest few moments of X and P. We will also make the simplifying assumption that the initial state is Gaussian, so that we only need to consider the simplest five moments: the means $\langle X\rangle$ and $\langle P\rangle$, the variances V_X and V_P, where $V_\alpha := \langle\alpha^2\rangle - \langle\alpha\rangle^2$, and the symmetrized covariance $C_{XP} := (1/2)\langle[X, P]_+\rangle - \langle X\rangle\langle P\rangle$. These moments completely characterize arbitrary Gaussian states (including mixed states).

9.1 *Position measurement*

In the case of a position measurement of the form $c = (2k)^{1/2}X$ as in equation (58), equation (94) becomes

$$\begin{aligned}
\mathrm{d}\langle A\rangle = &-\frac{\mathrm{i}}{\hbar}\langle[A, H]\rangle\mathrm{d}t - k\langle[X, [X, A]]\rangle\,\mathrm{d}t \\
&+ (2\eta k)^{1/2}\left[\langle[X, A]_+\rangle - 2\langle X\rangle\langle A\rangle\right]\mathrm{d}W. \quad (96)
\end{aligned}$$

Using this equation to compute the cumulant equations of motion, we find [5]

$$\begin{aligned}
\mathrm{d}\langle X\rangle &= \frac{1}{m}\langle P\rangle\mathrm{d}t + (8\eta k)^{1/2}V_X\,\mathrm{d}W, \\
\mathrm{d}\langle P\rangle &= -m\omega_0^2\langle X\rangle\mathrm{d}t + (8\eta k)^{1/2}C_{XP}\,\mathrm{d}W, \\
\partial_t V_X &= \frac{2}{m}C_{XP} - 8\eta k V_X^2, \\
\partial_t V_P &= -2m\omega_0^2 C_{XP} + 2\hbar^2 k - 8\eta k C_{XP}^2, \\
\partial_t C_{XP} &= \frac{1}{m}V_P - m\omega_0^2 V_X - 8\eta k V_X C_{XP}.
\end{aligned} \quad (97)$$

Notice that in the variance equations, the $\mathrm{d}W$ terms vanished, due to the assumption of a Gaussian state, which implies the following relations for the moments [48]:

$$\begin{aligned}
\langle X^3\rangle &= \langle X\rangle^3 + 3\langle X\rangle V_X, \\
\frac{1}{2}\langle[X, P^2]_+\rangle &= \langle X\rangle\langle P\rangle^2 + 2\langle P\rangle C_{XP} + \langle X\rangle V_P, \\
\frac{1}{2}\langle[X, [X, P]_+]_+\rangle &= \langle X\rangle\langle P\rangle^2 + 2\langle X\rangle C_{XP} + \langle P\rangle V_X.
\end{aligned} \quad (98)$$

For the reader wishing to become better acquainted with continuous measurement theory, the derivation of equations (97) is an excellent exercise. The derivation is straightforward, the only subtlety being the second-order Itô terms in the variances. For example, the equation of motion for the position variance starts as

$$\mathrm{d}V_X = \mathrm{d}\langle X^2\rangle - 2\langle X\rangle\mathrm{d}\langle X\rangle - (\mathrm{d}\langle X\rangle)^2. \quad (99)$$

The last, quadratic term is important in producing the effect that the measured quantity becomes more certain.

In examining equations (97), we can simply use the coefficients to identify the source and thus the interpretation of each term. The first term in each equation is due to the natural Hamiltonian evolution of the harmonic oscillator. Terms originating from the $\mathcal{D}[c]\rho$ component are proportional to $k\,\mathrm{d}t$ but not η; in fact, the only manifestation of this term is the $\hbar^2 k$ term in the equation of motion for V_P. Thus, a position measurement with rate constant k produces momentum diffusion (heating) at a rate $\hbar^2 k$, as is required to maintain the uncertainty principle as the position uncertainty contracts due to the measurement.

There are more terms here originating from the $\mathcal{H}[c]\rho$ component of the master equation, and they are identifiable since they are proportional to either $(\eta k)^{1/2}$ or ηk. The $\mathrm{d}W$

terms in the equations for $\langle X \rangle$ and $\langle P \rangle$ represent the stochastic nature of the position measurement. That is, during each small time interval, the wave function collapses slightly, but we do not know exactly where it collapses to. This stochastic behaviour is precisely the same behaviour that we saw in equation (26). The more subtle point here lies with the nonstochastic terms proportional to ηk, which came from the second-order term (for example, in equation (99)) where Itô calculus generates a nonstochastic term from $dW^2 = dt$. Notice in particular the term of this form in the V_X equation, which acts as a *damping* term for V_X. This term represents the certainty gained via the measurement process. The other similar terms are less clear in their interpretation, but they are necessary to maintain consistency of the evolution.

Note that we have made the assumption of a Gaussian initial state in deriving these equations, but this assumption is not very restrictive. Due to the linear potential and the Gaussian POVM for the measurement collapse, these equations of motion preserve the Gaussian form of the initial state. The Gaussian POVM additionally converts arbitrary initial states into Gaussian states at long times. Furthermore, the assumption of a Gaussian POVM is not restrictive—under the assumption of sufficiently high noise bandwidth, the central-limit theorem guarantees that temporal coarse-graining yields Gaussian noise for *any* POVM giving random deviates with bounded variance.

9.2 *Dissipation*

The position measurement above is an example of a Hermitian measurement operator. But what happens when the measurement operator is antihermitian? As an example, we will consider the annihilation operator for the harmonic oscillator by setting $c = \gamma^{1/2} a$, where

$$a = \frac{1}{2^{1/2} x_0} X + i \frac{x_0}{2^{1/2} \hbar} P \qquad (100)$$

and

$$x_0 := \left(\frac{\hbar}{m\omega_0} \right)^{1/2}. \qquad (101)$$

The harmonic oscillator with this type of measurement models, for example, the field of an optical cavity whose output is monitored via *homodyne detection*, where the cavity output is mixed on a beamsplitter with another optical field. (Technically, in homodyne detection, the field must be the same as the field driving the cavity; mixing with other fields corresponds to *heterodyne* detection.) A procedure very similar to the one above gives the following

cumulant equations for the conditioned evolution in this case:

$$d\langle X \rangle = \frac{1}{m} \langle P \rangle dt - \frac{\gamma}{2} \langle X \rangle dt$$
$$+ \left(2\eta\gamma \frac{m\omega_0}{\hbar} \right)^{1/2} \left(V_X - \frac{\hbar}{2m\omega_0} \right) dW,$$

$$d\langle P \rangle = -m\omega_0^2 \langle X \rangle dt - \frac{\gamma}{2} \langle P \rangle dt$$
$$+ \left(2\eta\gamma \frac{m\omega_0}{\hbar} \right)^{1/2} C_{XP} dW,$$

$$\partial_t V_X = \frac{2}{m} C_{XP} - \gamma \left(V_X - \frac{\hbar}{2m\omega_0} \right)$$
$$- 2\eta\gamma \frac{m\omega_0}{\hbar} \left(V_X - \frac{\hbar}{2m\omega_0} \right)^2, \qquad (102)$$

$$\partial_t V_P = -m\omega_0^2 C_{XP} - \gamma \left(V_P - \frac{m\omega_0\hbar}{2} \right)$$
$$- 2\eta\gamma \frac{m\omega_0}{\hbar} C_{XP}^2,$$

$$\partial_t C_{XP} = \frac{1}{m} V_P - m\omega_0^2 V_X - \gamma C_{XP}$$
$$- 2\eta\gamma \frac{m\omega_0}{\hbar} C_{XP} \left(V_X - \frac{\hbar}{2m\omega_0} \right).$$

The moment equations seem more complex in this case, but are still fairly simple to interpret.

First, consider the unconditioned evolution of the means $\langle X \rangle$ and $\langle P \rangle$, where we average over all possible noise realizations. Again, since $\langle\langle \rho \, dW \rangle\rangle = 0$, we can simply set $dW = 0$ in the above equations, and we will drop the double angle brackets for brevity. The Hamiltonian evolution terms are of course the same, but now we see extra damping terms. Decoupling these two equations gives an equation of the usual form for the damped harmonic oscillator for the mean position:

$$\partial_t^2 \langle X \rangle + \gamma \partial_t \langle X \rangle + \left(\omega_0^2 + \frac{\gamma^2}{4} \right) \langle X \rangle = 0. \qquad (103)$$

Note that we identify the frequency ω_0 here as the actual oscillation frequency ω_γ of the damped oscillator, given by $\omega_\gamma^2 = \omega_0^2 - \gamma^2/4$, and not the resonance frequency ω_0 that appears the usual form of the classical formula.

The noise terms in these equations correspond to *nonstationary diffusion*, or diffusion where the transport rate depends on the state of the system. Note that under such a diffusive process, the system will tend to come to rest in configurations where the diffusion coefficient vanishes, an effect closely related to the 'blowtorch theorem' [49]. Here, this corresponds to $V_X = \hbar/2m\omega_0$ and $C_{XP} = 0$.

The variance equations also contain unconditioned damping terms (proportional to γ but not η). These damping terms cause the system to equilibrate with the same variance values as noted above; they also produce the extra equilibrium value $V_P = m\omega_0\hbar/2$. The conditioning

terms (proportional to η) merely accelerate the settling to the equilibrium values. Thus, we see that the essential effect of the antihermitian measurement operator is to damp the energy from the system, whether it is stored in the centroids or in the variances. In fact, what we see is that this measurement process selects *coherent states*, states that have the same shape as the harmonic-oscillator ground state, but whose centroids oscillate along the classical harmonic-oscillator trajectories.

10. Physical model of a continuous measurement: atomic spontaneous emission

To better understand the nature of continuous measurements, we will now consider in detail an example of how a continuous measurement of position arises in a fundamental physical system: a single atom interacting with light. Again, to obtain weak measurements, we do not make projective measurements directly on the atom, but rather we allow the atom to become entangled with an auxiliary quantum system—in this case, the electromagnetic field—and then make projective measurements on the auxiliary system (in this case, using a photodetector). It turns out that this one level of separation between the system and the projective measurement is the key to the structure of the formalism. Adding more elements to the chain of quantum-measurement devices does not change the fundamental structure that we present here.

10.1 *Master equation for spontaneous emission*

We begin by considering the interaction of the atom with the electromagnetic field. In particular, treating the field quantum mechanically allows us to treat spontaneous emission. These spontaneously emitted photons can then be detected to yield information about the atom.

10.1.1. Decay of the excited state. We will give a brief treatment following the approach of Weisskopf and Wigner [50–52]. Without going into detail about the quantization of the electromagnetic field, we will simply note that the quantum description of the field involves associating a quantum harmonic oscillator with each field mode (say, each plane wave of a particular wave vector **k** and definite polarization). Then for a two-level atom with ground and excited levels $|g\rangle$ and $|e\rangle$, respectively, the uncoupled Hamiltonian for the atom and a *single* field mode is

$$H_0 = \hbar\omega_0\sigma^\dagger\sigma + \hbar\omega\left(a^\dagger a + \frac{1}{2}\right). \quad (104)$$

Here, ω_0 is the transition frequency of the atom, ω is the frequency of the field mode, $\sigma := |g\rangle\langle e|$ is the atomic lowering operator (so that $\sigma^\dagger\sigma = |e\rangle\langle e|$ is the excited-state

projector), and a is the field (harmonic oscillator) annihilation operator. The interaction between the atom and field is given in the dipole and rotating-wave approximations by the interaction Hamiltonian

$$H_{\text{AF}} = \hbar\big(g\sigma^\dagger a + g^*\sigma a^\dagger\big), \quad (105)$$

where g is a coupling constant that includes the volume of the mode, the field frequency, and the atomic dipole moment. The two terms here are the 'energy-conserving' processes corresponding to photon absorption and emission.

In the absence of externally applied fields, we can write the state vector as the superposition of the states

$$|\psi\rangle = c_e|e\rangle + c_g|g, 1\rangle, \quad (106)$$

where the uncoupled eigenstate $|\alpha, n\rangle$ denotes the atomic state $|\alpha\rangle$ and the n-photon field state, and the omitted photon number denotes the vacuum state: $|\alpha\rangle \equiv |\alpha, 0\rangle$. These states form an effectively complete basis, since no other states are coupled to these by the interaction (105). We will also assume that the atom is initially excited, so that $c_e(0) = 1$ and $c_g(0) = 0$.

The evolution is given by the Schrödinger equation,

$$\partial_t|\psi\rangle = -\frac{i}{\hbar}(H_0 + H_{\text{AF}})|\psi\rangle, \quad (107)$$

which gives, upon substitution of (106) and dropping the vacuum energy offset of the field,

$$\begin{aligned}\partial_t c_e &= -i\omega_0 c_e - igc_g, \\ \partial_t c_g &= -i\omega c_g - ig^*c_e.\end{aligned} \quad (108)$$

Defining the slowly varying amplitudes $\tilde{c}_e := c_e \exp(i\omega_0 t)$ and $\tilde{c}_g := c_g \exp(i\omega t)$, we can rewrite these as

$$\begin{aligned}\partial_t\tilde{c}_e &= -ig\tilde{c}_g \exp[-i(\omega - \omega_0)t], \\ \partial_t\tilde{c}_g &= -ig^*\tilde{c}_e \exp[i(\omega - \omega_0)t].\end{aligned} \quad (109)$$

To decouple these equations, we first integrate the equation for $\tilde{c}_g$:

$$\tilde{c}_g(t) = -ig^* \int_0^t dt'\,\tilde{c}_e(t') \exp[i(\omega - \omega_0)t']. \quad (110)$$

Substituting this into the equation for $\tilde{c}_e$,

$$\partial_t\tilde{c}_e = -|g|^2 \int_0^t dt'\,\tilde{c}_e(t') \exp[-i(\omega - \omega_0)(t - t')], \quad (111)$$

which gives the evolution for the excited state coupled to a single field mode.

Now we need to sum over all field modes. In free space, we can integrate over all possible plane waves, labelled by the wave vector $\mathbf{k}$ and the two possible polarizations ζ for each wave vector. Each mode has a different frequency $\omega_{\mathbf{k}} = ck$, and we must expand the basis so that a photon can be emitted into any mode:

$$|\psi\rangle = c_e |e\rangle + \sum_{\mathbf{k},\zeta} c_{\mathbf{k},\zeta} |g, 1_{\mathbf{k},\zeta}\rangle. \tag{112}$$

Putting in the proper form of the coupling constants $g_{\mathbf{k}}$ for each mode in the free-space limit, it turns out that the equation of motion becomes

$$\partial_t \tilde{c}_e = -\frac{d_{ge}^2}{6\epsilon_0 \hbar (2\pi)^3} \sum_\zeta \int d\mathbf{k} \, \omega_{\mathbf{k}} \int_0^t dt' \tilde{c}_e(t') \\ \times \exp\left[-i(\omega_{\mathbf{k}} - \omega_0)(t - t')\right], \tag{113}$$

where $\mathbf{d}_{ge} := \langle g|\mathbf{d}|e\rangle$ is the dipole matrix element characterizing the atomic transition strength. The polarization sum simply contributes a factor of 2, while carrying out the angular integration in spherical coordinates gives

$$\partial_t \tilde{c}_e = -\frac{d_{ge}^2}{6\pi^2 \epsilon_0 \hbar c^3} \int_0^\infty d\omega \, \omega^3 \int_0^t dt' \tilde{c}_e(t') \\ \times \exp\left[-i(\omega_{\mathbf{k}} - \omega_0)(t - t')\right]. \tag{114}$$

We can now note that $\tilde{c}_e(t')$ varies slowly on optical time scales. Also, ω^3 is slowly varying compared to the exponential factor in equation (114), which oscillates rapidly (at least for large times t) about zero except when $t \approx t'$ and $\omega \approx \omega_0$. Thus, we will get a negligible contribution from the ω integral away from $\omega = \omega_0$. We will therefore make the replacement $\omega^3 \to \omega_0^3$:

$$\partial_t \tilde{c}_e = -\frac{\omega_0^3 d_{ge}^2}{6\pi^2 \epsilon_0 \hbar c^3} \int_0^\infty d\omega \int_0^t dt' \tilde{c}_e(t') \\ \times \exp\left[-i(\omega_{\mathbf{k}} - \omega_0)(t - t')\right]. \tag{115}$$

The same argument gives

$$\int_0^\infty d\omega \exp\left[-i(\omega_{\mathbf{k}} - \omega_0)(t - t')\right] \\ \approx \int_{-\infty}^\infty d\omega \exp\left[-i(\omega_{\mathbf{k}} - \omega_0)(t - t')\right] = 2\pi\delta(t - t'). \tag{116}$$

We can see from this that our argument here about the exponential factor is equivalent to the Markovian approximation, where we assume that the time derivative of the quantum state depends only on the state at the present time. Thus,

$$\partial_t \tilde{c}_e = -\frac{\omega_0^3 d_{ge}^2}{3\pi \epsilon_0 \hbar c^3} \int_0^t dt' \tilde{c}_e(t')\delta(t - t') \\ = -\frac{\omega_0^3 d_{ge}^2}{3\pi \epsilon_0 \hbar c^3} \frac{\tilde{c}_e(t)}{2}. \tag{117}$$

Here, we have split the δ-function since the upper limit of the t' integral was t, in view of the original form (115) for the t' integral, where the integration limit is centred at the peak of the exponential factor. We can rewrite the final result as

$$\partial_t \tilde{c}_e = -\frac{\Gamma}{2} \tilde{c}_e, \tag{118}$$

where the spontaneous decay rate is given by

$$\Gamma := \frac{\omega_0^3 d_{ge}^2}{3\pi \epsilon_0 \hbar c^3}. \tag{119}$$

This decay rate is of course defined so that the *probability* decays exponentially at the rate Γ. Also, note that

$$\partial_t c_e = \left(-i\omega_0 - \frac{\Gamma}{2}\right) c_e \tag{120}$$

after transforming out of the slow variables.

10.1.2. Form of the master equation.

We now want to consider the *reduced* density operator for the evolution of the atomic state, tracing over the state of the field. Here we will compute the individual matrix elements

$$\rho_{\alpha\beta} := \langle \alpha|\rho|\beta\rangle \tag{121}$$

for the atomic state.

The easiest matrix element to treat is the excited-level population,

$$\rho_{ee} = c_e c_e^*. \tag{122}$$

Differentiating this equation and using (118) gives

$$\partial_t \rho_{ee} = -\Gamma \rho_{ee}. \tag{123}$$

The matrix element for the ground-state population follows from summing over all the other states:

$$\rho_{gg} := \sum_\zeta \int d\mathbf{k} \, \tilde{c}_{\mathbf{k},\zeta} \tilde{c}_{\mathbf{k},\zeta}^*. \tag{124}$$

Notice that the states $|e\rangle$ and $|g\rangle$ are effectively degenerate, but when we eliminate the field, we want $|e\rangle$ to have $\hbar\omega_0$ more energy than the ground state. The shortcut for doing this is to realize that the latter situation corresponds to the 'interaction picture' with respect to the field, where we use the slowly varying ground-state amplitudes $\tilde{c}_{\mathbf{k},\zeta}$ but the standard excited-state amplitude c_e. This explains why we use regular coefficients in equation (122) but the slow

variables in equation (124). Since by construction $\rho_{ee} + \rho_{gg} = 1$,

$$\partial_t \rho_{gg} = \Gamma \rho_{ee}. \tag{125}$$

Finally, the coherences are

$$\rho_{ge} := \sum_\zeta \int d\mathbf{k} \tilde{c}_{\mathbf{k},\zeta} c_e^*, \quad \rho_{eg} = \rho_{ge}^*, \tag{126}$$

and so the corresponding equation of motion is

$$\partial_t \rho_{ge} = \sum_\zeta \int d\mathbf{k} \tilde{c}_{\mathbf{k},\zeta} \left(i\omega_0 - \frac{\Gamma}{2} \right) c_e^* = \left(i\omega_0 - \frac{\Gamma}{2} \right) \rho_{ge}. \tag{127}$$

We have taken the time derivatives of the $\tilde{c}_{\mathbf{k},\zeta}$ to be zero here. From equation (109), the time derivatives, when summed over all modes, will in general correspond to a sum over amplitudes with rapidly varying phases, and thus their contributions will cancel.

Notice that what we have derived are exactly the same matrix elements generated by the master equation

$$\partial_t \rho = -\frac{i}{\hbar}[H_A, \rho] + \Gamma \mathcal{D}[\sigma]\rho, \tag{128}$$

where the form of $\mathcal{D}[\sigma]\rho$ is given by equation (85), and the atomic Hamiltonian is

$$H_A := \hbar \omega_0 |e\rangle\langle e|. \tag{129}$$

That is, the damping term here represents the same damping as in the optical Bloch equations.

10.2 *Photodetection: quantum jumps and the Poisson process*

In deriving equation (128), we have ignored the state of the field. Now we will consider what happens when we measure it. In particular, we will assume that we make *projective* measurements of the field photon number in every mode, not distinguishing between photons in different modes. It is this extra interaction that will yield the continuous measurement of the atomic state.

From equation (123), the transition probability in a time interval of length dt is $\Gamma \rho_{ee} dt = \Gamma \langle \sigma^\dagger \sigma \rangle dt$, where we recall that $\sigma^\dagger \sigma = |e\rangle\langle e|$ is the excited-state projection operator. Then assuming an ideal detector that detects photons at all frequencies, polarizations and angles, there are two possibilities during this time interval.

(1) *No photon detected.* The detector does not 'click' in this case, and this possibility happens with probability $1 - \Gamma \langle \sigma^\dagger \sigma \rangle dt$. The same construction as above for the master equation carries through, so we keep the equations of motion for ρ_{ee}, ρ_{eg} and ρ_{ge}. However, we do not keep the same equation for ρ_{gg}: no photo-detection implies that the atom does not return to the ground state. Thus, $\partial_t \rho_{gg} = 0$. This case is thus generated by the master equation

$$\partial_t \rho = -\frac{i}{\hbar}[H_A, \rho] - \frac{\Gamma}{2}[\sigma^\dagger \sigma, \rho]_+. \tag{130}$$

This evolution is unnormalized since $\mathrm{Tr}[\rho]$ decays to zero at long times. We can remedy this by explicitly renormalizing the state $\rho(t + dt)$, which amounts to adding one term to the master equation, as in equation (88):

$$\partial_t \rho = -\frac{i}{\hbar}[H_A, \rho] - \frac{\Gamma}{2}[\sigma^\dagger \sigma, \rho]_+ + \Gamma \langle \sigma^\dagger \sigma \rangle \rho. \tag{131}$$

(2) *Photon detected.* A click on the photodetector occurs with probability $\Gamma \langle \sigma^\dagger \sigma \rangle dt$. The interaction Hamiltonian H_{AF} contains a term of the form $\sigma a^\dagger$, which tells us that photon creation (and subsequent detection) is accompanied by lowering of the atomic state. Thus, the evolution for this time interval is given by the reduction

$$\rho(t + dt) = \frac{\sigma \rho(t) \sigma^\dagger}{\langle \sigma^\dagger \sigma \rangle}. \tag{132}$$

We can write this in differential form as

$$d\rho = \frac{\sigma \rho \sigma^\dagger}{\langle \sigma^\dagger \sigma \rangle} - \rho. \tag{133}$$

The overall evolution is stochastic, with either case occurring during a time interval dt with the stated probabilities.

We can explicitly combine these two probabilities by defining a stochastic variable dN, called the *Poisson process*. In any given time interval dt, dN is unity with probability $\Gamma \langle \sigma^\dagger \sigma \rangle dt$ and zero otherwise. Thus, we can write the average over all possible stochastic histories as

$$\langle\langle dN \rangle\rangle = \Gamma \langle \sigma^\dagger \sigma \rangle dt. \tag{134}$$

Also, since dN is either zero or one, the process satisfies $dN^2 = dN$. These last two features are sufficient to fully characterize the Poisson process.

Now we can add the two above possible cases together, with a weighting factor of dN for the second case:

$$d\rho = -\frac{i}{\hbar}[H_A, \rho]dt - \frac{\Gamma}{2}[\sigma^\dagger \sigma, \rho]_+ dt + \Gamma \langle \sigma^\dagger \sigma \rangle \rho \, dt$$
$$+ \left(\frac{\sigma \rho \sigma^\dagger}{\langle \sigma^\dagger \sigma \rangle} - \rho \right) dN. \tag{135}$$

It is unnecessary to include a weighting factor of $(1 - dN)$ for the first term, since $dN \, dt = 0$. It is easy to verify that

this master equation is equivalent to the stochastic Schrödinger equation

$$d|\psi\rangle = -\frac{i}{\hbar}H_A|\psi\rangle dt + \frac{\Gamma}{2}\left(\langle\sigma^\dagger\sigma\rangle - \sigma^\dagger\sigma\right)|\psi\rangle dt$$
$$+ \left(\frac{\sigma}{(\langle\sigma^\dagger\sigma\rangle)^{1/2}} - 1\right)|\psi\rangle dN, \qquad (136)$$

again keeping terms to second order and using $dN^2 = dN$. Stochastic Schrödinger equations of this form are popular for simulating master equations, since if the state vector has $O(n)$ components, the density matrix will have $O(n^2)$ components, and thus is much more computationally expensive to solve. If s solutions ('quantum trajectories') of the stochastic Schrödinger equation can be averaged together to obtain a sufficiently accurate solution to the master equation and $s \ll n$, then this Monte Carlo-type method is computationally efficient for solving the master equation. This idea is illustrated in figure 2, which shows quantum trajectories for the two-level atom driven by a field according to the Hamiltonian (169) in section 10.4.1. As many trajectories are averaged together, the average converges to the master-equation solution for the ensemble average. (About 20 000 trajectories are necessary for the Monte Carlo average to be visually indistinguishable from

the master-equation solution on the time scale plotted here.) Note that the 'Rabi oscillations' apparent here are distorted slightly by the nonlinear renormalization term in equation (136) from the usual sinusoidal oscillations in the absence of spontaneous emission. However, the damping rate in figure 2 is small, so the distortion is not visually apparent. 'Unravellings' [33] of this form are much easier to solve computationally than 'quantum-state diffusion' unravellings involving dW. Of course, it is important for more than just a numerical method, since this gives us a powerful formalism for handling photodetection.

To handle the case of photodetectors with less than ideal efficiency η, we simply combine the conditioned and unconditioned stochastic master equations, with weights η and $1 - \eta$, respectively:

$$d\rho = -\frac{i}{\hbar}[H_A, \rho]dt + \eta\frac{\Gamma}{2}\left[\langle\sigma^\dagger\sigma\rangle - \sigma^\dagger\sigma, \rho\right]_+ dt$$
$$+ (1-\eta)\Gamma\mathcal{D}[\sigma]\rho\, dt + \left(\frac{\sigma\rho\sigma^\dagger}{\langle\sigma^\dagger\sigma\rangle} - \rho\right)dN$$
$$= -\frac{i}{\hbar}[H_A, \rho]\, dt + \Gamma\mathcal{D}[\sigma]\rho\, dt + \eta\Gamma\langle\sigma^\dagger\sigma\rangle\rho\, dt$$
$$- \eta\Gamma\sigma\rho\sigma^\dagger\, dt + \left(\frac{\sigma\rho\sigma^\dagger}{\langle\sigma^\dagger\sigma\rangle} - \rho\right)dN. \qquad (137)$$

The Poisson process is modified in this case such that

$$\langle\langle dN\rangle\rangle = \eta\Gamma\langle\sigma^\dagger\sigma\rangle dt \qquad (138)$$

to account for the fact that fewer photons are detected.

10.3 *Imaged detection of fluorescence*

10.3.1. Centre-of-mass dynamics. Now we want to consider how the evolution of the atomic internal state influences the atomic centre-of-mass motion. To account for the external atomic motion, we use the centre-of-mass Hamiltonian

$$H_{CM} = \frac{p^2}{2m} + V(x) \qquad (139)$$

in addition to the internal atomic Hamiltonian H_A. We also need to explicitly include the spatial dependence of the field by letting

$$g_\mathbf{k} \longrightarrow g_\mathbf{k}\exp(i\mathbf{k}\cdot\mathbf{r}) \qquad (140)$$

in the interaction Hamiltonian (105). In the weak-excitation limit, we can take $\mathbf{k}$ to have the value $\mathbf{k}_L$ of an externally applied probe field (the emitted photons are elastically scattered from the incident field).
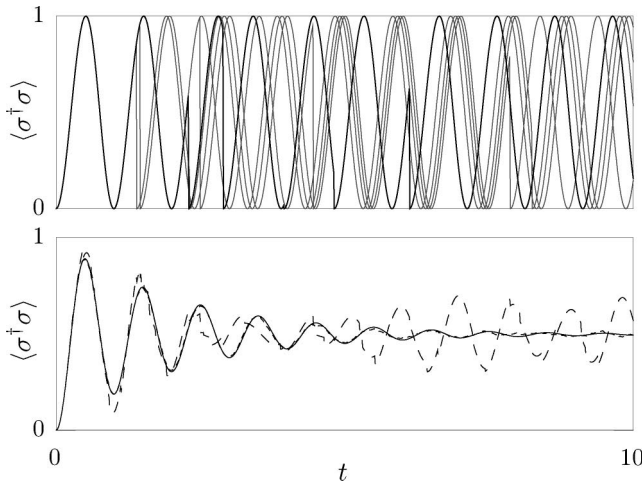
Figure 2. Quantum jumps in a driven two-level atom. Top: evolution of the excited-state probability for a single atom (quantum trajectory) with jumps to the ground state, corresponding to a detected photon. Four other trajectories are included to illustrate the dephasing due to the random nature of the jumps. Bottom: ensemble-averaged excited-state probability computed from the master equation (solid line), an average of 20 trajectories (dashed line), and an average of 2000 trajectories (dotted line). Time is measured in units of $2\pi/\Omega$ (see equation (169)) and the decay rate is $\Gamma = 0.1$ in the same units.

To include the centre of mass in the atomic state, we can explicitly write the state in terms of momentum-dependent coefficients as

$$|\psi\rangle = \int d\mathbf{p}\psi_e(\mathbf{p})|\mathbf{p},e\rangle + \sum_{\mathbf{k},\zeta} \psi_{\mathbf{k},\zeta}(\mathbf{p})|\mathbf{p},g,1_{\mathbf{k},\zeta}\rangle. \quad (141)$$

Notice that the new interaction Hamiltonian

$$H_{AF} = \sum_{\mathbf{k},\zeta} \hbar(g_{\mathbf{k},\zeta}a_{\mathbf{k},\zeta}\sigma^\dagger \exp(i\mathbf{k}\cdot\mathbf{r}) + g_{\mathbf{k},\zeta}^* a_{\mathbf{k},\zeta}^\dagger \sigma \exp(-i\mathbf{k}\cdot\mathbf{r}))$$

$$(142)$$

couples the state $|\mathbf{p},e\rangle$ to the states $|\mathbf{p}-\hbar\mathbf{k},g,1_{\mathbf{k}},\zeta\rangle$ (in the momentum basis), giving rise to the atomic momentum recoil from spontaneous emission. (The additional recoil due to the absorption of the photon comes about by examining the coupling to the driving field.) The derivation of the last section carries through here with the replacement

$$\sigma \to \sigma \exp(-i\mathbf{k}_L \cdot \mathbf{r}). \quad (143)$$

Summing over all possible emission directions, the unconditioned master equation (128) becomes

$$\partial_t \rho = -\frac{i}{\hbar}[H_A + H_{CM}, \rho] + \Gamma \int d\Omega f(\theta,\phi)\mathcal{D}\times[\sigma\exp(-i\mathbf{k}_L\cdot\mathbf{r})]\rho,$$

$$(144)$$

where $f(\theta,\phi)$ is the normalized classical angular distribution for the radiated light, which here represents the angular probability distribution for the emitted photons.

Applying the same reasoning here as for the quantum-jump master equation (135), we obtain

$$d\rho = -\frac{i}{\hbar}[H_A + H_{CM}, \rho]dt + \frac{\Gamma}{2}[\langle\sigma^\dagger\sigma\rangle - \sigma^\dagger\sigma, \rho]_+ dt$$

$$+ \int d\Omega \left(\frac{\sigma\exp(-i\mathbf{k}\cdot\mathbf{r})\rho\sigma^\dagger\exp(i\mathbf{k}\cdot\mathbf{r})}{\langle\sigma^\dagger\sigma\rangle} - \rho\right)\frac{dN(\theta,\phi)}{d\Omega},$$

$$(145)$$

where

$$\left\langle\!\left\langle\frac{dN(\theta,\phi)}{d\Omega}\right\rangle\!\right\rangle = \Gamma\langle\sigma^\dagger\sigma\rangle f(\theta,\phi)dt \quad (146)$$

as before. We can simplify this equation by carrying out the angular integral, defining dN to be one whenever $\max[dN(\theta,\phi)] = 1$. The result is

$$d\rho = -\frac{i}{\hbar}[H_A + H_{CM}, \rho]dt + \frac{\Gamma}{2}[\langle\sigma^\dagger\sigma\rangle - \sigma^\dagger\sigma, \rho]_+ dt$$

$$+ \left(\frac{\sigma\exp(-i\mathbf{k}_L\cdot\mathbf{r})\rho\sigma^\dagger\exp(i\mathbf{k}\cdot\mathbf{r})}{\langle\sigma^\dagger\sigma\rangle} - \rho\right)dN \quad (147)$$

with

$$\langle\!\langle dN\rangle\!\rangle = \Gamma\langle\sigma^\dagger\sigma\rangle\,dt \quad (148)$$

as before. The angles θ and ϕ are then stochastic variables with probability density $f(\theta,\phi)\sin\theta$.

10.3.2. Imaging. The above master equation (145) is for an *angle-resolving* detector. What we see is that angle-resolved detection keeps explicit track of the atomic momentum kicks due to spontaneous emission. An *imaging* detector, on the other hand, gives up resolution of the direction of the emitted photon wave vector $\mathbf{k}$, thus obtaining instead some position information about the atom. An imaging system operates by summing fields from many directions together and then detecting the resulting interference pattern. The procedure for obtaining the measurement operators for the imaging system is as follows [53,54]. Notice that we can regard the master equation (145) as a normal jump process of the form (135), with measurement operators

$$\sigma(\theta,\phi) = [f(\theta,\phi)]^{1/2}\sigma\exp(ik_L z\cos\theta), \quad (149)$$

where we sum over all possible emission angles. In writing down this operator, we are specializing to one-dimensional motion along the z axis ($x = y = 0$), so we only require the z component $k\cos\theta$ of $\mathbf{k}$. This operator ranges from -1 to 1 in $\cos\theta$ and from 0 to 2π in ϕ. Thus, we can write down Fourier coefficients, since these functions are defined on a bounded domain, with two indices α and β:

$$\tilde{\sigma}_{\alpha\beta} = \frac{\sigma}{(4\pi)^{1/2}}\int_0^{2\pi}d\phi\int_{-1}^1 d(\cos\theta)[f(\theta,\phi)]^{1/2}\exp(ik_L z\cos\theta)$$

$$\times \exp(-i\alpha\pi\cos\theta)\exp(-i\beta\phi). \quad (150)$$

If we consider an atom whose radiation pattern is axially symmetric, then performing the ϕ integral amounts to letting $f(\theta,\phi) \to f(\theta)/2\pi$, since the integral is nonzero only for $\beta = 0$. Carrying this out and suppressing the β dependence,

$$\tilde{\sigma}_\alpha = \frac{\sigma}{2^{1/2}}\int_{-1}^1 d(\cos\theta)[f(\theta)]^{1/2}\exp[ik_L(z-\alpha\lambda/2)\cos\theta]. \quad (151)$$

Notice that with the normalization convention for the Fourier coefficients here,

$$\int d\Omega\sigma^\dagger(\theta,\phi)\sigma(\theta,\phi) = \sum_\alpha \tilde{\sigma}_\alpha^\dagger\tilde{\sigma}_\alpha, \quad (152)$$

so that the set of measurement operators is complete and properly normalized in either basis.

Notice that the $\tilde{\sigma}_\alpha$ operators contain localized functions of the position z, and thus correspond to position measurements. For example, a radiating atomic dipole oriented along the z axis has

$$f(\theta) = \frac{3}{4}\sin^2\theta, \tag{153}$$

which gives measurement operators of the form

$$\tilde{\sigma}_\alpha = \sigma\left(\frac{3\pi^2}{8}\right)^{1/2}\frac{J_1(k_L z_\alpha)}{k_L z_\alpha}, \tag{154}$$

where $z_\alpha := z - \alpha\lambda/2$ and $J_1(x)$ is an ordinary Bessel function. Notice also that the set of possible measurement values is not continuous, but rather is discretely spaced by $\lambda/2$.

10.3.3. Gaussian aperture. For the ideal imaging system we have considered here, the aperture extends over the full 4π solid angle (requiring, for example, arbitrarily large lenses on either side of the atom), though in practice it is rare to come anywhere close to this extreme. Thus, we will include the effects of an aperture that only allows the imaging system to detect radiated light within a limited solid angle (figure 3). For mathematical convenience, we will choose an aperture with a Gaussian spatial profile. We consider the above case of motion along the z axis, with the atomic dipole oriented along the z axis. Then photons going into *any* azimuthal angle ϕ are equivalent as far as providing position information about the atom, since the form of $\sigma(\theta,\phi)$ is independent of ϕ. Thus, it suffices to consider only the θ dependence of the aperture, as any ϕ dependence
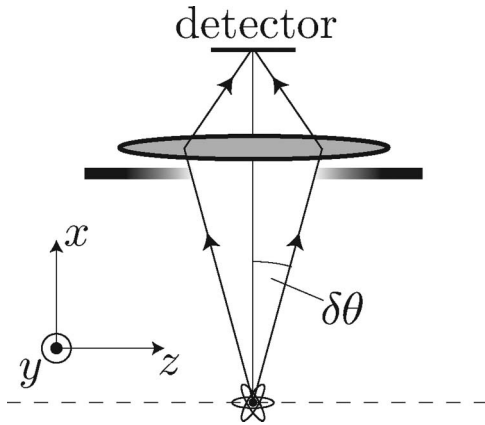

Figure 3. Basic setup for imaging resonance fluorescence from a single atom as a continuous position measurement. Light scattered from a probe laser (not shown) is collected by a Gaussian aperture of angular half-width $\delta\theta$ and focused by a lens on a position-sensitive detector, such as a photodiode array. The atom is constrained to move along the z axis.

contributes only by reducing the effective detection efficiency of the photodetector. Intuitively, one expects a camera imaging system to be most effective when oriented normal to the z axis, so we choose the aperture to be centred about $\theta = \pi/2$. We thus take the *intensity* transmission function of the aperture to be

$$T(\theta) = \exp\left[-\frac{2(\theta - \pi/2)^2}{(\delta\theta)^2}\right]. \tag{155}$$

The generalization of equation (151) to this case is

$$\tilde{\sigma}_\alpha = \frac{\sigma}{2^{1/2}}\int_{-1}^{1}d(\cos\theta)[T(\theta)f(\theta)]^{1/2}\exp\left[ik_L(z - \alpha\lambda/2)\cos\theta\right]. \tag{156}$$

If $\delta\theta$ is small, then the integrand is only appreciable for θ near $\pi/2$ due to the Gaussian factor. Recentring the integrand, making the small-angle approximation in the rest of the integrand, and extending the limits of integration, we find

$$
\begin{aligned}
\tilde{\sigma}_\alpha &= \sigma\left(\frac{3}{8}\right)^{1/2}\int_{-\pi/2}^{\pi/2}d\theta\cos^2\theta \\
&\quad \times \exp\left[-ik_L(z - \alpha\lambda/2)\sin\theta\right] \\
&\quad \times \exp\left[-\frac{\theta^2}{(\delta\theta)^2}\right] \\
&\approx \sigma\left(\frac{3}{8}\right)^{1/2}\int_{-\infty}^{\infty}d\theta\exp\left[-ik_L(z - \alpha\lambda/2)\theta\right]\exp\left[-\frac{\theta^2}{(\delta\theta)^2}\right] \\
&= \sigma\left(\frac{3\pi}{8}\right)^{1/2}\delta\theta\exp\left[-\left(\frac{k_L\delta\theta}{2}\right)^2\left(z - \frac{\alpha\lambda}{2}\right)^2\right].
\end{aligned}
\tag{157}
$$

Thus, the measurement operator in this case is actually Gaussian. We can write the fraction of photons transmitted by the aperture as an efficiency

$$\eta_\theta := \int_{-1}^{1}d(\cos\theta)T(\theta)f(\theta) \approx \frac{3}{4}\left(\frac{\pi}{2}\right)^{1/2}\delta\theta \tag{158}$$

in the same regime of small $\delta\theta$. Then the Gaussian measurement operators $\tilde{\sigma}_\alpha$ satisfy

$$\sum_\alpha \tilde{\sigma}_\alpha^\dagger\tilde{\sigma}_\alpha = \eta_\theta\sigma^\dagger\sigma. \tag{159}$$

This normalization is sensible, although as we will see later, η_θ turns out *not* to be the actual measurement efficiency.

10.3.4. Spatial continuum approximation. If an atom is initially completely delocalized, after one photon is

detected and the collapse operator $\tilde{\sigma}_\alpha$ applies, the atom is reduced to a width of order

$$\delta\alpha = \frac{1}{k_L \delta\theta} = \frac{\lambda}{2\pi\delta\theta}. \tag{160}$$

Since this is much larger than the spacing

$$\Delta\alpha = \frac{\pi}{k_L} = \frac{\lambda}{2}, \tag{161}$$

it is effectively impossible to 'see' the discreteness of the measurement record, and it is a good approximation to replace the set of measurement operators with a set corresponding to a continuous range of possible measurement outcomes. Since in the limit of small spacing Δx, it is a good approximation to write an integral as a sum

$$\sum_n f(n\Delta x)\Delta x = \int dx f(x) \tag{162}$$

for an arbitrary function $f(x)$, we can make the formal identification

$$\tilde{\sigma}_\alpha \longrightarrow \frac{\tilde{\sigma}(\alpha)}{(\Delta\alpha)^{1/2}} \tag{163}$$

to obtain the continuum limit of the position collapse operators. Thus, we have

$$\tilde{\sigma}(\alpha) = \int dz |z\rangle\langle z| \sigma\eta_\theta^{1/2} \frac{1}{[(2\pi)^{1/2}\delta\alpha]^{1/2}} \exp\left[-\frac{(z-\alpha)^2}{4(\delta\alpha)^2}\right]. \tag{164}$$

We have inserted the identity here to make this expression a proper operator on the atomic centre-of-mass state. Again, α is now a continuous index with dimensions of length, rather than an integer index.

Thus, from the form of equation (137), we can deduce the following form of the master equation for imaged photodetection through the Gaussian aperture:

$$d\rho = -\frac{i}{\hbar}[H_A + H_{CM}, \rho]dt + \Gamma \int d\Omega \mathcal{D}[\sigma(\theta,\phi)]\rho\, dt$$
$$+ \eta_\theta \Gamma\langle\sigma^\dagger\sigma\rangle\rho\, dt$$
$$- \Gamma \int d\Omega T(\theta)\sigma(\theta,\phi)\rho\sigma^\dagger(\theta,\phi)dt$$
$$+ \left[\frac{\tilde{\sigma}(\alpha)\rho\tilde{\sigma}^\dagger(\alpha)}{\langle\tilde{\sigma}^\dagger(\alpha)\tilde{\sigma}(\alpha)\rangle} - \rho\right]dN. \tag{165}$$

Recalling the normalization

$$\int d\Omega T(\theta)\sigma^\dagger(\theta,\phi)\sigma(\theta,\phi) = \int d\alpha\tilde{\sigma}^\dagger(\alpha)\tilde{\sigma}(\alpha) = \eta_\theta\sigma^\dagger\sigma, \tag{166}$$

we have for the Poisson process

$$\langle\!\langle dN\rangle\!\rangle = \Gamma\, dt \int d\alpha\langle\tilde{\sigma}^\dagger(\alpha)\tilde{\sigma}(\alpha)\rangle = \eta_\theta\Gamma\langle\sigma^\dagger\sigma\rangle\, dt. \tag{167}$$

Again, α is a random real number corresponding to the result of the position measurement for a given spontaneous emission event. The probability density for α is

$$P(\alpha) = \frac{\langle\tilde{\sigma}^\dagger(\alpha)\tilde{\sigma}(\alpha)\rangle}{\eta_\theta\langle\sigma^\dagger\sigma\rangle}$$
$$= \frac{1}{\langle\sigma^\dagger\sigma\rangle}\int dz|\psi_e(z)|^2 \frac{1}{(2\pi)^{1/2}\delta\alpha}\exp\left[-\frac{(z-\alpha)^2}{2(\delta\alpha)^2}\right], \tag{168}$$

that is, in the case of a localized atomic wave packet, a Gaussian probability density with variance $(\delta\alpha)^2$.

10.4 *Adiabatic approximation*

So far, we have seen how the internal and external dynamics of the atom are intrinsically linked. Now we would like to focus on the external atomic dynamics. To do so, we will take advantage of the natural separation of time scales of the dynamics. The internal dynamics are damped at the spontaneous emission rate Γ, which is typically on the order of $\sim 10^7$ s^{-1}. The external dynamics are typically much slower, corresponding to kHz or smaller oscillation frequencies for typical laser dipole traps. The adiabatic approximation assumes that the internal dynamics equilibrate rapidly compared to the external dynamics, and are thus always in a quasi-equilibrium state with respect to the external state.

10.4.1. Internal quasi-equilibrium. In treating the internal dynamics, we have noted that the atom decays, but not why it was excited in the first place. A resonant, driving (classical) laser field enters in the form [55]

$$H_D = \frac{\hbar\Omega}{2}(\sigma + \sigma^\dagger), \tag{169}$$

where the *Rabi frequency* Ω characterizes the strength of the laser–atom interaction. In writing down this interaction, we have implicitly made the standard unitary transformation to a rotating frame where $H_A = 0$. We have also assumed the driving field propagates along a normal to the z axis, so we have not written any spatial dependence of the field in H_D.

The usual unconditioned master equation with this interaction, but neglecting the external motion (that is equivalent to the usual, on-resonance optical Bloch equations) is

$$\partial_t\rho = -\frac{i}{\hbar}[H_D, \rho] + \Gamma\mathcal{D}[\sigma]\rho. \tag{170}$$

This equation implies that the expectation value of an operator A evolves as

$$\partial_t \langle A \rangle = -\frac{i}{\hbar} \langle [A, H_D] \rangle + \Gamma \left\langle \sigma^\dagger A \sigma - \frac{1}{2} [\sigma^\dagger \sigma, A]_+ \right\rangle. \quad (171)$$

This gives the following equations of motion for the density-matrix elements:

$$\partial_t \rho_{ee} = \partial_t \langle \sigma^\dagger \sigma \rangle = \frac{i\Omega}{2} (\langle \sigma \rangle - \langle \sigma^\dagger \rangle) - \Gamma \langle \sigma^\dagger \sigma \rangle,$$
$$\partial_t \rho_{eg} = \partial_t \langle \sigma \rangle = \frac{i\Omega}{2} (\langle \sigma^\dagger \sigma \rangle - \langle \sigma \sigma^\dagger \rangle) - \frac{\Gamma}{2} \langle \sigma \rangle. \quad (172)$$

The remaining matrix elements are determined by $\rho_{ge} = \rho_{eg}^*$ and $\rho_{gg} = \langle \sigma \sigma^\dagger \rangle = 1 - \langle \sigma^\dagger \sigma \rangle$. Setting the time derivatives to zero, we can solve these equations to obtain

$$\langle \sigma^\dagger \sigma \rangle \beta = \frac{\Omega^2/\Gamma^2}{1 + 2\Omega^2/\Gamma^2},$$
$$\langle \sigma \rangle \beta = \frac{-i\Omega/\Gamma}{1 + 2\Omega^2/\Gamma^2}, \quad (173)$$

for the internal steady-state of the atom.

10.4.2. External master equation. To make the adiabatic approximation and eliminate the internal dynamics, we note that there is no effect on the external dynamics apart from the slow centre-of-mass motion in the potential $V(x)$ and the collapses due to the detection events. When the internal timescales damp much more quickly than the external timescales, we can make the replacement

$$\langle \sigma^\dagger \sigma \rangle \rightarrow \langle \sigma^\dagger \sigma \rangle_{ss} \quad (174)$$

in the master equation (165). Also, in steady state, the internal equations of motion (172) give

$$\langle \sigma^\dagger \sigma \rangle = \frac{\Omega^2/\Gamma^2}{1 + \Omega^2/\Gamma^2} \langle \sigma \sigma^\dagger \rangle, \quad (175)$$

so that the ground- and excited-state populations are proportional. When we also account for the atomic spatial dependence, this argument applies at each position z, so that we can write

$$|\psi_e(z)|^2 = \frac{\Omega^2/\Gamma^2}{1 + \Omega^2/\Gamma^2} |\psi_g(z)|^2, \quad (176)$$

where we are using the general decomposition

$$\langle z|\psi \rangle = \psi_e(z)|e\rangle + \psi_g(z)|g\rangle \quad (177)$$

for the atomic state vector. Thus, the spatial profile of the atom is independent of its internal state, so we need not assign multiple wave functions $\psi_g(z)$ and $\psi_e(z)$ to different internal states of the atom.

Furthermore, we will take a partial trace over the internal degrees of freedom by defining the external density operator

$$\rho_{ext} := \langle e|\rho|e\rangle + \langle g|\rho|g\rangle. \quad (178)$$

The result of applying the same partial trace on the master equation is

$$d\rho_{ext} = -\frac{i}{\hbar} [H_{CM}, \rho_{ext}] \, dt$$
$$+ \gamma \int d\Omega [1 - T(\theta)] f(\theta, \phi) \mathcal{D}[\exp(-ik_L z \cos\theta)] \rho_{ext} \, dt$$
$$+ \left[\frac{A(\alpha)\rho_{ext}A^\dagger(\alpha)}{\langle A^\dagger(\alpha)A(\alpha) \rangle} - \rho_{ext} \right] dN, \quad (179)$$

where

$$\langle\langle dN \rangle\rangle = \eta_\theta \gamma \, dt,$$
$$\tilde{\sigma}(\alpha) =: \sigma A(\alpha), \quad (180)$$
$$\gamma := \Gamma \langle \sigma^\dagger \sigma \rangle.$$

The form (179) follows from the fact that the density operator ρ factorizes into external and internal parts, as we saw in equation (177). Also, equation (168) becomes

$$P(\alpha) = \int dz |\psi(z)|^2 \frac{1}{(2\pi)^{1/2} \delta\alpha} \exp\left[-\frac{(z-\alpha)^2}{2(\delta\alpha)^2} \right], \quad (181)$$

where $\psi(z)$ is the effective state-independent wave function for the atom. When the external state is not pure, we simply make the substitution $|\psi(z)|^2 \rightarrow \langle z|\rho_{ext}|\{z\} \rangle$ in equation (181) to handle this.

Now we have what we want: a master equation for the atomic centre-of-mass state that exhibits localizing collapses due to a physical measurement process. What we essentially have is continuous evolution, with the end of each interval of mean length $(\eta_\theta \gamma)^{-1}$ punctuated by a POVM-type reduction of the form $\rho \rightarrow A(\alpha)\rho A^\dagger(\alpha)$. But note that here there is extra disturbance for the amount of information we gain, because the aperture only picks up a fraction of the available information. We will return to this point shortly.

10.5 White-noise limit

We now have a POVM with a form similar to equation (22), but we still have a quantum-jump master equation for a position measurement that does not look like equation (32). However, we can note that the Gaussian form of the collapse operator $A(\alpha)$ is applied to the state after every

time interval of average length $\Delta t = (\eta_\theta \gamma)^{-1}$. In the regime of slow atomic centre-of-mass motion, the collapses come quickly compared to the motion. Then it is a good approximation to take the formal limit $\Delta t \to 0$, while keeping the rate of information gain constant. (Note that the same result arises in homodyne detection, where the emitted light interferes with a strong phase-reference field, without any coarse-graining approximation.)

10.5.1. Quantum-state diffusion. Comparing equation (181) with equation (24), we see that they are the same if we identify

$$4k\Delta t = \frac{1}{2(\delta\alpha)^2}. \tag{182}$$

Note that k here refers to the *measurement strength*, not the wave number k_L of the scattered light. Solving for the measurement strength,

$$k = \frac{\eta_\theta \gamma}{8(\delta\alpha)^2} = \frac{\pi^2 \eta_\theta \gamma (\delta\theta)^2}{2\lambda^2}. \tag{183}$$

Repeating the procedure of section 4, we can take the limit $\Delta t \to 0$ with k fixed. The resulting master equation, in 'quantum-state diffusion' form, is

$$
\begin{aligned}
\mathrm{d}\rho_{\text{ext}} = &-\frac{\mathrm{i}}{\hbar}[H_{\text{CM}}, \rho_{\text{ext}}]\mathrm{d}t \\
&+ \gamma \int \mathrm{d}\Omega [1 - T(\theta)]f(\theta, \phi)\mathcal{D}[\exp(-\mathrm{i}k_L z \cos\theta)]\rho_{\text{ext}}\,\mathrm{d}t \\
&+ 2k\mathcal{D}[z]\rho_{\text{ext}}\,\mathrm{d}t + (2\eta_\phi k)^{1/2}\mathcal{H}[z]\rho_{\text{ext}}\,\mathrm{d}W.
\end{aligned}
\tag{184}
$$

The form here is the same as in equation (32), except for an extra 'disturbance term' representing the undetected photons. We have also added an extra efficiency η_ϕ to model aperturing in the ϕ direction and other effects such as the intrinsic (quantum) efficiency of the imaging detector.

10.5.2. Diffusion rates. To simplify the master equation (184), we will analyse the diffusion rates due to the second and third terms (proportional to γ and k, respectively). From the analysis of equations (97), recall that the term $2k\mathcal{D}[z]\rho_{\text{ext}}\,\mathrm{d}t$ causes diffusion in momentum at the rate

$$D_k = 2\hbar^2 k = \frac{\eta_\theta}{4}\gamma\hbar^2 k_L^2(\delta\theta)^2. \tag{185}$$

This is the disturbance corresponding to the information gain. The relation $k = D_k/(2\hbar^2)$ will be useful below.

We can compute the *total* diffusion rate due to the spontaneously emitted photons as follows. Each photon emission causes a momentum kick of magnitude $\hbar k_L \cos\theta$,

and the spontaneous emission rate is γ. Averaging over the angular photon distribution, the diffusion rate becomes

$$D_{\text{SE}} = \gamma\hbar^2 k_L^2 \int \mathrm{d}\Omega f(\theta, \phi)\cos^2\theta = \frac{\gamma\hbar^2 k_L^2}{5}. \tag{186}$$

On the other hand, the diffusion rate due only to the detected photons is

$$
\begin{aligned}
D_\theta &= \gamma\hbar^2 k_L^2 \int \mathrm{d}\Omega\, T(\theta)f(\theta, \phi)\cos^2\theta \\
&= \gamma\hbar^2 k_L^2 \frac{3}{4}\int_0^\pi \mathrm{d}\theta \sin^3\theta \cos^2\theta \exp\left[-\frac{2(\theta - \pi/2)^2}{(\delta\theta)^2}\right] \\
&\approx \frac{\eta_\theta}{4}\gamma\hbar^2 k_L^2(\delta\theta)^2,
\end{aligned}
\tag{187}
$$

where we used the fact that $\delta\theta$ is small. This is precisely the same rate as D_k, since they are two different representations of the same physical process.

We see now that the second and third terms of equation (184) have the same effect of momentum diffusion, but at different rates. We can formally combine them to obtain

$$
\begin{aligned}
\mathrm{d}\rho_{\text{ext}} = &-\frac{\mathrm{i}}{\hbar}[H_{\text{CM}}, \rho_{\text{ext}}]\mathrm{d}t \\
&+ 2k_{\text{eff}}\mathcal{D}[z]\rho_{\text{ext}}\,\mathrm{d}t + (2\eta_{\text{eff}}k_{\text{eff}})^{1/2}\mathcal{H}[z]\rho_{\text{ext}}\,\mathrm{d}W,
\end{aligned}
\tag{188}
$$

where the effective measurement strength is

$$k_{\text{eff}} = \frac{D_{\text{SE}}}{2\hbar^2} = \frac{\gamma k_L^2}{10}, \tag{189}$$

and the effective measurement efficiency is

$$\eta_{\text{eff}} = \frac{\eta_\phi k}{k_{\text{eff}}} = \frac{5}{4}\eta_\phi\eta_\theta(\delta\theta)^2. \tag{190}$$

Notice that since $\delta\theta$ is assumed small, the *apparent* efficiency η_{eff} derived from comparing the information rate to the disturbance rate, is much smaller than the photon-detection efficiency of $\eta_\phi\eta_\theta$. Evidently, the photons radiated near $\theta = \pi/2$ are much less effective compared to the photons radiated near $\theta = 0$ or π. This result is counter-intuitive when considering typical imaging setups as we have considered here, but suggests that other ways of processing the radiated photons (e.g. measuring the phase of photons radiated closer to the z axis) are more effective than camera-like imaging.

11. Conclusion

We have presented what we hope is a readily accessible introduction to continuous measurements in quantum systems. If you have read and digested most of the above,

you should have a good basic understanding of how to treat such measurements and manipulate the equations that describe them. There is now a considerable literature discussing such measurements in a variety of systems, and here we give a brief overview of this literature so as to provide a pointer to further reading. We have already mentioned that continuous measurement has many applications in areas such as feedback control and metrology, and references on these topics have been given in the introduction. The early pioneering work on continuous measurement may be found in [1,56 – 62]. Derivations of continuous measurements driven by Gaussian noise in quantum-optical systems are given in [5,63,64], and further applications in quantum optics may be found in [16,32,33,65 – 67]. Derivations and applications of stochastic Schrödinger equations with jump (Poisson) processes— developed originally in quantum optics as a tool for the simulation of master equations using the 'Monte Carlo' method, as in section 10.2—may be found in [54,65,66,68 – 74]. A treatment of continuous measurement in a solid-state system is given in [75], and further applications in these systems may be found in [6,76 – 81]. Last, but not least, if the reader is interested in treatments of quantum continuous measurements using the rigorous mathematical language of filtering theory, these may be found in [1,82 – 84]. Other rigorous treatments are given in [42,57].

Acknowledgments

References

[1] V.P. Belavkin, in *Information, Complexity and Control in Quantum Physics*, edited by A. Blaquiere, S. Diner and G. Lochak (Springer-Verlag, New York, 1987).

[2] P. Bushev, D. Rotter, A. Wilson, *et al.*, Phys. Rev. Lett. **96** 043003 (2006).

[3] J. Combes and K. Jacobs, Phys. Rev. Lett. **96** 010504 (2006).

[4] C. D'Helon and M.R. James, Phys. Rev. A **73** 053803 (2006).

[5] A.C. Doherty and K. Jacobs, Phys. Rev. A **60** 2700 (1999).

[6] A. Hopkins, K. Jacobs, S. Habib, *et al.*, Phys. Rev. B **68** 235328 (2003).

[7] P. Rabl, V. Steixner and P. Zoller, eprint: quant-ph/0506185 (2005).

[8] D. Steck, K. Jacobs, H. Mabuchi, *et al.*, Phys. Rev. Lett. **92** 223004 (2004).

[9] D. Steck, K. Jacobs, H. Mabuchi, *et al.*, Phys. Rev. A **74** 012322 (2006).

[10] V. Steixner, P. Rabl and P. Zoller, eprint: quant-ph/0506187 (2005).

[11] H.M. Wiseman and A.C. Doherty, Phys. Rev. Lett. **94** 070405 (2005).

[12] D.W. Berry and H.M. Wiseman, Phys. Rev. A **65** 043803 (2002).

[13] J. Geremia, J.K. Stockton and H. Mabuchi, Phys. Rev. Lett. **94** 203002 (2005).

[14] D.T. Pope, H.M. Wiseman and N.K. Langford, Phys. Rev. A **70**, 043812 (2004).

[15] J.K. Stockton, J.M. Geremia, A.C. Doherty, *et al.*, Phys. Rev. A **69** 032109 (2004).

[16] H.M. Wiseman, Phys. Rev. Lett. **75** 4587 (1995).

[17] S. Dolinar, *Tech. Rep. 111, Research Laboratory of Electronics* (MIT, Cambridge, 1973).

[18] J.M. Geremia, Phys. Rev. A **70** 062303 (2004).

[19] K. Jacobs, Quant. Informat. Comp. **7** 127 (2007).

[20] C. Ahn, A.C. Doherty and A.J. Landahl, Phys. Rev. A **65** 042301 (2002).

[21] M. Sarovar, C. Ahn, K. Jacobs, *et al.*, Phys. Rev. A **69** 052324 (2004).

[22] R. van Handel and H. Mabuchi, eprint: quant-ph/0511221 (2006).

[23] T. Bhattacharya, S. Habib and K. Jacobs, Phys. Rev. Lett. **85** 4852 (2000).

[24] T. Bhattacharya, S. Habib and K. Jacobs, Phys. Rev. A **67** 042103 (2003).

[25] M.J. Everitt, T.D. Clark, P.B. Stiffell, *et al.*, New J. Phys. **7** 64 (2005).

[26] S. Ghose, P. Alsing, I. Deutsch, *et al.*, Phys. Rev. A **69** 052116 (2004).

[27] S. Ghose, P.M. Alsing, B.C. Sanders, *et al.*, Phys. Rev. A **72** 014102 (2005).

[28] S. Habib, K. Jacobs, H. Mabuchi, *et al.*, Phys. Rev. Lett. **88** 040402 (2002).

[29] S. Habib, K. Jacobs and K. Shizume, Phys. Rev. Lett. **96** 010403 (2006).

[30] T.A. Brun, Am. J. Phys. **70** 719 (2002).

[31] V.B. Braginsky, F.Y. Khalili and K.S. Thorne, *Quantum Measurement* (Cambridge University Press, Cambridge, 1995).

[32] H.M. Wiseman, Quant. Semiclass. Opt. **8** 205 (1996).

[33] H. Carmichael, *An Open Systems Approach to Quantum Optics* (Springer-Verlag, Berlin, 1993).

[34] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[35] B. Schumacher, Phys. Rev. A **54** 2614 (1996).

[36] C.A. Fuchs and K. Jacobs, Phys. Rev. A **63** 062305 (2001).

[37] P.E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, Berlin, 1992).

[38] C.M. Caves and J.G. Milburn, Phys. Rev. A **36** 5543 (1987).

[39] C.W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 1985).

[40] D.T. Gillespie, Am. J. Phys. **64** 225 (1996).

[41] K. Jacobs and P.L. Knight, Phys. Rev. A **57** 2301 (1998).

[42] A. Barchielli, Int. J. Theor. Phys. **32** 2221 (1993).

[43] J. Dziarmaga, D.A.R. Dalvit and W.H. Zurek, Phys. Rev. A **69** 022109 (2004).

[44] H.M. Wiseman and G.J. Milburn, Phys. Rev. A **47** 642 (1993).

[45] S.L. Adler, Phys. Lett. A **265** 58 (2000).

[46] H.M. Wiseman and L. Diosi, Chem. Phys. **91** 268 (2001).

[47] G. Lindblad, Comm. Math. Phys. **48** 199 (1976).

[48] S. Habib, eprint: quant-ph/0406011 (2004).

[49] R. Landauer, Physica A **194** 551 (1993).

[50] P.W. Milonni, *The Quantum Vacuum* (Academic Press, New York, 1994).

[51] M.O. Scully and M.S. Zubairy, *Quantum Optics* (Cambridge University Press, Cambridge, 1997).

[52] V. Weisskopf and E. Wigner, Z. Phys. **63** 54 (1930).

[53] W. Greenwood, P. Pax and P. Meystre, Phys. Rev. A **56** 2109 (1997).

[54] M. Holland, S. Marksteiner, P. Marte, *et al.*, Phys. Rev. Lett. **76** 3683 (1996).

[55] R. Loudon, *The Quantum Theory of Light,* 2nd edition (Oxford University Press, Oxford, 1983).

[56] A. Barchielli, L. Lanz and G.M. Prosperi, Nuovo Cimento B **72** 79 (1982).

[57] A. Barchielli and G. Lupieri, J. Math. Phys. **26** 2222 (1985).

[58] V.P. Belavkin, Phys. Lett. A **140** 355 (1989).

[59] L. Diosi, Phys. Lett. A **114** 451 (1986).

[60] L. Diosi, Phys. Lett. A **129** 419 (1986).

[61] N. Gisin, Phys. Rev. Lett. **52** 1657 (1984).

[62] M.D. Srinivas and E.B. Davies, Optica Acta **28** 981 (1981).

[63] H.J. Carmichael, S. Singh, R. Vyas, *et al.*, Phys. Rev. A **39** 1200 (1989).

[64] H.M. Wiseman and G.J. Milburn, Phys. Rev. A **47** 642 (1993).

[65] B.M. Garraway and P.L. Knight, Phys. Rev. A **50** 2548 (1994).

[66] M.B. Plenio and P.L. Knight, Rev. Mod. Phys. **70** 101 (1998).

[67] H.M. Wiseman, Phys. Rev. A **47** 5180 (1993).

[68] J. Dalibard, Y. Castin and K. Mølmer, Phys. Rev. Lett. **68** 580 (1992).

[69] C.W. Gardiner, A.S. Parkins and P. Zoller, Phys. Rev. A **46** 4363 (1992).

[70] G.C. Hegerfeldt and M.B. Plenio, Phys. Rev. A **53** 1164 (1996).

[71] G.C. Hegerfeldt and T.S. Wilser, in *Proceedings of the II International Wigner Symposium*, edited by H.D. Doebner, W. Scherer and F. Schroeck (World Scientific, Singapore, 1991), p. 104.

[72] M.J. Holland, D.F. Walls and P. Zoller, Phys. Rev. Lett. **67** 1716 (1991).

[73] M.B. Plenio, Novel structures in the emission spectrum of laser driven atoms, PhD thesis, University of Göttingen (1994).

[74] P. Zoller, M. Marte and D.F. Walls, Phys. Rev. A **35** 198 (1987).

[75] A.N. Korotkov, Phys. Rev. B **63** 115403 (2001).

[76] T.A. Brun and H.-S. Goan, Phys. Rev. A **68** 032301 (2003).

[77] H.-S. Goan and G.J. Milburn, Phys. Rev. B **64** 235307 (2001).

[78] A.N. Jordan and A.N. Korotkov, Phys. Rev. B **74** 085307 (2006).

[79] R. Ruskov, K. Schwab and A.N. Korotkov, Phys. Rev. B **71** 235407 (2005).

[80] D.H. Santamore, A.C. Doherty and M.C. Cross, Phys. Rev. B **70** 144301 (2004).

[81] M. Sarovar, H.-S. Goan, T.P. Spiller, *et al.*, Phys. Rev. A **72** 062327 (2005).

[82] V.P. Belavkin, Prob. Theory App. **38** 742 (1993).

[83] V.P. Belavkin, Prob. Theory App. **39** 640 (1994).

[84] L. Bouten, R. van Handel and M.R. James, eprint: math.PR/0606118 (2006).

Kurt Jacobs is an Assistant Professor of Physics at the University of Massachusetts at Boston. He completed a master's degree in physics with Dan Walls at Auckland University in 1995, and a PhD with Peter Knight at Imperial College, London, in 1998. During his time at Auckland he was introduced to classical continuous measurement theory by Sze Tan, and was exposed to the recently developed field of quantum continuous measurement particularly in interactions with Howard Wiseman who was then a postdoc at Auckland. He continued to work on this subject during his PhD, and then joined Los Alamos National Laboratory where he worked with Salman Habib on problems in quantum feedback control and the quantum-to-classical transition. His work currently focuses on these two areas, including applications of continuous measurement and control in atom optics and quantum-nano-electro mechanics.

Daniel Steck is an Assistant Professor of Physics at the University of Oregon. He performed his PhD work, involving experiments in quantum chaos and transport, with Mark Raizen at the University of Texas at Austin. He was then a postdoctoral fellow at Los Alamos National Laboratory with Salman Habib, working on the theory of quantum measurement and feedback control. His work now focuses on the study of continuous quantum measurements and the transition between quantum and classical dynamics in ultracold-atom experiments.