1 Positive Operator-Valued Measures

To help handle generalized measurements, we will now introduce the somewhat mathematical concept of a positive operator-valued measure (POVM). By referring to generalized measurements, we mean to differentiate these measurements from the usual projective, or von Neumann, measurements, which is what you normally find in introductory quantum-mechanics texts. The usual description goes like this: for a quantum system in state $|\psi\rangle$, a measurement of the observable $Q$ leaves the system in an eigenstate $|q\rangle$ of $Q$ with probability $\langle\psi|Q|\psi\rangle$, in which case the “result” of the measurement is the eigenvalue $q$. We can see that this notion of a measurement is lacking in two situations. First, it does not properly describe the situation in photodetection of atomic radiation, where each detection event results in the loss of energy (i.e., the atom is always found to be in the ground state), and we gain information even during instants when a photon is not detected. Thus, POVMs are crucial to the formal definition of a continuous measurement. The second situation is when the observable is the position operator, where eigenstate collapse is unphysical: a position eigenstate is a state of infinite energy. POVMs allow us to define an imprecise or partial measurement of an observable, which will be a stepping stone on the way to defining a continuous measurement of position.

1.1 Discrete, Finite Spaces

Consider a discrete, finite Hilbert space of dimension $N$. That is, the Hilbert space is spanned by the set of eigenstates

$$\{ |q\rangle : q = 1, \ldots, N \}$$

of the observable $Q$. Then we can define a positive operator-valued measure (POVM) as a set of positive-semidefinite operators $\Omega_q^\dagger \Omega_q$ that sum to the identity operator:

$$\sum_{q=1}^{N_q} \Omega_q^\dagger \Omega_q = 1.$$  \hspace{1cm} (2)

Note that we are writing the $q$th positive operator $\Omega_q^\dagger \Omega_q$ as a factorization in terms of the Kraus operator $\Omega_q$, since any positive operator always has such a factorization. We also note that the number $N_q$ of positive operators is not necessarily the same as the dimension $N$ of the Hilbert space.

Now the important physical point here is that a POVM defines a quantum measurement on the Hilbert space. The $q$th possible outcome of the measurement is that the state vector changes according to the replacement

$$|\psi\rangle \longrightarrow \frac{\Omega_q |\psi\rangle}{\sqrt{\langle \Omega_q^\dagger \Omega_q \rangle}}.$$  \hspace{1cm} (3)

or in terms of the density operator,

$$\rho \longrightarrow \frac{\Omega_q \rho \Omega_q^\dagger}{\text{Tr}[\Omega_q \rho \Omega_q^\dagger]} = \frac{\Omega_q \rho \Omega_q^\dagger}{\langle \Omega_q^\dagger \Omega_q \rangle}.$$  \hspace{1cm} (4)
That is, in the \( q \)th outcome, the state is “hit” by the operator \( \Omega_q \) and then renormalized if necessary. The probability that the \( q \)th outcome occurs is

\[
P(q) = \text{Tr}[\Omega_q \rho \Omega_q^\dagger] = \langle \Omega_q^\dagger \Omega_q \rangle.
\] (5)

The (classical) “result” of the quantum measurement in this case is simply \( q \) (or some physically meaningful function of \( q \)). This notion may seem rather abstract, but we can note that the usual projective measurement comes out as a special case of the POVM-based measurement. In particular, the usual measurement arises from a projection-valued measure, where we partition the Hilbert space according to a set of (Hermitian) projection operators

\[
P_q := |q\rangle \langle q|
\] (6)

that also sum to the identity:

\[
\sum_{q=1}^{N} P_q^2 = 1.
\] (7)

Of course, \( P_q^2 = P_q \), but we have written the sum in this form to emphasize the similarity with Eq. (2) by taking \( \Omega_q = P_q \) and \( N_q = N \). Then the standard projective measurement of the observable \( Q \) results in the \( q \)th outcome of a reduction to the \( q \)th eigenstate \(|q\rangle\),

\[
|\psi\rangle \rightarrow \frac{P_q |\psi\rangle}{\sqrt{\langle P_q^2 \rangle}} = |q\rangle,
\] (8)

or in terms of the density operator,

\[
\rho \rightarrow \frac{P_q \rho P_q^\dagger}{\text{Tr}[P_q \rho P_q^\dagger]} = P_q \rho P_q^\dagger = |q\rangle \langle q|.
\] (9)

This outcome happens with probability

\[
P(q) = \text{Tr}[P_q \rho P_q^\dagger] = \langle P_q^2 \rangle = \langle P_q \rangle,
\] (10)

which for a pure state \(|\psi\rangle\) becomes the familiar Born rule

\[
P(q) = |\langle q | \psi \rangle|^2.
\] (11)

Thus, the POVM-based measurement above is a reasonably straightforward generalization of the usual projective measurements, at least when the standard measurements are cast in the proper way.

1.2 Measure

Why is a POVM called a “POVM”? The answer requires an excursion into mathematics, and so the short answer, if you feel the need to skip forward, is that a measure is usually something that assigns numbers to sets, and so a positive operator-valued measure is a measure that instead associates positive operators with sets, and thence probabilities to the same sets via the expectation value as above. To really answer this question, we need to define what we usually mean by a measure, and then adapt it to the operator case. Informally, a measure is a rule for assigning numbers to subsets of some set, or space. This is a very useful notion in probability theory, where you would consider the set of all possible outcomes or events, and the measure would assign probabilities to each outcome or collection of outcomes. Alternately, a measure is an abstraction of the notion of volume, where the measure represents the “volume” of subsets of the main set.

Before formally defining a measure, though, we should first note that for a given space, it is problematic to try to define a measure on every subset. Instead, we will define the measure on only a limited collection of subsets, chosen to make the definition of the measure consistent. Formally, this collection is a \( \sigma \)-algebra, which we define as a collection \( \mathcal{S} \) of subsets of the space \( X \) such that:

2
1. The empty set is included: $\emptyset \in \mathcal{F}$.

2. Countable, disjoint unions are included (with countable here meaning finite or countably infinite): if $\mathcal{U} \subset \mathcal{F}$ with $A \cap B = \emptyset$ for any $A, B \in \mathcal{U}$, and $\mathcal{U}$ is countable, then
\[
\bigcup_{A \in \mathcal{U}} A \in \mathcal{F}.
\] (12)

3. Complements are included: if $A \in \mathcal{F}$, then $X - A \in \mathcal{F}$.

Any element of a $\sigma$-algebra is said to be a measurable set. This definition can be contrasted with the possibly familiar definition for a topology on a space $X$, which is a collection $\mathcal{F}$ of subsets of $X$ such that:

1. The empty set and the whole space are included: $\emptyset, X \in \mathcal{F}$.

2. Arbitrary unions are included: if $\mathcal{U} \subset \mathcal{F}$, then
\[
\bigcup_{A \in \mathcal{U}} A \in \mathcal{F}.
\] (13)

3. Finite intersections are included: if $\mathcal{U} \subset \mathcal{F}$ with $\mathcal{U}$ finite, then
\[
\bigcap_{A \in \mathcal{U}} A \in \mathcal{F}.
\] (14)

Any element of a topology is said to be an open set, while the complement of an open set is said to be a closed set. Thus, while topologies contain in general only open sets, $\sigma$-algebras contain both open and closed sets. For example, on the real line $\mathbb{R}$, the standard topology is the topology consisting of all open intervals of the form $(a, b)$ and all possible unions of such intervals (and the empty set). It turns out there is a unique $\sigma$-algebra associated with the standard topology, which is the smallest $\sigma$-algebra containing it. This is called the Borel $\sigma$-algebra on $\mathbb{R}$, which would contain all open intervals as well as all closed intervals of the form $[a, b]$ (and many other sets). The notion of a $\sigma$-algebra may not be intuitively clear at this stage, but the definition is basically concocted to make the definition of measure work out, as we will now see.

A measure is a function $\mu : \mathcal{F} \rightarrow [0, \infty]$ defined on a $\sigma$-algebra $\mathcal{F}$ on a space $X$, which satisfies

1. The empty set has zero measure: $\mu(\emptyset) = 0$.

2. The measure for countable, disjoint unions adds: if $\mathcal{U} \subset \mathcal{F}$ with $A \cap B = \emptyset$ for any $A, B \in \mathcal{U}$, and $\mathcal{U}$ is countable, then
\[
\mu\left(\bigcup_{A \in \mathcal{U}} A\right) = \sum_{A \in \mathcal{U}} \mu(A).
\] (15)

These two requirements are sensible considering our analogies to probabilities and volumes, and we can also see how the requirements for a $\sigma$-algebra guarantee that we don’t have any problems in defining a measure (the last axiom for a $\sigma$-algebra imposes the sensible constraint that if $A$ is a measurable subset, then so is $X - A$). Note that the point $\infty$ is explicitly included in the range of a measure, which is intuitively a “good” measure for something like the entire real line. Also, strictly speaking, we have defined a positive measure, since we have only allowed nonnegative values in the range of $\mu$. As an example of measure, the Lebesgue measure on the real line is defined on the Borel $\sigma$-algebra. We can define it in several cases as follows:

---

1. It turns out that any open set $A$ can be written as the union of a countable set of open intervals $(a_j, b_j)$, in which case the Lebesgue measure of $A$ is the sum of the interval lengths:

$$\mu(A) := \sum_j (b_j - a_j).$$

(16)

2. It turns out that any closed set $B$ can be written as a closed interval $[a, b]$ with the union of a countable set of open intervals $(a_j, b_j)$ removed from it,

$$B = [a, b] - \bigcup_j (a_j, b_j),$$

(17)

where every $a_j > a$ and every $b_j < b$, in which case the Lebesgue measure of $B$ is the length of the closed interval minus the Lebesgue measure of the removed component:

$$\mu(A) := (b - a) - \sum_j (b_j - a_j).$$

(18)

3. For any other set $C$ in the Borel $\sigma$-algebra, the Lebesgue measure is the infimum (greatest lower bound) of the set of Lebesgue measures of all open sets containing $C$:

$$\mu(C) := \inf \{ \mu(A) : A \text{ is open and } C \subset A \}.$$

(19)

Note that there exist sets that do not have Lebesgue measures according to the above definitions, and thus they are excluded by considering only the $\sigma$-algebra. The Lebesgue measure is useful in that it extends the notion of length to more complicated and subtle sets: the set of rational numbers, being countable, is a set of Lebesgue measure zero on the real line; and the Cantor middle-thirds set, a fractal set constructed by starting with the interval $[0, 1]$, removing the open “middle third” interval $(1/3, 2/3)$, removing the middle-thirds of the two remaining closed intervals, and so on $ad$ $infinitum$, is an uncountable set but of zero Lebesgue measure.

For measurements, the concept of a probability measure is more useful, and it is simply that of a measure, but where the range of the measure is $[0, 1]$ rather than $[0, \infty]$, with a measure of the whole space being unity. For example, the Lebesgue measure on the space $[0, 1]$ is a probability measure, and corresponds to a uniform probability density on the same interval.

### 1.3 General Definition

Now with the above mathematical concepts, we can now give a more general definition of a POVM than in the finite case above. In more general terms, a positive operator-valued measure (POVM) defined on a $\sigma$-algebra $\mathcal{S}$ on a space $X$ is a function $\Pi$ that takes as values positive semidefinite, Hermitian operators on a Hilbert space $\mathcal{H}$ such that for any $|\psi\rangle \in \mathcal{H}$, the function $\mu : X \rightarrow [0, 1]$, defined by

$$\mu(A) := \langle \psi | \Pi(A) | \psi \rangle$$

(20)

for any measurable subset $A$ of $X$, defines a probability measure on $\mathcal{S}$. In particular, this implies that $\Pi(X)$ is the identity operator, which is the generalization of the sum rule (2). Thus, the POVM associates positive operators with measurable subsets of the space of outcomes, which are then associated with probabilities by appropriate expectation values. In this way, we can define a family of probability measures, “parameterized” by the quantum state.
We could, of course, write the probability measure more generally in terms of the density operator as
\[ \mu(A) = \text{Tr}[\Pi(A)\rho]. \] (21)

Incidentally, a trace of this form is (for a Hilbert space of dimension larger than two) the only way to construct a quantum probability measure; this is essentially the content of Gleason’s theorem.²

1.4 Realization

It is important to note that measurements induced by POVMs, while generalizing projective measurements, don’t introduce anything fundamentally new to quantum mechanics: any of these more general measurements can be realized by introducing an auxiliary system (ancilla), performing a unitary transformation on the combined system, and then perform a projective measurement on the ancilla. Thus, generalized measurements correspond to indirect measurements, where information about a system comes from projective measurements on the “environment” with which the system has interacted (and thus become entangled with).

This result is known as Naimark’s theorem (or Neumark’s theorem),³ and we will only sketch the argument for the finite case here. Starting with the system in the state \( |\psi\rangle \), we will extend the Hilbert space to contain the environment, whose dimension is equal to the number of Kraus operators defining the POVM, \( |\psi\rangle \rightarrow |\psi\rangle|0_E\rangle \equiv |\psi_0E\rangle \). Note that we are assuming a pure state for the system, which we may as well do as long as we are extending the Hilbert space by invoking purification. We can thus define an operator \( U \) that acts on the composite state as
\[ U|\psi_0E\rangle = \sum_q (\Omega_q|\psi\rangle)|q_E\rangle = \sum_q \sqrt{\langle\psi|\Omega_q\Omega_q^\dagger|\psi\rangle} \left( \frac{\Omega_q|\psi\rangle}{\sqrt{\langle\psi|\Omega_q\Omega_q^\dagger|\psi\rangle}} \right)|q_E\rangle, \] (22)

where the Kraus operators \( \Omega_q \) only operate on the original system. Now computing the norm of the transformed composite state,
\[ \langle\psi_0E|U^\dagger U|\psi_0E\rangle = \sum_{qq'} \langle q_E|\langle\psi|\Omega_q\Omega_q^\dagger|\psi\rangle|q'_E\rangle \] (23)

so that \( U \) preserves the norm of states in the subspace of the original system. The operator \( U \) is thus unitary on this subspace, but is not fixed uniquely by the above argument. In principle, the action of \( U \) on the environment can be chosen to make \( U \) unitary on the composite Hilbert space. Basically, this is because taken as a matrix, the columns of \( U \) span the subspace of the original system (i.e., a subset of them form an orthonormal basis), and the extra degrees of freedom (elements of the extra rows) in expanding \( U \) to the composite Hilbert space may then be chosen to make the columns of \( U \) form an orthonormal basis on the entire composite space. Now after the transformation, a projective measurement of the state of the environment leads to the result \( |q_E\rangle \) with probability
\[ \text{Tr} \left[ |q_E\rangle\langle q_E| U|\psi_0E\rangle\langle\psi_0E|U^\dagger \right] = \langle\psi|\Omega_q\Omega_q^\dagger|\psi\rangle. \] (24)


³Asher Peres, Quantum Theory: Concepts and Methods (Springer, 1995), Section 9-6, p. 285. For a similar argument to what we present here for the unitary representation of linear positive maps, see Benjamin Schumacher, op. cit.
Furthermore, the projection of the environment into state $|qE\rangle$ induces the transformation
\[
|\psi\rangle \longrightarrow \frac{\Omega_q|\psi\rangle}{\sqrt{\langle\psi|\Omega_q^\dagger\Omega_q|\psi\rangle}}
\]
on the original system. Thus we have constructed the POVM-based measurement based on the larger projective measurement.

1.5 Example: Spontaneous Emission

As an example of a POVM, we return to the stochastic master equation for photodetection of atomic resonance fluorescence with quantum jumps:
\[
d\rho = -\frac{i}{\hbar}[H,\rho]dt - \frac{\Gamma}{2}[\sigma^\dagger\sigma,\rho]_+ dt + \Gamma\langle\sigma^\dagger\sigma\rangle\rho dt + \left(\frac{\sigma\rho\sigma^\dagger}{\langle\sigma^\dagger\sigma\rangle} - \rho\right) dN.
\]
In any given time interval of duration $dt$, there are only two possible outcomes: no photon is detected, or one photon is detected. We can define this evolution in terms of a POVM as follows. Let $U(dt)$ denote the evolution operator for the combined atom–field system. Before each infinitesimal time interval, the field starts in the vacuum state $|0\rangle$, and after each infinitesimal time interval, the detector projectively measures the field and registers a detection event if a photon is emitted into any mode. Since the detector does not distinguish modes, we will simply denote the field state as $|1\rangle$ in the case of an emitted photon. Then the two “jump operators” for the two measurement outcomes are
\[
\Omega_0(dt) = \langle 0|U(dt)|0\rangle = 1 - \frac{H}{\hbar} dt - \frac{\Gamma}{2}[\sigma^\dagger\sigma,\rho]_+ dt + \Gamma\langle\sigma^\dagger\sigma\rangle\rho dt + \left(\frac{\sigma\rho\sigma^\dagger}{\langle\sigma^\dagger\sigma\rangle} - \rho\right) dN.
\]
\[
\Omega_1(dt) = \langle 1|U(dt)|0\rangle = \sqrt{\Gamma} dt \sigma.
\]
In the case of no photon detected, the state is transformed according to
\[
\rho \longrightarrow \frac{\Omega_0(dt)\rho\Omega_0^\dagger(dt)}{\text{Tr}[\Omega_0(dt)\rho\Omega_0^\dagger(dt)]} = \rho - \frac{i}{\hbar}[H,\rho] dt - \frac{\Gamma}{2}[\sigma^\dagger\sigma,\rho]_+ dt + \Gamma\langle\sigma^\dagger\sigma\rangle\rho dt,
\]
keeping terms to first order in $dt$, and in the case of a detector click the state is transformed according to
\[
\rho \longrightarrow \frac{\Omega_1(dt)\rho\Omega_1^\dagger(dt)}{\text{Tr}[\Omega_1(dt)\rho\Omega_1^\dagger(dt)]} = \frac{\sigma\rho\sigma^\dagger}{\langle\sigma^\dagger\sigma\rangle}.
\]
These two transformations correspond exactly to the transformations induced by the SME (26) in the cases $dN = 0$ and $dN = 1$, respectively.

Notice that this POVM tends to drive the atom towards the ground state, as compared to the unconditioned Hamiltonian evolution (and for either possible outcome $\Omega_{0,1}$). By involving the atomic annihilation operator, we see in this case that the POVM generalizes projective measurements by modeling dissipation due to the measurement process. In the case at hand, the physical origin of the dissipation in the case at hand is absorption of radiated photons by the photodetector.

1.6 Example: Gaussian Projectors

POVMs can also generalize projective measurements to model partial or imprecise measurements. Partial measurements leave some uncertainty in the measured observable, whereas projective measurements leave the system in a state where the observable is perfectly defined—that is, an eigenstate of the observable. As a simple example, we can model partial measurements by defining the measurement operators $\Omega_q$ to be Gaussian-weighted sums over projection operators for the discrete set of eigenstates $|q\rangle$ ($q \in \mathbb{Z}$) of the observable $Q$:

$$\Omega_q = \frac{1}{\mathcal{N}} \sum_{j=-\infty}^{\infty} e^{-\kappa(j-q)^2/4} |j\rangle\langle j|.$$  

(30)

Here,

$$\mathcal{N}^2 := \sum_{j=-\infty}^{\infty} e^{-\kappa(j-q)^2/2},$$  

(31)

so that

$$\sum_{q=-\infty}^{\infty} \Omega_q^\dagger \Omega_q = 1,$$  

(32)

as required for the operators to form a POVM. The Gaussian weights lead to having only partial information about $Q$ after the measurement. For example, in a highly uncertain mixed state, where $\langle q|\rho|q\rangle$ is approximately the same for any $q$ and $\langle q|\rho|q'\rangle = 0$ for any $q \neq q'$, the measurement leads to the collapse

$$\rho \rightarrow \frac{\Omega_q \rho \Omega_q^\dagger}{\text{Tr}[\Omega_q \rho \Omega_q]} \approx \frac{1}{\mathcal{N}} \sum_{j} e^{-\kappa(j-q)^2/2} |q\rangle\langle q|.$$  

(33)

The $q$th possible final state is thus peaked about the eigenvalue $q$, and additionally has an uncertainty $\Delta Q = 1/\sqrt{\kappa}$.

In the limit $\kappa \rightarrow \infty$, the measurements here reduce to the usual projective measurements. Thus, for large $\kappa$, the variance in the measurement results (taken over an ensemble of measurements on identically prepared systems) is dominated by the uncertainty in the quantum state, while for small $\kappa$, the measurement variance is dominated by the uncertainty introduced by the measurement operators $\Omega_q$. This distinction divides two categories of measurements, strong measurements where $\kappa$ is large, and weak measurements, where $\kappa$ is small.5

We can also generalize these Gaussian projectors to the continuous-variable case. For example, for a position measurement, the properly normalized measurement operators have the form

$$\Omega(\alpha) = \left(\frac{\kappa}{2\pi}\right)^{1/4} \int_{-\infty}^{\infty} dx e^{-\kappa(x-\alpha)^2/4} |x\rangle\langle x|.$$  

(34)

Again, if this operator is applied to an initially uncertain state (such as a momentum eigenstate), the resulting position variance in the collapsed state is $1/\kappa$ (i.e., the uncertainty is $1/\sqrt{\kappa}$). In what follows, we will consider sequences of weak position measurements of this form, and thus construct continuous quantum

---

measurements of position. For this it is useful to consider the product of two operators,

\[ \Omega(\alpha'; \kappa')\Omega(\alpha; \kappa) = \left( \frac{\kappa \kappa'}{2\pi} \right)^{1/4} \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dx e^{-\kappa'(x'-\alpha')^2/4} |x'\rangle\langle x'| e^{-\kappa(x-\alpha)^2/4} |x\rangle\langle x| \]

which corresponds to a sequence of two Gaussian position measurements, the first of strength \( \kappa \) and the second of strength \( \kappa' \), with measurement outcomes \( \alpha \) and then \( \alpha' \), respectively. This operator product is still Gaussian, but it is not normalized properly in the sense that \( \Omega(\alpha) \) is normalized (note that the norm vanishes if \( \alpha - \alpha' \) becomes large), but we can see from its form that applying this operator to an initially uncertain state gives \( 1/(\kappa + \kappa') \) for the resulting position variance of the state. Hence, a sequence of two Gaussian measurements is effectively equivalent to a single Gaussian measurement, where the strength is the sum of the individual measurement strengths, as long as no other transformation or evolution occurs between the two measurements.

Notice how the information from the second measurement is incorporated with that of the first. After the first measurement, the best estimate for the position of the quantum system is \( \alpha \), with uncertainty \( 1/\sqrt{\kappa} \). After the second measurement (where the result is \( \alpha' \)), the new best position estimate is an average of the old estimate and the new measurement result,

\[ \langle x \rangle = \frac{\alpha \kappa + \alpha' \kappa'}{\kappa + \kappa'}, \]

weighted by the respective uncertainties. The new uncertainty of the estimate is reduced to \( 1/\sqrt{\kappa + \kappa'} \).