

Chapter 1: The Measurement Problem in Quantum Information

Quantum information theory is about information stored, processed, transmitted, and extracted from quantum systems. The word extracted is important. A quantum computer, a quantum communication device, or a quantum sensor is not useful merely because it evolves according to beautiful equations. It becomes useful when we finally ask it a question and obtain classical data: a bit string, a detector click, a measurement record, or an estimate of an unknown parameter.

This chapter begins with the basic tension:

> Quantum systems are described by mathematical states, but our experimental records are classical outcomes.

A measurement is the bridge between the two. The bridge is subtle because quantum measurement is not just passive observation. In the standard mathematical formalism of quantum mechanics, a measurement is represented by a rule that assigns probabilities to possible outcomes and may also change the state of the system after an outcome is obtained (von Neumann, 1955; Nielsen and Chuang, 2010).

In quantum information, we often focus on a practical version of the measurement problem:

> Given a quantum state, what measurements are mathematically allowed, what probabilities do they produce, and how can those measurements be implemented physically?

Naimark's dilation theorem answers a major part of this question. It says, in finite dimensions, that the generalized measurements used throughout quantum information can always be viewed as ordinary projective measurements on a larger Hilbert space. Before we can prove that theorem, we need to understand why generalized measurements are needed at all.

1.1 Classical observation: a helpful but incomplete starting point

Let us begin with a simple classical example.

Suppose a box contains a ball that is either red or blue. You do not know which. In classical probability, we may describe your uncertainty by assigning probabilities:

$$\Pr(\text{red}) = 0.7, \quad \Pr(\text{blue}) = 0.3.$$

The ball itself is assumed to have a definite color. Your probability describes your lack of knowledge. If you open the box and look, you discover the color. In an idealized classical model, the act of looking does not need to change the color of the ball.

This kind of model has three features:

1. The property is already definite.

The ball is red or blue before you look.

2. The measurement reveals the property.

The observation tells you which case holds.

3. The measurement can be nondisturbing in principle.

Looking at the ball need not change red into blue.

This picture is extremely useful in ordinary life. It is also useful in many classical information tasks. But it does not describe quantum measurement in general.

Quantum theory does not usually allow us to say that all measurable quantities have definite pre-existing values independent of the measurement context. The standard formalism represents states and measurements in a way that makes probabilities fundamental, not merely a sign of ignorance (Peres, 1995; Nielsen and Chuang, 2010).

1.2 Quantum states as probability-generating objects

In quantum theory, the state of a finite-dimensional system is represented by a vector or by a matrix, depending on how general we need to be.

For now, imagine a two-level quantum system called a qubit. A qubit is the quantum analogue of a classical bit, but it is not simply “0 or 1 with unknown probability.” A pure qubit state can be written as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle,$$

where α and β are complex numbers satisfying

$$|\alpha|^2 + |\beta|^2 = 1.$$

The symbols $|0\rangle$ and $|1\rangle$ form a standard orthonormal basis. The notation $|\psi\rangle$ is called ket notation or Dirac notation, widely used in quantum mechanics and quantum information (Nielsen and Chuang, 2010).

If we measure this qubit in the standard basis, the Born rule gives

$$\Pr(0) = |\alpha|^2, \quad \Pr(1) = |\beta|^2.$$

The Born rule is the rule that converts quantum states and measurements into probabilities. It is one of the central rules of quantum theory (von Neumann, 1955; Nielsen and Chuang, 2010).

Example: a balanced qubit

Consider the state

$$|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle.$$

If we measure in the standard basis $\{|0\rangle, |1\rangle\}$, then

$$\Pr(0) = \left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2},$$

and

$$\Pr(1) = \left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2}.$$

So the measurement outcome is random.

But this does not mean that $|+\rangle$ is merely a classical mixture that is “really $|0\rangle$ ” half the time and “really $|1\rangle$ ” half the time. The state $|+\rangle$ behaves differently if we measure it in another basis. For example, if we measure in the basis

$$\{|+\rangle, |-\rangle\},$$

where

$$|-\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle,$$

then the outcome corresponding to $|+\rangle$ occurs with probability 1, and the outcome corresponding to $|-\rangle$ occurs with probability 0.

Thus the same quantum state can look random under one measurement and deterministic under another.

This is our first warning: in quantum theory, probabilities are not attached only to hidden classical alternatives. They depend on the relation between the state and the measurement.

1.3 Measurement depends on the question asked

A quantum measurement is not just “looking.” It is more like choosing a question that the physical system will answer probabilistically.

For a qubit, two common measurement bases are:

$$\{|0\rangle, |1\rangle\}$$

and

$$\{|+\rangle, |-\rangle\}.$$

These correspond to different experimental arrangements. Measuring in one basis is generally not the same as measuring in another basis.

Example: measuring $|0\rangle$ in two different bases

Let the state be $|0\rangle$.

If we measure in the standard basis $\{|0\rangle, |1\rangle\}$, then

$$\Pr(0) = 1, \quad \Pr(1) = 0.$$

So the result is certain.

But if we measure in the $\{|+\rangle, |-\rangle\}$ basis, we first write

$$|0\rangle = \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle.$$

Therefore,

$$\Pr(+)=\frac{1}{2}, \quad \Pr(-)=\frac{1}{2}.$$

The same state produces different probability distributions depending on the measurement.

This is a central idea: a quantum state is not a list of answers to every possible measurement. It is a mathematical object that generates probabilities once a measurement is specified.

1.4 The first mathematical model: projective measurement

The earliest standard mathematical model of quantum measurement is the projective measurement, also called a projection-valued measurement in more general settings. This model is closely connected to the spectral theory of self-adjoint operators, which was foundational in von Neumann's mathematical formulation of quantum mechanics (von Neumann, 1955).

We will study projective measurements carefully in Chapter 5. For now, here is the basic finite-dimensional idea.

A projective measurement is described by a collection of orthogonal projections

$$P_1, P_2, \dots, P_m$$

such that

$$P_i P_j = 0 \quad \text{for } i \neq j,$$

and

$$P_1 + P_2 + \dots + P_m = I.$$

Here:

- I is the identity operator.
- Each P_i projects onto a subspace.
- The condition $P_i P_j = 0$ means the corresponding subspaces are orthogonal.
- The condition $\sum_i P_i = I$ means the outcomes cover the whole space.

If the quantum state is represented by a unit vector $|\psi\rangle$, then the probability of outcome i is

$$\Pr(i) = \langle \psi | P_i | \psi \rangle.$$

This is a version of the Born rule.

Example: standard measurement of a qubit

For a qubit measured in the standard basis, define

$$P_0 = |0\rangle\langle 0|, \quad P_1 = |1\rangle\langle 1|.$$

These are projections onto the one-dimensional subspaces spanned by $|0\rangle$ and $|1\rangle$. They satisfy

$$P_0 + P_1 = I,$$

and

$$P_0 P_1 = 0.$$

If

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle,$$

then

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

and

$$\Pr(1) = \langle\psi|P_1|\psi\rangle = |\beta|^2.$$

So projective measurements correctly describe many familiar quantum measurements.

But they are not enough for quantum information.

1.5 Why projective measurements are not the whole story

A projective measurement has a rigid structure. Its outcomes correspond to mutually orthogonal subspaces. This is mathematically elegant, but many realistic and useful measurement procedures do not appear in this form on the original system.

There are at least four reasons.

First, real detectors may be noisy or inefficient. A detector might sometimes fail to click. It might confuse one outcome with another. It might combine several microscopic outcomes into one macroscopic record. Such procedures are naturally modeled by more general measurement rules than sharp projections (Busch, Lahti, and Mittelstaedt, 1996; Nielsen and Chuang, 2010).

Second, a system may be measured indirectly. In the laboratory, one often couples the system of interest to an auxiliary system, allows them to interact, and then measures the auxiliary system. The effective measurement on the original system need not be projective, even if the final measurement on the larger system is projective. This viewpoint is central in quantum information and quantum measurement theory (Nielsen and Chuang, 2010; Watrous, 2018).

Third, some information-processing tasks require measurements with more outcomes than can be represented as rank-one orthogonal projectors on the original space. For example, a two-dimensional qubit Hilbert space cannot contain three nonzero mutually orthogonal one-dimensional subspaces. Yet useful qubit measurements with three outcomes exist, such as the trine measurement used in state discrimination and communication examples (Peres, 1995; Nielsen and Chuang, 2010).

Fourth, optimal information extraction may require generalized measurements. In quantum detection and estimation theory, generalized measurements are essential for describing optimal strategies for distinguishing quantum states and estimating parameters (Helstrom, 1976; Holevo, 1982).

These reasons lead us from projective measurements to POVMs.

1.6 POVMs: the generalized measurement of quantum information

A POVM is a positive operator-valued measure. In finite-dimensional quantum information, a POVM with outcomes $1, 2, \dots, m$ is a collection of operators

$$E_1, E_2, \dots, E_m$$

such that:

$$E_i \geq 0 \quad \text{for every } i,$$

and

$$E_1 + E_2 + \dots + E_m = I.$$

The condition $E_i \geq 0$ means that E_i is a positive semidefinite operator. Informally, this means it never gives a negative expectation value:

$$\langle \psi | E_i | \psi \rangle \geq 0$$

for every vector $|\psi\rangle$.

The condition

$$\sum_i E_i = I$$

is called normalization. It guarantees that the total probability of all outcomes is 1.

Given a state $|\psi\rangle$, the probability of outcome i is

$$\Pr(i) = \langle \psi | E_i | \psi \rangle.$$

More generally, if the state is represented by a density matrix ρ , then

$$\Pr(i) = \text{Tr}(\rho E_i).$$

We will define density matrices and trace carefully in Chapter 4. For now, you can think of $\text{Tr}(\rho E_i)$ as the general version of the Born rule for possibly mixed states.

The POVM formalism is standard in quantum information theory because it captures all measurement outcome probabilities obtainable by general quantum measurement procedures on finite-dimensional systems (Nielsen and Chuang, 2010; Watrous, 2018).

1.7 Every projective measurement is a POVM

A POVM is a generalization of a projective measurement.

If

$$P_1, \dots, P_m$$

are orthogonal projections satisfying

$$\sum_i P_i = I,$$

then they automatically form a POVM.

Why?

Each projection P_i is positive semidefinite. To see this, remember that a projection satisfies

$$P_i^2 = P_i$$

and projects a vector onto a subspace. The quantity

$$\langle \psi | P_i | \psi \rangle$$

is the squared length of the projected part of $|\psi\rangle$, so it is nonnegative. Also, the projections sum to I . Therefore the projective measurement satisfies the POVM axioms.

So we have the inclusion:

$$\text{projective measurements} \subseteq \text{POVMs.}$$

But the inclusion is strict: there are POVMs that are not projective measurements on the original Hilbert space.

1.8 A simple non-projective POVM

Consider a qubit and define two operators

$$E_0 = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|,$$

$$E_1 = \frac{1}{4}|0\rangle\langle 0| + \frac{3}{4}|1\rangle\langle 1|.$$

They are positive because their eigenvalues are nonnegative. Also,

$$E_0 + E_1 = \left(\frac{3}{4} + \frac{1}{4}\right) |0\rangle\langle 0| + \left(\frac{1}{4} + \frac{3}{4}\right) |1\rangle\langle 1| = |0\rangle\langle 0| + |1\rangle\langle 1| = I.$$

Therefore E_0, E_1 is a POVM.

But E_0 and E_1 are not projections. For example,

$$E_0^2 = \left(\frac{3}{4}\right)^2 |0\rangle\langle 0| + \left(\frac{1}{4}\right)^2 |1\rangle\langle 1| = \frac{9}{16}|0\rangle\langle 0| + \frac{1}{16}|1\rangle\langle 1|.$$

This is not equal to

$$E_0 = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|.$$

So $E_0^2 \neq E_0$, meaning E_0 is not a projection.

What kind of measurement is this?

It is an unsharp or noisy version of the standard basis measurement. If the state is $|0\rangle$, then

$$\Pr(0) = \langle 0|E_0|0\rangle = \frac{3}{4},$$

$$\Pr(1) = \langle 0|E_1|0\rangle = \frac{1}{4}.$$

If the state is $|1\rangle$, then

$$\Pr(0) = \langle 1|E_0|1\rangle = \frac{1}{4},$$

$$\Pr(1) = \langle 1|E_1|1\rangle = \frac{3}{4}.$$

So the measurement tends to report the standard-basis value, but it sometimes makes an error. This is exactly the kind of measurement model that appears naturally when detectors are imperfect.

1.9 Measurement is not only about probabilities

A POVM tells us the probabilities of the possible outcomes. But a full physical measurement description should also tell us what happens to the quantum state after the outcome is obtained.

For example, suppose a qubit is measured in the standard basis and the outcome is 0. In the simplest projective measurement model, the post-measurement state becomes $|0\rangle$. If the outcome is 1, it becomes $|1\rangle$. This is often called state update or collapse, though the exact interpretation depends on one's view of quantum mechanics (von Neumann, 1955; Peres, 1995).

In quantum information, we often use a more operational language. A measurement outcome is associated not only with a POVM element E_i , but also with a transformation of the state. These transformations are described by measurement operators, Kraus operators, or more generally quantum instruments. We will study these in Chapter 14.

For now, remember this distinction:

> A POVM describes outcome probabilities. > A quantum instrument describes both outcome probabilities and post-measurement states.

This distinction matters because two different physical measurement procedures can have the same POVM but disturb the system in different ways. Thus, the POVM answers the question "How likely is each outcome?" but not always the question "What is the state after the outcome?"

1.10 The main puzzle leading to Naimark dilation

At this point, POVMs may look like a mathematical invention. We have relaxed projections into positive operators. But why is this relaxation physically legitimate?

Naimark's dilation theorem gives a powerful answer.

Roughly, it says:

> Every POVM can be realized by embedding the system into a larger Hilbert space and performing a projective measurement there.

In quantum information language, this means:

1. Add an auxiliary system, called an ancilla.
2. Let the original system and ancilla be treated as one larger quantum system.
3. Perform an ordinary projective measurement on the larger system.
4. When viewed only from the original system, the measurement statistics are described by the desired POVM.

This is why generalized measurements are not arbitrary. They are exactly the kind of measurement statistics that can arise from ordinary projective measurements on a larger space, at least in the finite-dimensional setting we will study.

This is the central idea of the book.

1.11 A small preview: compression from a larger space

The word dilation means enlarging. In mathematics, a dilation often represents a complicated object on a smaller space as the shadow or compression of a simpler object on a larger space.

Here is the rough shape of Naimark dilation.

Suppose \mathcal{H} is the original Hilbert space. A POVM on \mathcal{H} is a collection

$$E_1, \dots, E_m.$$

Naimark's theorem says that there exists a larger Hilbert space \mathcal{K} , an embedding of \mathcal{H} into \mathcal{K} , and projections

$$P_1, \dots, P_m$$

on \mathcal{H} such that the POVM probabilities on \mathcal{H} are reproduced by the projective measurement probabilities on \mathcal{H}' .

In a common finite-dimensional formulation, the relation looks like

$$E_i = V^* P_i V,$$

where:

- $V: \mathcal{H} \rightarrow \mathcal{H}'$ is an isometry, meaning it preserves inner products and lengths;
- V^\wedge is the adjoint of V ;
- P_i are projections on the larger space \mathcal{H}' ;
- E_i are the POVM elements on the original space \mathcal{H} .

This equation says that each generalized measurement operator E_i is obtained by going up into the larger space using V , applying the projection P_i , and then compressing back down using V^\wedge .

We will not prove this yet. The proof requires linear algebra, Hilbert spaces, projections, positive operators, and square roots of positive operators. These tools are built carefully in the next chapters.

But already, the philosophical message is clear:

> POVMs are generalized measurements on the original system, but they can be understood as projective measurements on an enlarged system.

This is why Naimark dilation is so important in quantum information.

1.12 Why quantum information needs generalized measurements

Let us finish the chapter by naming some major quantum information tasks where POVMs naturally appear.

State discrimination

In quantum state discrimination, someone prepares one of several possible quantum states, and we try to guess which one was prepared. If the possible states are not mutually orthogonal, no measurement can distinguish them perfectly with zero error. This limitation is a basic feature of quantum theory and is central to quantum detection theory (Helstrom, 1976; Holevo, 1982).

POVMs are often needed to describe optimal discrimination strategies. Projective measurements may be too restrictive.

Quantum communication

In quantum communication, a sender encodes classical information into quantum states, and a receiver performs a measurement to decode the message. The receiver's measurement is generally described by a POVM. The theory of accessible information and optimal decoding uses generalized measurements in an essential way (Holevo, 1982; Nielsen and Chuang, 2010).

Quantum cryptography

In quantum cryptography, measurement choices affect what information an honest receiver obtains and what information an eavesdropper might extract. POVMs are used to describe general measurement attacks and realistic detector behavior (Nielsen and Chuang, 2010).

Quantum tomography

In quantum tomography, we try to reconstruct an unknown quantum state from measurement data. Some tomography procedures use POVMs whose outcome probabilities contain enough information to determine the state. Such POVMs are called informationally complete. Informationally complete measurements are standard tools in quantum information and quantum statistical inference (Nielsen and Chuang, 2010; Watrous, 2018).

These applications will return later in the book. For now, the important point is that POVMs are not optional decoration. They are part of the natural working language of quantum information.

1.13 Chapter summary

Quantum measurement is mathematically subtle because quantum states do not simply list pre-existing classical answers. A state gives probabilities only after a measurement has been specified.

Projective measurements are the first and most familiar model. They use orthogonal projections and the Born rule. But quantum information requires a more flexible model because realistic detectors, indirect measurements, noisy measurements, and optimal information-extraction tasks often go beyond projective measurements on the original system.

POVMs provide that flexible model. A finite-outcome POVM is a collection of positive operators summing to the identity. Each POVM element gives one outcome probability through the Born rule.

Naimark dilation explains why POVMs are physically and mathematically natural: every finite-dimensional POVM can be represented as a projective measurement on a larger Hilbert space. The rest of the book builds the tools needed to prove this statement rigorously and use it in quantum information.

References

Busch, P., Lahti, P. J., and Mittelstaedt, P. (1996). *The Quantum Theory of Measurement*. Springer.

Helstrom, C. W. (1976). *Quantum Detection and Estimation Theory*. Academic Press.

Holevo, A. S. (1982). *Probabilistic and Statistical Aspects of Quantum Theory*. North-Holland.

Nielsen, M. A., and Chuang, I. L. (2010). *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press.

Peres, A. (1995). *Quantum Theory: Concepts and Methods*. Kluwer Academic Publishers.

von Neumann, J. (1955). *Mathematical Foundations of Quantum Mechanics*. Princeton University Press.

Watrous, J. (2018). *The Theory of Quantum Information*. Cambridge University Press.

Document information

Chapter 1: The Measurement Problem in Quantum Information

Project	Naimark Dilation from First Principles
Document	Document 1.5
Author	mujirin
Verifier	Not verified
Downloaded	July 06, 2026 01:25 KST
Status	Working
Document link	https://www.theorytrace.com/projects/naimark-dilation-from-first-principles/documents/-chapter-1-the-measurement-problem-in-quantum-information/