Review by Stephen Parrott
of
Quantum Computation and Quantum Information
by
Michael A. Nielson and Isaac L. Chuang

1 General review

My first acquaintance with this book came from a copy which I ordered through interlibrary loan after seeing favorable comments on the internet. The loan period was only two weeks, so I wasn't able to study this 600-page book in detail. But I learned quite a bit just by skimming it. After I saw that it was a book that would repay study, I purchased it.

The first chapter of 58 pages nicely introduces many of the important ideas, leaving the more difficult details to later chapters. For example, I learned about quantum teleportation, which I had never understood from popular accounts.

I read it from cover to cover and was able to follow almost all of it in detail. Since I read it as someone learning this material for the first time, I'll review it from a student's perspective.

Chapter 2 gives a nice (with one possible exception[1]) summary of basic quantum mechanics. It includes an introduction to necessary concepts from abstract linear algebra, including important specific applications (e.g., the Schmidt decomposition) which are not likely to be covered even in advanced linear algebra courses.

The authors wisely limit the treatment to finite-dimensional Hilbert spaces. This is adequate for nearly all the book's applications, and it vastly simplifies the mathematics. The treatment is mathematically accurate and more rigorous than most physics texts, but is generally too coordinate-dependent to full exploit the conceptual simplifications of abstract linear algebra.

For example, I never got used to the authors' notation for the (orthogonal) projector from a Hilbert space H to a subspace E with orthonormal basis $\{e_i\}_{i=1}^m$. Most mathematicians would use some notation like P_E, but most physics texts (including this one) use the more cumbersome (expressed in the Dirac "ket" notation which is used throughout the book):

$$\sum_{i=1}^{m} |e_i\rangle\langle e_i|. \tag{1}$$

And then if v is a vector expressed in some other orthonormal basis $\{h_k\}_{k=1}^n$ for H as $v = \sum_{k=1}^n |h_k\rangle\langle h_k|v\rangle$, one arrives at hard-to-parse expressions for the projection of v on E such as:

$$\sum_{i,k} |e_i\rangle\langle e_i|h_k\rangle\langle h_k|v\rangle \quad . \tag{2}$$

[1]The exception is their untraditional axiom concerning "measurement operators", which I'll discuss in detail later in this review.

Surely $P_E v$ is simpler and easier to understand than (2)! It's true that (1) and (2) give more information than P_E and $P_E v$ because they effectively contain a recipe for the construction of P_E in terms of given orthonormal bases, but the reader must already know this. There is no point to repeat it every time a projector arises! It is simpler and more enlightening to think of P_E as the unique operator which fixes all vectors in E and annihilates all vectors orthogonal to E instead of as a basis-dependent formula like (1). That frees up mental space for more important things.

The third chapter gives an introduction to computer science concepts. This gives a conceptual framework within which to present the ideas of quantum computation. More material is included here than is necessary to understand the rest of the book. Readers may find it efficient to skim this chapter initially and return for more detail when necessary.

The next three chapters present the essentials of quantum circuits, the quantum Fourier transform, and quantum search algorithms. Here there is perhaps room for a little improvement. I thought that important details were sometimes omitted from the exposition, and I occasionally had to go to the original literature to understand the ideas.

Also, there is a bad misuse of the "Big-O" notation, startling in a book so generally carefully written. The text defines the notation in more or less the usual way:

> "Suppose $f(n)$ and $g(n)$ are functions on the non-negative integers. We say . . . '$f(n)$ is $O(g(n))$' if there are constants c and n_0 such that for all values of n greater than n_0, $f(n) \leq cg(n)$."

(I imagine that the authors probably intended that f, g, and c be positive, but this would be only a minor slip.) Thus to write $f(n) = O(1)$ means simply that f is bounded above. In their analysis of various algorithms (e.g., the "quantum order-finding" algorithm on p. 252) one of the conclusions is often that the algorithm "succeeds with probability $O(1)$". Since all probabilities are bounded above by 1, this doesn't say a lot! Presumably, the authors mean that the probability $p(n)$ of success with input of size n is bounded away from zero for large n. Sophisticated readers will understand what must have been meant, but this sort of gross error might demoralize an undergraduate. Fortunately, though I did spot an occasional mathematical error, they were few, and none were as elementary and glaring as this.

The mathematics of quantum computation is easy compared to the problems of physically realizing it. Chapter 7 gives an extensive discussion of these problems and various proposals for overcoming them. This concludes the "quantum computation section of the book, which is a little more than half of the 600-odd pages.

The rest deals with quantum information theory. This is presented in less detail than the quantum computation chapters, and demands more from the reader. A summary of classical information theory is included, with sketches of proofs of important results. I found this very helpful in refreshing my memory of Khinchin's book on information theory, which I read decades ago.

Some of the more complicated proofs of quantum information results are only sketched. I didn't get as much from the quantum information section of the book as from the quantum computation section. I think it gives a useful overview of the field, but if I wanted to learn quantum information in detail, I would look for a book dedicated to this topic, perhaps reading Nielsen/Chuang first as an introduction.

The book concludes with a 12-page introduction to quantum cryptography. I couldn't follow this section in detail. Perhaps it could be followed with enough work, but I wasn't motivated. I imagine that a proper treatment of cryptography would require many more than 12 pages. Again, if I wanted to learn this material, I would seek an expository text dedicated to it.

In summary, this is an exceptionally fine text which can be read on many levels. The first chapter, 58 pages, gives an overview of quantum computation, much of which should be comprehensible to anyone familiar with the basic ideas of quantum mechanics. The rest of the book may possibly be readable with great effort by well-prepared undergraduates, but I think a graduate-level background in quantum mechanics and linear algebra would be more realistic prerequisites, and also more efficient. These prerequites will have to be mastered anyway for anyone who wants to work in a field dependent on quantum theory. Those who lack the prerequisites may still be able to get a feel for the problems of quantum computation and information from the book, even if the details seem too difficult.

Although this is a serious book suitable for obtaining a professional knowledge of its subjects, it is unusually carefully written in an expository style. There are many exercises interspersed with the exposition, but no solutions are provided. Most of them should be solvable on sight by anyone following the presentation, so they provide a useful check on one's understanding of the material. I could do most of them in my head, or at least sense how the solutions might go. (I am a professional mathematician; students may find the exercises less easy.)

For those which puzzled me, I sometimes wished for an appendix of solutions. When one can't do an exercise, one wants to know if it is because some important concept has been missed, or if one simply hasn't thought of some solution trick. Glancing at a solution can often determine this.

Each chapter ends with "History and further reading" sections, often extensive. I found these very helpful.

I hope that this general review may be helpful to those thinking of investing time or money in the book. The rest of the review consists of comments on the authors' approach, and mathematical comments. Most of these will probably be of interest only to those who have already read or are reading parts of the book.

One possible exception is the section on "measurement operators". It advises most readers to ignore the text's introduction of measurement operators. This introduction seems questionable, and the measurement operators are hardly used in the rest of the text anyway. For nearly all applications, the traditional formulation in which the measurement operators are orthogonal projectors (a special case of the text's approach) seems sufficient.

2 Remarks on "measurement operators"

2.1 The text's axiomatization of quantum mechanics

The text develops quantum mechanics from four postulates. I don't think that they are sufficiently precise to be meaningful to someone who has never studied quantum mechanics, nor do they seem complete (e.g., observables are never mentioned), but they provide a useful summary for someone who is already familiar with the subject.

The first postulate specifies that a *state* of an "isolated system" (which the text does not define) is a unit vector in a Hilbert space. This is what is usually called a *pure state*. Later, the text introduces mixed states, without a formal postulate.

The second postulate states that evolution through a given time of the state ψ of a "closed" quantum system is given by a unitary operator U: $\psi \mapsto U\psi$.

I found the third postulate extremely puzzling. In fact, I puzzled over it through about 500 pages of the book. The text introduces it as follows:

[BEGIN QUOTE]

> "We postulated that closed quantum systems evolve according to unitary evolution. The evolution of systems which don't interact with the rest of the world is all very well, but there must also be times when the experimentalist and their experimental equipment — an external physical system in other words — observes the system to find out what is going on inside the system, an interaction which makes the system no longer closed, and thus not necessarily subject to unitary evolution. To explain what happens when this is done, we introduce Postulate 3, which provides a means for describing the effects of measurements on quantum systems.
>
> **Postulate 3:** Quantum measurements are described by a collection $\{M_m\}$ of *measurement operators*. These are operators acting on the state space of the system being measured. The index m refers to measurement outcomes that may occur in the measurement. If the state of the quantum system is ψ immediately before the measurement then the probability that result m occurs is given by
>
> $$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$$
>
> and the state of the system after the measurement is
>
> $$\frac{M_m|\psi\rangle}{\sqrt{\langle|M_m^\dagger M_m|\psi\rangle}} \quad . \tag{3}$$

The measurement operators satisfy the *completeness equation*

$$\sum_m M_m^\dagger M_m = I \quad .$$

[END QUOTE]

If the measurement operators M_m are orthogonal projectors, then they may be considered the spectral measure of some observable, and the experimenter is simply measuring that observable. But what if the M_m are *not* orthogonal projectors? How does the measurement differ from measuring an observable? If one is not measuring an observable, what is one measuring?

Only much later did I find a partial answer to this question, and I am still somewhat puzzled by it. I will describe this partial answer below. But before doing this, let me reassure potential readers that in nearly all of the text's applications, the measurement operators do turn out to be orthogonal projectors. Thus the reader who shares my puzzlement will lose little by assuming that the measurement operators are orthogonal projectors, in which case the text's treatment would be compatible with more usual treatments of elementary quantum mechanics.

At the end of the chapter the authors comment

> "Many texts on quantum mechanics deal only with projective measurements. For applications to quantum computing and quantum information it is more convenient — and, we believe, easier for novices — to start with the general description of measurements of which projective measurements can be regarded as a special case. Of course, ultimately, as we have shown, the two approaches are equivalent."

I certainly did not find it easier to start with the text's "measurement operators". I had a lot of trouble relating the text's approach to the treatments of most quantum mechanics texts.

3 Remarks on measurement operators

This section presents the interpretation of measurement operators at which I eventually arrived, in the hope that it might help other readers. I puzzled over this through about 400 pages of the book. When I finally partially resolved the problem, I realized that much of the puzzlement was because the book had presented some concepts in earlier chapters which were not properly explained until later chapters, without warning that the proper explanation would come later. Also, the reader has to figure out how to interpret the later explanations

Suppose we have a quantum system which if treated as an isolated system would be studied as described in beginning quantum mechanics texts. For example, evolution through time t of a pure state f would be given by

$$f \mapsto U(t)f, \quad \text{where } U(t) \text{ is a unitary operator.}$$

And time evolution of a mixed state ρ (represented by a positive Hermitian operator of trace 1) would be given by

$$\rho \mapsto U(t)\rho U(t)^\dagger \quad .$$

That elementary quantum mechanics approach is similar to studying Newtonian mechanics assuming that there are no frictional forces. This approach works well for celestial mechanics, but not so well if we are studying the motion of a creaky old wagon that needs lubrication.

Similarly, the idealizations of elementary quantum mechanics may not well describe the behavior of “real-world” quantum systems. A commonly used attempt to introduce “decoherence” (a sort of quantum analog of Newtonian friction) into the mathematics is the following.

Assume that the system S which we are studying is coupled to an ”environment” E. The environment is a separate quantum system which is regarded as not fully known, just as frictional forces are usually too complicated to be described in full detail.

The Hilbert space for the composite system (“ideal system” S plus “environment” E) is $H_S \otimes H_E$, where H_S and H_E are respectively the Hilbert spaces for S and E. A mixed state of the composite system is represented by a positive operator of trace 1 on $H_S \otimes H_E$. (By “mixed state”, I mean a state which is not necessarily a pure state, but can be.)

The text assumes that the starting state of the composite system is a product state $\sigma \otimes \epsilon$, where σ is a state of S and ϵ a pure state of the environment E (represented in the present “mixed state” notation by a one-dimensional projector on H_E. That ϵ may be assumed a pure state of the environment is justified by a fictitious mathematical construction known as “purification” discussed in Chapter 2.

The following mathematics does not require that ϵ be pure, but it does require that the initial state of the composite sysem be a product state. The text (p. 358) gives the impression that the product state assumption can be justified. Perhaps so, but I have not been able find a convincing justification.

It considers the evolution of the starting state into an “unnormalized state”

$$T(\sigma \otimes \epsilon)T^\dagger, \tag{4}$$

where T is a given operator on the $H_S \otimes H_E$. I introduce the nonstandard term “unnormalized state” because (4) is not necessarily a genuine state: it is a positive operator but need not have trace 1. In the text, (p. 363 ff.) T is of the form $T = UP$ with U unitary and P a projector, but the mathematics works for any T. The special cases $T = U$ and $T = P$ correspond, respectively, to unitary time evolution and projective measurement.

To get the final state of the original system S corresponding to (4), we take the partial trace tr_E of (4) over the environment:

$$\text{unnormalized final state of } S = \mathrm{tr}_E\ (T(\sigma \otimes \epsilon)T^\dagger) \quad . \tag{5}$$

The result is not obvious, but calculation reveals that

$$\begin{aligned}\text{unnormalized final state of } S &= \operatorname{tr}_E\,(T(\sigma\otimes\epsilon)T^\dagger)\\ &= \sum_{i=1}^{n} A_i\sigma A_i^\dagger \qquad (6)\end{aligned}$$

where $A_1, A_2, \ldots, A_n$ are operators on H_S. The operation $\sigma \mapsto (6)$ is trace-preserving if (and only if) $\sum_{i=1}^{n} A_i^\dagger A_i = I$, where I is the identity operator on S. If T is unitary, then (4) is trace-preserving, and since partial tracing is trace-preserving, so are (5) and (6).

The text calls the operation which sends a state σ of S into (6) a "quantum operation" (for which there is also an axiomatic definition). Thus unitary time evolution of S is replaced by the quantum operation (6). Also, usual projective measurement operations on S are replaced by (6).

If we agree that one can never separate a system S from its "environment", and that one must use the formalism just described, then we see that the text is really replacing elementary quantum mechanics by a more general formulation in which time evolution is not necessarily unitary, and in which both time evolution and measurement are given by the same kind of operation, namely (6).

The operation (6) is reminiscent of the text's formulation of measurement in terms of "measurement operators". Suppose we are doing a projective measurement on the composite system corresponding to orthogonal projectors $P_1, P_2, \ldots P_q$ with $\sum_{k=1}^{q} P_k = I$. According to the usual rules of projective measurement, the kth measurement result (corresponding to P_k) is realized with probability $\operatorname{tr}\, P_k(\sigma\otimes\epsilon)P_k$, and the resulting state of the composite system is

$$\frac{P_k(\sigma\otimes\epsilon)P_k}{\operatorname{tr}\, P_k(\sigma\otimes\epsilon)P_k} \quad .$$

This corresponds to the kth measurement result being realized in the system S with the same probability

$$\operatorname{tr}\, P_k(\sigma\otimes\epsilon)P_k = \operatorname{tr}\, \sum_{i=1}^{n_k} A_{ki}\sigma A_{ki} \quad ,$$

where the A_{ki} are the operators in (6) corresponding to $T = P_k$, and the resulting state of S:

$$\begin{aligned}&\text{state of S conditional on measurement result } k =\\ &= \frac{\operatorname{tr}_E\, P_k(\sigma\otimes\epsilon)P_k}{\operatorname{tr}\, P_k(\sigma\otimes\epsilon)P_k}\\ &= \frac{\sum_{i=1}^{n_k} A_{ki}\sigma A_{ki}^\dagger}{\operatorname{tr}\, \sum_{i=1}^{n_k} A_{ki}\sigma A_{ki}^\dagger} \quad . \qquad (7)\end{aligned}$$

This is the state of S assuming that we know that the k'th measurement result occured. If we only know that a measurement has been made, but do not know the result, the messy probabilities cancel, and the resulting state of S is more simply expressed as:

$$\begin{aligned}
&\text{state of S when measurement result is unknown} \\
&\qquad = \sum_{k=1}^{q} \mathrm{tr}_E \; P_k(\sigma \otimes \epsilon) P_k \\
&\qquad = \sum_{k=1}^{q} \sum_{i=1}^{n_k} A_{ki}(\sigma \otimes \epsilon) A_{ki}^\dagger \quad . \qquad (8)
\end{aligned}$$

Reindexing (replace ki by a single index) converts this to form (6).

This seemingly artificial situation arises, for example, if the measurement is made by someone else, who informs us that this has been done, but does not tell us the result. We know that the measurement will probably have changed the state of the system, but we don't know how. From the measurer's point of view, the post-measurement state is given by (7), but from our point of view it is (8).

Explained in this way, to extend elementary quantum mechanics in this way seems not unreasonable. But in that case, it seems that one ought to abandon not only projective measurement, but also unitary time evolution.

Also several consistency problems arise. For example, even if the starting state σ for (6) is a product state, the result need not be a product state, which was the premise for the proposal of (6) to replace time evolution and projective measurement..

Another possible inconsistency arises when we try to relate (7) to the text's measurement Postulate 3. This was given above, and the part of concern here is (3):

$$\text{post-measurement state conditional on measurement result } m = \frac{M_m|\psi\rangle}{\sqrt{\langle| M_m^\dagger M_m |\psi\rangle}} \quad .$$

That postulate refers to pure states, but when reformulated in terms of mixed states described by positive, unit-trace operators σ, (7) becomes:

$$\text{post-measurement state conditional on measurement result } m = \frac{M_m \sigma M_m^\dagger}{\mathrm{tr}\; M_m \sigma M_m^\dagger} . \qquad (9)$$

Here σ is the pre-measurement state. When σ is the pure state associated with state vector ψ (i.e., σ is the projector on ψ), then (9) is the projector on $M_m\psi$.[2] The reformulation (9) of Postulate 3 is (7) with just one A_{ki}, e.g., $n_k = 1$ in (7). Thus the measurement postulate is formally (and actually) less general than the result (7) obtained by adjoining an "environment" to the original system S

[2]To see this without calculation, note that the range of (9) is one-dimensional (except in the degenerate case in which the numerator and denominator are identically zero), and (9) is a positive Hermitian operator, so it must be the one-dimensional projector on the vector $M_m\psi$ in its range. (The degenerate case occurs with probability zero, and so can be ignored.)

under study, performing a projective measurement in the enhanced system, and partial tracing over the environment to obtain the state of the original system. So shouldn't one use (7) in place of (9) for the measurement postulate?[3]

This question is not addressed in the text, but it is mentioned in a paper of the text's author Nielsen and C. Caves (Phys. Rev. A **55** (1997), 2547-2556). That paper discusses general measurements given by (7) and defines an *ideal measurement* to be the special case (9). In explanation for this terminology, it states:

> "It can be shown that ideal measurements correspond in a certain sense to doing a perfect readout of the state of the apparatus to which the system is coupled. This is the reason we call such a measurement ideal."

No further explanation is given, and I am not sure I have correctly guessed the "certain sense" which might make the authors' statement true.

I have wondered if it might have something to do with the fact that for a pure starting state, the measurement outcomes will be pure states if and only if it can be assumed that for each measurement outcome k, only one A_{ki} occurs in (7). If this should be the correct guess, it would be very helpful if the text made it explicit. In any case, I think further explanation of the text's measurement postulate is needed.

3.1 The text's explanation of the relation of Postulate 3 to more usual approaches

What I think of as the "traditional" formulation of quantum mechanics assumes a one-to-one correspondence between projectors on the Hilbert space of pure states and binary observables. A binary observable is one which takes on only two values, so it may be considered as a question with "yes" or "no" answers. It is assumed that the probability of a "yes" answer for the question associated with a projector P when the system is in pure state ψ is $\langle\psi|P|\psi\rangle$.

A general observable which can take on a finite set of distinct real values, say $v_1, v_2, ..., v_n$ may then be identified with the n questions: "Was the observable measured as having value v_k, $k = 1, \ldots n$. [4] The traditional approach assumes that these n binary observables correspond to *orthogonal* projectors, so via the spectral theorem, every observable corresponds to a Hermitian operator, and conversely.

Before examining carefully the text's attempt to relate the measurement operators of Postulate 3 to the traditional approach, we need to add the text's

[3] At first glance, it may look as if one could convert (7) into (9) by reindexint the A_{ki}, but this is an illusion because the index k refers to a particular measurement result, as does the index m in (9), while the double index ki is an artifice of the mathematics. The problem is that a mapping $\sigma \mapsto \sum_j B_j \sigma B_j^\dagger$ with more than one j cannot necessarily be expressed as a mapping $\sigma \mapsto B\sigma B^\dagger$ with only one term.

[4] Recall that the text only considers finite-dimensional state spaces, and correspondingly an observables is assumed to take on only a finite set of of values.

last Postulate 4:

Postulate 4: "The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if we have systems numbered 1 through n and system number i is prepared in the state $|\psi_i\rangle$, then the joint state of the total sysem is $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots \otimes |\psi_n\rangle$."

After a brief discussion of this postulate, the text begins its explanation of the relation of the measurement operators of postulate 3 with the traditional approach, in which the measurement operators are orthogonal projectors (the text calls the latter a "projective measurement"):

> "In Section 2.2.5 we claimed that projective measurements together with unitary dynamics are sufficient to implement a general measurement. The proof of this statement makes use of composite quantum systems, and is a nice illustration of Postulate 4 in action."

I think that most people would read this as claiming that the measurement Postulate 3 for measurement operators which happen to be (orthogonal) projectors together with the assumption that time evolution is given by a unitary operator as described in Postulate 2 (together with Postulate 4) imply Postulate 3 for arbitary measurement operators. But the proof following in the text does not prove this.

Instead, it shows that:

> Given any measurement operators M_m with $\sum_m M_m = I$, it is possible to embed the original system S with Hilbert space H_S in a composite system with Hilbert space $H_S \otimes H_E$, such that for some collection $\{P_m\}$ of orthogonal projectors on $H_S \otimes H_E$ and some unitary operator U on $H_S \otimes H_E$,
>
> $$M_m = \mathrm{tr}_E \; U^\dagger P_m U \quad .$$

(That is not a quote from the book; it is what I think was actually proved in the text, as opposed to what was claimed to be proved.)

If $H_S \otimes H_E$ happens to be physically realizable, and if U implements an evolution of that system, one might reasonably conclude that the measurement operators M_m are physically realizable. But the question of the physical realizability of U is addressed only much later in the book, and then only incompletely.

Let's examine the short proof given in the text to see exactly what was proved. It continues:

> "Suppose we have a quantum system with state space Q, and we want to perform a measurement described by measurement operators M_m on the system Q. To do this, we introduce an *ancilla* system with state space M, having an orthonormal basis $|m\rangle$ in one-to-one

correspondence with the possible outcomes of the measurement we wish to implement. This ancilla system can be regarded as merely a mathematical device appearing in the construction, or it can be interpreted physically as an extra quantum system introduced into the problem, which we assume has a state space with the required properties.

Letting $|0\rangle$ be any fixed state of M, define an operator U on products $|\psi\rangle|0\rangle$ of states $|\psi\rangle$ from Q with the state $|0\rangle$ by

$$U|\psi\rangle|0\rangle \equiv \sum_m M_m|\psi\rangle|m\rangle.$$

Following is a routine calculation showing that

> "... U can be extended to a unitary operator on the space $Q \otimes M$, which we also denote by U."

The proof continues:

> "Next, suppose we perform a projective measurement on the two systems described by projectors $P_m \equiv I_Q \otimes |m\rangle\langle m|$. Outcome m occurs with probability
>
> $$\begin{aligned} p(m) &= \langle\psi|\langle 0|U^\dagger P_m U|\psi\rangle|0\rangle \\ &= \sum_{m'm''} \langle\psi|M^\dagger_{m'}|(I_Q \otimes |m\rangle\langle m|)\langle m|)M_{m''}|\psi\rangle|m''\rangle \\ &= \langle\psi|M^\dagger_m M_m|\psi\rangle, \end{aligned}$$
>
> as given in Postulate 3. The joint state of the system QM after measurement conditional on result m occurring is given by
>
> $$\frac{P_m U|\psi\rangle|0\rangle}{\sqrt{\langle|U^\dagger P_m U|\psi\rangle}} = \frac{M_m|\psi\rangle|m\rangle}{\sqrt{\langle\psi|M^\dagger M_m|\psi\rangle}}.$$
>
> It follows that the state of system M after the measurement is $|m\rangle$, and the state of system Q is
>
> $$\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M^\dagger_m M_m|\psi\rangle}},$$
>
> just as prescribed by Postulate 3. Thus unitary dynamics, projective measurements, and the ability to introduce ancillary systems , together allow any measurement of the form described in Postulate 3 to be realized."

Remarks on the text's proof

1. This proof, in Chapter 2, implicitly assumes that the unitary operator U can be physically implemented. This assumption is more or less justified in Chapter 4, over a hundred pages after the above proof. Chapter 4 sketches a highly technical proof that every unitary operator on a finite-dimensional quantum system can be *approximately* implemented by quantum gates which can be physically realized.

   The Chapter 2 proof does not warn the reader that its (partial) justification will be deferred to Chapter 4. The result in my case was that I puzzled over the Chapter 2 proof for a hundred pages, before I finally realized how it fits into the general scheme of the book.

2. The text claims that "projective measurements together with unitary dynamics are sufficient to implement a general measurement". I read this as saying that the existence of projective measurements togetner with the fact that time evolution is implemented by unitary operators imply that general measurement operators M_m can be implemented. But the text does not prove this. Instead, it proves that projective measurements together with the physical implementability of an arbitrary unitary operator imply that general measurement operators can be implemented.

   And only the *approximate* implementability of arbitrary unitary operators is proved later in the book, and that only for finite-dimensional systems. I think that the text's claim that standard quantum mechanics justifies Postulate 3 is open to question.

4 A minor error?

For such a large book, this one is remarkably free of errors, both typographical and substantive. I did notice a few here and there, but very few which would seriously bother an experienced reader. This section mentions one possible error because it occurs in the statement of a fundamental result and might bother mathematicians.

Section 8.2.4 gives an axiomatization of "quantum operations". Its main result is Theorem 8.1, which states that any operation which satisfies the axioms is of the form (6).

The text defines a "quantum operation" $\mathcal{E}$ as a mapping from the set of density operators (i.e., positive operators of unit trace) on one Hilbert space to the corresponding set on another Hilbert space. This mapping is required to be convex-linear (along with other properties). However, the proof of Theorem 8.1 (cf. their equation (8.58)) requires applying $\mathcal{E}$ to operators which are not Hermitian (and so cannot be density operators).

I think that to fix this, the definition of "quantum operation" should probably be changed to allow them to be applied to arbitrary (not just density) operators. It could be fixed more easily if there were a canonical way to extend an operation on density operators to an operation on all operators, but this seems unlikely.

An operation on Hermitian operators can be trivially extended to all operators, but it seems problematic to extend a convex-linear map on density operators to Hermitian operators. The problem is that although every Hermitian operator H is a unique difference $H = H_+ - H_-$ of positive operators H_+ and H_-, the convex-linearity of the "natural" extension $\mathcal{E}(H) := \mathcal{E}(H_+) - \mathcal{E}(H_-)$ is not obvious, and may not be true. This is because if, say, $H = (1/2)H_1 + (1/2)H_2$, then it is usually not true that $H_+ = (1/2)(H_1)_+ + (1/2)(H_2)_+$, which would presumably be needed to verify that the extended $\mathcal{E}$ is convex-linear.