Pergamon

Stud. Hist. Phil. Mod. Phys., Vol. 31, No. 1, pp. 117–125, 2000

Printed in Great Britain
1355-2198/00 $ - see front matter

Essay Review

Operational Quantum Physics

*Gordon N. Fleming**

Paul Busch, Marian Grabowski and Pekka Lahti, *Operational Quantum Physics* (Berlin: Springer Verlag, 1995), xi + 230 pp., ISBN: 3-540-59358-6, $49.95.

1. Introduction

For those at least, like our authors, who embrace an interpretation of quantum theory as realistic, objective and applicable to individual systems, the quantum measurement *problem* has not been solved and remains a thorn in the side. But quantum measurement *theory*, i.e. the theory of measurements performable on quantum systems, has evolved considerably from the classic version formulated long ago by Von Neumann (1955), and this book presents a version of the generalised theory and a variety of examples of its application. It can be regarded as a sequel to an earlier book, *The Quantum Theory of Measurement* by the first and last of the present authors, along with Peter Mittelstaedt (Busch *et al*., 1991). But while the earlier book focused on the authors' philosophical interpretation and the basic physics underlying quantum measurement theory, the present book is more concerned with the resulting mathematical and structural features and with applications. Still, ideology is very present here and the reader should be aware that the book is part of the tradition growing out of the earlier work of Ludwig (1983) and Krause (1983), Ali (1985) and Prugovecki (1986), and Davies (1976) and Holevo (1982). Ludwig and Krause, and the Marburg school generally, sought to understand and interpret quantum theory ultimately in terms of the manipulation of macroscopic systems alone. Ali,

(*Received 2 March 1999; revised 7 April 1999*)
* 325 Davey Laboratory, Pennsylvania State University, University Park, PA 16802, U.S.A. (*e-mail*: gnf1@earthlink.net).

PII: S1355-2198(99)00025-8

Prugovecki and others (Schroeck, 1996) have sought to construct, in quantum theory, the concept of joint probability distributions for sets of non-commuting observables (and thereby, phase space) by generalising the traditional concept of (sharp) observable to include one of unsharp, or fuzzy observable. The classic open system study of Davies and the probability study of Holevo were less driven by ideological concerns.

A large amount of rather technical material is covered in this slim volume. To introduce the reader to *some* of that material, and prepare him for this reviewer's evaluation, I will first present a survey of the core ideas of the generalised quantum theory of measurement given in the book.

2. Operational Quantum Measurement Theory: Effects and Observables

For mathematical simplicity my discussion will be confined to the case of measurements that have a finite number of possible outcomes, O_n ($n = 1, 2, \ldots, N$). The state of the quantum system being measured (just prior to the measurement) will be represented by a positive trace class operator, $\hat{\rho}$. To each possible outcome O_n is associated a positive operator $\hat{P}_n$, called an *effect*, such that for every possible $\hat{\rho}$,

$$0 \leqslant [\mathrm{tr}(\hat{\rho}\hat{P}_n)/\mathrm{tr}(\hat{\rho})] \leqslant 1. \tag{2.1}$$

The set of all the $\hat{P}_n$ satisfy

$$\sum_n \hat{P}_n = \hat{I}, \tag{2.2}$$

where $\hat{I}$ is the identity operator. All partial sums of the $\hat{P}_n$ are also called effects and the set of all these effects is called a *positive operator valued measure* or POV measure, for short. If an effect, which is, after all, a self-adjoint operator, has an eigenvalue greater than $\frac{1}{2}$ and an eigenvalue less than $\frac{1}{2}$, then it is called a *property*. If at least one of the effects generated by the $\hat{P}_n$ is a property, then the POV measure is called an *observable* and the probability of obtaining the outcome O_n from a system in the state $\hat{\rho}$ is given by

$$P(O_n/\hat{\rho}) = \mathrm{tr}(\hat{\rho}\hat{P}_n)/\mathrm{tr}(\hat{\rho}). \tag{2.3}$$

If the $\hat{P}_n$ are all non-trivial projection operators, then they and all their effects are properties and we have a standard observable of the classic von Neumann theory. In the present context such observables are called *sharp observables*. The others, the new ones, are called *unsharp observables* and for them the $\hat{P}_n$ not only need not be projection operators, but they need not even commute with one another! Furthermore, for the effects of an unsharp observable which are not properties, it follows from (2.3) that the associated outcome, or collection of outcomes, is either unlikely (probability $\leqslant \frac{1}{2}$) regardless of the quantum state $\hat{\rho}$, or it is likely (probability $\geqslant \frac{1}{2}$) regardless of the quantum state $\hat{\rho}$! Only effects that are properties can yield probabilities that vary from less than $\frac{1}{2}$ for some states to more than $\frac{1}{2}$ for others.

By now the reader may be puzzled and have some questions. Why have these unsharp observables been introduced at all? Why are not all the effects required to be properties? Why is the POV measure itself called an observable, rather than a single self-adjoint operator built from the $\hat{P}_n$, as in the case of sharp observables?

The most physical reason for introducing the unsharp observables is that the more sophisticated and realistic the account one gives of the actual processes that occur in real measurements, the clearer it becomes that the probability distributions that result for the object system being observed correspond to unsharp observables. The recognition that all measurements beyond mere counting are inaccurate and approximate to some degree is the simplest indicator of the need for unsharp observables. In this regard an important subclass of unsharp observables is that comprised of effects generated by *coarse graining* over a sharp observable. In general, if

$$\hat{Q}_m = \sum_n p_{mn} \hat{P}_n, \tag{2.4}$$

where

$$\sum_n p_{mn} = 1, \tag{2.5}$$

for all n, then the $\hat{Q}_m$ are effects that generate an observable which is a *coarse grained* version of the observable generated by the $\hat{P}_n$. We will see further into this physically motivated need for unsharp observables when we consider the measurement process in the next section.

The association of sharp observables with single self-adjoint operators, with the $\hat{P}_n$ as the projection operators that comprise their spectral resolution, is possible because from that single self-adjoint operator one can uniquely recover the $\hat{P}_n$. For unsharp observables that is not possible. Not being mutually orthogonal projection operators, the *semi-spectral* resolution of a self-adjoint operator that such $\hat{P}_n$ would provide is very non-unique. So to define an unsharp observable, the $\hat{P}_n$ must be independently specified.

The remaining question is the most delicate. Why should not all the effects be properties? The reason is the desire to find unsharp observables that can, in some sense, be regarded as representing approximate joint measurements of incompatible (non-commuting) sharp observables. It is much harder (not always impossible, however) to do that if one requires all the effects of a POV measure to be properties! And if each of the $\hat{P}_n$ themselves are properties then all the other effects will be properties. But there is a downside to this liberality and we will consider it in Section 4.

How does one identify joint observables in the present scheme? Let the $\hat{P}_n$ be effects generating the observable P and let $\hat{Q}_m$ denote effects generating the observable Q. If there exist effects $\hat{R}_{mn}$ generating an observable R, and such that

$$\hat{P}_n = \sum_m \hat{R}_{mn}, \tag{2.6a}$$

and

$$\hat{Q}_m = \sum_n \hat{R}_{mn}, \tag{2.6b}$$

then R is a *joint observable* for P and Q. If P and Q are sharp observables, with projection operators for effects, then a corresponding R exists iff P and Q are compatible, i.e. the $\hat{P}_n$ and $\hat{Q}_m$ commute, in which case,

$$\hat{R}_{mn} = \hat{P}_n \, \hat{Q}_m = \hat{Q}_m \, \hat{P}_n. \tag{2.7}$$

For unsharp observables, however, (2.7) need not hold. But it is much harder to find such R's if all their effects must be properties. Furthermore, by relaxing the properties requirement one can find observables for which the probability distributions, $P(O_n/\hat{\rho})$, are unique signatures of the states $\hat{\rho}$, i.e. no two inequivalent states give the same distribution. Such observables are called *informationally complete*.

3. Operational Quantum Measurement Theory: Operations and Measurements

The other principal ingredient of the theory presented in this book is the account of the changes the quantum states undergo as a consequence of the measurements. This material being mathematically more complex than the preceding account of observables, my presentation will be accordingly more sketchy.

We begin with *operations*, which are positive linear maps Φ, from trace class operators to trace class operators, such that for any quantum state $\hat{\rho}$, we have

$$0 \leqslant \mathrm{tr}[\Phi(\hat{\rho})] \leqslant \mathrm{tr}[\hat{\rho}]. \tag{3.1}$$

An operation uniquely defines an effect $\hat{B}$, by the association

$$\mathrm{tr}[\Phi(\hat{\rho})] = \mathrm{tr}[\hat{\rho}\hat{B}], \; \forall \hat{\rho}, \tag{3.2}$$

which can always be satisfied. Any effect can be so determined by infinitely many distinct operations. As an important example of the operation–effect connection, we have the *Luders operation* Φ^B_{L}, for the effect $\hat{B}$:

$$\Phi^B_{\mathrm{L}}(\hat{\rho}) := \hat{B}^{1/2}\hat{\rho}\hat{B}^{1/2}. \tag{3.3}$$

But

$$\Phi^B_{U\mathrm{L}}(\hat{\rho}) := \hat{U}^{\dagger}\hat{B}^{1/2}\hat{\rho}\hat{B}^{1/2}\hat{U}, \tag{3.4}$$

where $\hat{U}$ is any unitary operator, defines the same effect.

Now consider a set O_n of possible outcomes of a measurement and let N be the set of values of the index, n. If we then assign an operation Φ_Δ to each subset Δ of N, such that

$$\Phi_\Delta(\hat{\rho}) = \sum_{n\in\Delta} \Phi_n(\hat{\rho}), \tag{3.5}$$

and

$$\mathrm{tr}[\Phi_N(\hat{\rho})] = \mathrm{tr}[\hat{\rho}], \tag{3.6}$$

then the set of all the Φ_Δ constitute a *state transformer*, and the associated set of effects, if it includes a property, defines an observable. In the measurement context, of course, the defined observable is the measured observable and the operations for the individual outcomes yield the corresponding state reductions, though the latter terminology is largely avoided in the book.

The measurement context itself is modelled with the usual product state space $H := H_S \otimes H_A$, where H_S is the object system state space and H_A the apparatus state space, the usual initial product state $\hat{\rho}^0 := \hat{\rho}_S^0 \otimes \hat{\rho}_A^0$ and a *measurement coupling* V, which yields the resulting entangled state operator and is usually a unitary evolution,

$$V(\hat{\rho}^0) = \hat{U}^\dagger \hat{\rho}^0 \hat{U} := \hat{\rho}^{\mathrm{f}}. \tag{3.7}$$

An apparatus observable Z is identified, with effects $\hat{I} \otimes \hat{Z}_n$, associated with the individual outcomes and such that the *probability reproducibility condition*

$$\begin{aligned} P_S(O_n/\hat{\rho}_S^0) &= \mathrm{tr}_S[\hat{\rho}_S^0 \hat{P}_n]/\mathrm{tr}_S[\hat{\rho}_S^0] = \mathrm{tr}[\hat{\rho}^{\mathrm{f}}(\hat{I} \otimes \hat{Z}_n)]/\mathrm{tr}[\hat{\rho}^{\mathrm{f}}] \\ &= \mathrm{tr}_A[\rho_A^{\mathrm{f}} \hat{Z}_n]/\mathrm{tr}_A[\rho_A^{\mathrm{f}}] = P_A(O_n/\hat{\rho}_A^{\mathrm{f}}) \end{aligned} \tag{3.8}$$

is satisfied, where $\hat{\rho}_A^{\mathrm{f}} := \mathrm{tr}_S[\hat{\rho}^{\mathrm{f}}]$. It is in this scheme that one can see in detail the physically most important motivation for generalising the concept of observable from that of sharp observable. For even if the apparatus observable Z is sharp, realistic measurement couplings will not allow a sharp system observable $\{\hat{P}_n\}$ to satisfy the probability reproducibility condition. And so, in effect (no pun intended), one is said to be measuring an unsharp system observable. Furthermore, the generalisation does not need further extension in the sense that even if Z were unsharp, a corresponding unsharp system observable $\{\hat{P}_n\}$ always exists.

Lastly, for any single outcome O_n, let the state transforming operation on the system-apparatus composite be Φ_n. Then a system state operation, Φ_{S_n}, is induced by tracing out the apparatus:

$$\Phi_{S_n}(\hat{\rho}_S^{\mathrm{f}}) := \mathrm{tr}_A[\Phi_n(\hat{\rho}^{\mathrm{f}})], \tag{3.9}$$

and the measurement is called *repeatable* iff $P_S(O_n/\hat{\rho}_S^0) \neq 0$ implies

$$P_S(O_n/\Phi_{S_n}(\hat{\rho}_S^{\mathrm{f}})) = 1. \tag{3.10}$$

Remembering from (3.2) that

$$\mathrm{tr}_S[\Phi_{S_n}(\hat{\rho}_S^{\mathrm{f}})] = \mathrm{tr}_S[\hat{\rho}_S^{\mathrm{f}}\hat{P}_n], \tag{3.11}$$

repeatability is equivalent to

$$\mathrm{tr}_S[\Phi_{S_n}(\Phi_{S_n}(\hat{\rho}_S^{\mathrm{f}}))] = \mathrm{tr}_S[\Phi_{S_n}(\hat{\rho}_S^{\mathrm{f}})]. \tag{3.12}$$

It then turns out that only observables with countable outcomes can be repeatably measured, and of them only those whose individual outcome effects have 1 as an eigenvalue. This excludes continuous observables, which we have not discussed, and also informationally complete observables which were mentioned above.

4. Critique

There are many topics addressed in the book which I have not commented on. The relevance of symmetry considerations and covariance to the concept of observable. The many ways of quantitatively assessing unsharpness and the uncertainty or imprecision associated with the joint measurement of non-commuting, unsharp observables. The study of special observables including screen observables, time-of-arrival observables, angle observables, phase observables, informationally complete observables and phase space observables. The examination of the application of the general formalism to real measurements. All of these and more are in the book and worthy of study.

The treatment throughout, unfortunately, is rather abstract in character and many readers, particularly physicists of a less mathematical bent than the authors, will wish for a more physical discussion. To that end I recommend as a supplement the book *Quantum Measurement* by Braginsky and Khalili (1992). It is not nearly as mathematically general or rigorous as the book under review but, since Braginsky and Khalili have been involved in the design of real, extremely sensitive experiments, their physically broader account and their use of much of the same theory Busch *et al.* are presenting is a cornucopia of relevant material.

Some of the material presented in the book is controversial. The objection has been raised by Uffink (1994) that the concept of observable presented in the literature contributing to this book is so general as to lead to absurdities. He showed that the definition of observable allows one to identify many observables the probability distributions for which could be obtained by measurement sequences in which one simply flips coins to decide which of many ordinary observables one would measure in each instance. Among the examples he presented were informationally complete observables. Unfortunately, there are a number of instances in the book to which this criticism still applies. Oddly, Uffink's argument is not responded to nor is his paper even cited.

The book fares better with respect to a second of Uffink's objections. He had pointed to instances in the earlier literature in which workers erroneously inferred the existence of joint unsharp observables corresponding to non-commuting *sharp* observables. The book uses the language of joint observables more carefully than the earlier works Uffink considers and I found no instances in which the error in question occurred. Nevertheless, the terminology of the book is, in my opinion, conducive towards the error and no warnings are stated. For that reason I will paraphrase Uffink's argument here.

Suppose A is a joint observable for the observables B and C, i.e. in the notation of Eqns (2.6a) and (2.6b),

$$\hat{B}_n = \sum_m \hat{A}_{mn} \quad \text{and} \quad \hat{C}_m = \sum_n \hat{A}_{mn}. \tag{4.1}$$

Suppose B is a coarse-grained version of the observable D, and C is a coarse-grained version of the observable E, i.e. in the notation of Eqns (2.4) and (2.5)

$$\hat{B}_n = \sum_r p_{nr}\hat{D}_r \quad \text{and} \quad \hat{C}_m = \sum_{\mathrm{s}} p'_{m\mathrm{s}}\hat{E}_{\mathrm{s}}. \tag{4.2}$$

Then the offending inference was to regard A as a joint observable for D and E. Indeed this inference was endorsed in the earlier literature to identify unsharp joint observables for position and momentum or for two orthogonal components of angular momentum. The intuitive motivation for the inference is that since measuring B provides partial information about D, and measuring C provides partial information about E, and measuring A provides joint information about B and C, surely the last measurement must provide joint partial information about D and E (see Uffink, 1994, pp. 208–210). But the inference requires the existence, among the effects comprising the observable A, of effects $\hat{A}'_{rs}$ that satisfy

$$\hat{D}_r = \sum_s \hat{A}'_{rs} \quad \text{and} \quad \hat{E}_s = \sum_r \hat{A}'_{rs}, \tag{4.3}$$

and the existence of such effects is just not implied by the assumption of Eqns (4.1) and (4.2). As a counterexample suppose D is the sharp joint observable $J_{1\mathrm{x}}\&J_{2\mathrm{x}}$, where $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ are independent angular momenta. In other words the projection valued spectral resolution of D consists of the products of the spectral resolution operators for $J_{1\mathrm{x}}$ and $J_{2\mathrm{x}}$. Similarly suppose E is the sharp joint observable $J_{1\mathrm{y}}\&J_{2\mathrm{y}}$. Then $B := J_{1\mathrm{x}}$ is a coarse graining of D, and $C := J_{2\mathrm{y}}$ is a coarse graining of E, although here the coarse grainings still yield sharp observables. Finally,

$$A := B\&C = J_{1\mathrm{x}}\&J_{2\mathrm{y}}$$

is a joint observable for B and C, with A still a sharp observable. But then it is absurd to think of A as a joint observable for D and E for they are all sharp and no two of them commute.

As stated above, the language of coarse graining and joint observables is used very carefully in this book. It is, for example, never position and momentum, per se, for which one finds unsharp joint observables, but rather, unsharp coarse grainings of position and momentum only. However, the joint observable thus found are often subject to Uffink's first criticism.

So what assessment is to be made here? Is Operational Quantum Measurement Theory so broad as to be trivial? I think not. The mathematics and formal deductions are quite rigorous and the structural relationships between operations, effects and probability distributions in quantum theory are both interesting and important. Much of the analysis presented and reported on here has seen use in quantum optics, macroscopic quantum interference experiments, gravitational wave detection experiments, quantum non-demolition experiments, etc. (cf. Braginsky and Khalili, 1992).

The problem, it seems to me, is the terminology and the interpretations. What is really being studied here are measurable probability distributions that can be sustained by quantum systems. Unfortunately, the explicit statement to that effect does not appear until the middle of the book (p.118), where the authors write: 'From the operational point of view observables are representations of the *totality of probability distributions for measurement outcomes*' (my italics). That is an important subject. But to associate all such possible probability distributions with the concept of observable is to disconnect that concept from its traditional quantum and classical association with *possessable properties*. If, for example, the technical concept of observable was restricted to POV measures, *all* of whose effects were properties in the technical sense (see definition between (2.2) and (2.3)), then all of Uffink's absurd coin tossing examples would be eliminated. The discarded POV measures would, in a sense, still define *experiments*, but not observables. And there would be informationally complete experiments, but not informationally complete observables.

For myself the most interesting passages of this book examined models of joint measurement couplings, Eqn (3.7), of non-commuting sharp observables to commuting pointer observables of independent apparatus systems (pp. 150–155, 174–177, 195–197, 206–212.). This kind of study is what one must undertake to find out what present theory dictates will happen when one attempts to jointly measure two non-commuting sharp observables. But it seems prejudicial to decide, at the outset, to call the result, whatever it may be, a measurement of some unsharp observable. In some cases that terminology may make sense. In others it seems very forced.

There are lesser problems. I found the book hard to read. The text is dense on the page and much of the mathematics is dense in the text, not visually separated from the text. This frequently made it difficult to find a passage I wanted to return to if I did not remember exactly where it had appeared. The 'realistic' examples discussed (Ch.VII) were disappointingly confined to spin and photon measurements when a wider range of examples is available and interesting (see Braginsky *et al.*, 1992). And finally, the discussion of observables defining ordinary phase space (Ch.VI) suffered from the absence of any account of how to proceed when an external magnetic field is turned on and the canonical momentum becomes gauge-dependent and, therefore, not observable. This would be doubly interesting since the kinetic momentum, which is gauge-invariant, acquires non-commuting components. See Ballentine (1990, Ch. 11) for a nice discussion, from a more traditional perspective, of this situation.

To sum up, I regard the material covered here as important, and likely to become more so as the interpretation scheme is clarified. But I strongly recommend the reader to have a copy of Uffink and a copy of Braginsky and Khalili near at hand.

References

Ali, S. T. (1985) 'Stochastic Localization, Quantum Mechanics on Phase Space and Quantum Space-Time', *La Rivista del Nuovo Cimento* **8**(11), 1–128.

Ballentine, L. E. (1990) *Quantum Mechanics* (London: Prentice Hall).

Braginsky, V. B. and Khalili, F. Y. (1992) *Quantum Measurement*, edited by K. S. Thorne (Cambridge: Cambridge University Press).

Busch, P., Lahti, P. and Mittelstaedt, P. (1991) *The Quantum Theory of Measurement* (Berlin: Springer Verlag).

Krause, K. (1983) *States, Effects and Operations* (Berlin: Springer Verlag).

Ludwig, G. (1983). *Foundations of Quantum Mechanics*, *Vol. 1* (Berlin: Springer Verlag).

Prugovecki, E. (1986) *Stochastic Quantum Mechanics and Quantum Space Time,* 2nd edn (Dordrecht: Reidel).

Schroeck, F. E. Jr., (1996) *Quantum Mechanics on Phase Space* (Dordrecht: Kluwer).

Uffink, J. (1994) 'The Joint Measurement Problem', *International Journal of Theoretical Physics* **33**(1), 199–212.

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

Well, classical physics isn't going to cut it. You'll need quantum physics to explain things like atomic energy levels, photons, and the wavelength of electrons.Â About this unit. Do you want to describe the microscopic world of atoms and nuclei? Well, classical physics isn't going to cut it. You'll need quantum physics to explain things like atomic energy levels, photons, and the wavelength of electrons. Our mission is to provide a free, world-class education to anyone, anywhere. Khan Academy is a 501(c)(3) nonprofit organization.