TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

The uncertainty principle

Citation for published version (APA):
Martens, H. (1991). *The uncertainty principle*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Applied Physics and Science Education]. Technische Universiteit Eindhoven. https://doi.org/10.6100/IR348359

DOI:
10.6100/IR348359

Document status and date:
Published: 01/01/1991

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 04. Oct. 2023

The Uncertainty Principle

Hans Martens

The Uncertainty Principle

The Uncertainty Principle

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de Rector Magnificus, prof. ir. M. Tels, voor een commissie aangewezen door het College van Dekanen in het openbaar te verdedigen op dinsdag 19 maart 1991 te 16:00 uur

door

HANS MARTENS

geboren te Roermond.

Dit proefschrift is goedgekeurd door de promotoren, prof. F. Sluijter en prof. J. de Graaf.

Druk: FEBO, Enschede

The night world can't be represented
in the language of the day

James Joyce

Voor mijn ouders

Contents

Introduction

With the discovery of quantum mechanics (QM) in the early twenties[1], the cornerstone for most of modern physics was laid. The new theory, however, remained at first highly abstract. Instead of characterizing objects by means of a velocity, or rather a momentum p, and a position q, it employs a complex function $\psi(q)$, to be chosen such that

$$(1) \qquad \int_{-\infty}^{\infty} |\psi(q)|^2 \, dq = 1 \quad .$$

These functions can even be *superposed*, as in $c_1\psi_1(q) + c_2\psi_2(q)$, leading to the possibility of interference for material objects. Insight into the meaning of these *wave functions* was obtained through Born's "statistical" interpretation[2]. It appeared that $|\psi(q)|^2$ represented the probability density of finding the particle at position q. Condition (1) thus reduces to a probability normalization. The function $\psi(q)$ is uniquely related to its Fourier transform $\phi(p)$, given by

$$(2) \qquad \phi(p) = (2\pi\hbar)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \psi(q) \exp(i\frac{pq}{\hbar}) \, dq \quad .$$

Here $\hbar$ is Planck's constant divided by 2π, which we shall in the following take to be equal to 1. The whole theory can be alternatively framed in terms of the momentum representation $\phi(p)$ rather than $\psi(q)$. Indeed, as became clear with the the advent of *transformation theory*[3], many more such equivalent representations of QM exist. Therefore it is profitable to denote the state in a representation-free way by the abstract vector $|\psi\rangle$ (Dirac notation). The vector space consisting of these vectors is called a Hilbert space[4]. The inner product of two vectors $|\psi\rangle$ and $|\varphi\rangle$ is denoted by $\langle\psi|\varphi\rangle$. The position representation $\psi(q)$ and momentum representation $\phi(p)$ of the

---

[1]For a detailed history see: J. Mehra & H. Rechenberg (1982): *The Historical Development of Quantum Theory* (6 vols., Springer, Berlin); M. Jammer (1989): *The Conceptual Development of Quantum Mechanics* (2nd. ed., Tomash/American Institute of Physics).

[2]M. Born (1926a): *Zs. f. Phys.* **37**, p. 863; (1926b): *ibid.* **38**, p. 803

[3]P. Dirac (1927): *Proc. R. Soc. A* **113**, p. 621; P. Jordan (1927): *Zs. f. Phys.* **40**, p. 809; D. Hilbert, J. von Neumann & L. Nordheim (1927): *Math. Annal.* **98**, p. 1.

[4]J. von Neumann (1932): *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin). This work was done mainly in the period 1927-1929 [J. von Neumann (1961): *Collected Works*, vol. 1 (ed. by A. Taub; Pergamon, NY)]

vector $|\psi\rangle$ can be expressed as $\langle q|\psi\rangle$ and $\langle p|\psi\rangle$, respectively. The "vectors" $|q\rangle$ and $|p\rangle$ are eigenvectors of the self-adjoint operators $\boldsymbol{Q}$ and $\boldsymbol{P}$ with eigenvalues q and p, respectively:

$$\boldsymbol{Q}\,|q\rangle = q|q\rangle \quad \text{and} \quad \boldsymbol{P}\,|p\rangle = p|p\rangle \tag{3}$$

In terms of the wave function $\psi(q)$ these operators correspond to

$$\boldsymbol{Q}[\psi(q)] = q\,\psi(q) \quad \text{and} \quad \boldsymbol{P}[\psi(q)] = \tfrac{1}{i}\frac{\partial\psi(q)}{\partial q} \quad . \tag{4}$$

For each of the possible representations of the wave-function, a probabilistic interpretation can be set up. Since, e.g., position and momentum representation are connected through (2), it is clear that the position and momentum probability distributions cannot be chosen independently. This is brought out most clearly by the *uncertainty principle* (UP), discovered in 1927 by Heisenberg[5]. He showed that (for position $\boldsymbol{Q}$ and momentum $\boldsymbol{P}$)

$$\langle\Delta^2\boldsymbol{Q}\rangle\,\langle\Delta^2\boldsymbol{P}\rangle \geq \tfrac{1}{4} \; . \tag{5}$$

Here the expectation value $\langle f(\boldsymbol{Q})\rangle$ for some function f is defined as

$$\langle f(\boldsymbol{Q})\rangle = \int_{-\infty}^{\infty} f(q)\,|\psi(q)|^2\,dq \;, \tag{6}$$

and in particular the variance $\langle\Delta^2\boldsymbol{Q}\rangle$ is given by

$$\langle\Delta^2\boldsymbol{Q}\rangle := \langle(\boldsymbol{Q}-\langle\boldsymbol{Q}\rangle)^2\rangle = \int_{-\infty}^{\infty}(q-\langle\boldsymbol{Q}\rangle)^2\,|\psi(q)|^2\,dq \quad . \tag{7}$$

The momentum quantities are analogously defined. The variance, as is well-known from probability theory, characterizes the spread of a probability distribution. Thus, ineq. (5) says that the position and momentum probability distributions cannot both be arbitrarily narrow. It implies, roughly speaking, that one cannot at the same time attribute velocity and position to an object. Position and momentum and, more generally, pairs of quantities that satisfy relations like (5), are termed *incompatible*.

---

[5]W. Heisenberg (1927): *Zs. f. Phys.* **43**, p. 172

Note that no characteristics of the position measurement device are inherent in $\langle \Delta^2 Q \rangle$. In particular its accuracy is not involved. In fact the probability distribution $|\psi(q)|^2$ only results from an ideally accurate position measurement (in a certain sense it *defines* such a measurement).

The UP has been called "the most important principle of twentieth century physics". Laplacean determinism, the assertion that the future of the whole universe is determined by the specification of the positions and velocities of its constituents at a certain time, seemed at an end. The UP engendered a flurry of philosophizing about possible new world pictures[6]. It was even claimed that the existence of free will, which appeared to contradict the Laplacean world picture, was saved by QM. From a physicist's point of view, however, its importance is limited. In quantum mechanical calculations it is incorporated automatically. Explicit consideration of the UP is superfluous. Similarly, Lorentz contraction need not be explicitly introduced into relativistic calculations, as these incorporate the effect automatically. But, whereas the UP is perhaps little used in actual practical calculations, its importance from a pedagogical point of view remains substantial. From the point of view of the conceptually familiar classical physics, the UP highlights one of the ways in which QM is fundamentally "different". But precisely what it means, philosophically and otherwise, is not as simple as suggested at the beginning of this paragraph. The QM evolution equation (Schrödinger's equation) replacing Newton's laws, on which Laplace's views were based, is just as deterministic as the latter. To what extent this implies a deterministic world, depends on the meaning of $|\psi\rangle$ itself. In other words, the philosophical significance of QM can be judged only when the state vectors $|\psi\rangle$ are interpreted satisfactorily, and not through (5) alone. This interpretation problem, closely connected to the notorious "measurement problem" is very complex and controversial, however. Since this work is not directly concerned with it, we shall (apart from an occasional remark) not go into it any further.

---

[6]M. Jammer (1974): *The Philosophy of Quantum Mechanics* (Wiley, NY); p. 75ff; E. McMullin (1954): *The Principle of Uncertainty* (PhD thesis, University of Louvain, Belgium, unpublished).

But even on a pragmatic level, to which we shall limit ourselves, the meaning of the UP is less clear than it may seem. We saw how Born's probabilistic interpretation gives (5) the meaning of a limit to the width of probability distributions. But Heisenberg, judging by the illustrations in his 1927 paper, and Bohr, who subsequently studied the UP in depth, intended the UP to have a much wider significance. Thus, it has been suggested that there are as many as three or four uncertainty principles[7]. Most notably, the UP was interpreted as a limit to the accuracy with which incompatible observables can be measured jointly. A consequence of this latter version of the UP was assumed to be the fact that a position meter must "disturb" incompatible observables, e.g. momentum, to an extent at least reciprocally related to its accuracy. However, as Born's interpretation presupposes the measurement to be ideally accurate, so does (5). It does not at all address questions involving accuracy or disturbance.

In the early days of QM the expansion of the domain of applicability of the new formalism was most important. It is therefore understandable that conceptual issues without direct practical relevance, were not thoroughly investigated (except by Bohr and Einstein). Moreover, measurement devices were for many years so inaccurate that a detailed consideration of quantum induced bounds to accuracy were academic. In recent years, especially the demand for accuracy by gravitational wave detectors and the rapid development of the field of quantum optics, have brought the (alleged ?) quantum bounds into sight. Indeed a number of investigations into these bounds have appeared[8]. It turned out that the conventional QM formalism, though suitable for all calculations, showed deficiencies as regards the description and characterization of measurements. An extended formalism was developed[9].

---

[7]Y. Yamamoto *et al.* (1990): *Progress in Opt.* (ed. by E. Wolf, North Holland, Amsterdam) **28**, p. 87 (see esp. p. 101); McMullin, *op. cit.*

[8]See e.g. Yamamoto *et al.*, *op. cit.*.

[9]E. Davies (1976): *Quantum Theory of Open Systems* (Academic, NY); G. Ludwig (1983): *Foundations of Quantum Mechanics*, 2 vols. (Springer, Berlin); A. Holevo (1982): *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam).

We will start this thesis with a concise overview of the inception of the standard formalism in general, and of the UP in particular. We will also study the relevance of (5) more closely, and see that it is indeed limited when compared to the intended meaning of the UP. Certain implicit assumptions in the setting up of the formalism are traced, assumptions that (may have) led to its later inadequacy for the description of measurements. The subject of ch. II is Bohr's *complementarity*. Bohr developed this philosophy in the years 1927-1939, and we shall study it with special regard for Bohr's views on the UP.

Bohr, as we noted earlier, gave the UP a significance far beyond (5). This discrepancy between the content of the UP and its formalistic status needs clarification. Therefore we proceed with a mathematical investigation in ch. III, using the aforementioned extended formalism. We show that, giving 'inaccuracy' a mathematically well-defined content, an inaccuracy bound can be derived. Next, in ch. IV the inaccuracy notion and the quantum inaccuracy bound are applied in certain experiments, e.g. from quantum optics. In particular certain well-known results, such as Heisenberg's γ-microscope, are treated as consequences of the inaccuracy principle. Inequalities of the type (5) can be shown to have highly analogous consequences, but for devices other than meters: for preparators, i.e. object sources. Thus a dualistic UP is proposed, consisting of a cluster of relations like (5) on the one hand, and of relations like the inaccuracy inequality of ch. III on the other. These two sub-principles appear sufficient to justify the UP in its full Bohr/Heisenberg content. The results are summarized and evaluated in ch. V.

In this thesis equations are numbered in each chapter separately. When an equation in another chapter is referred to, the chapter number is stated explicitly. For example, (III.20) means equation (20) of chapter III. As in the introduction, short remarks and references can be found in the footnotes, indicated by Arabic numerals. Roman numerals indicate longer comments, which can be found at the end of each chapter. The appendices contain a more detailed justification of ch. II and ch. III, but are not directly involved in the line of argumentation of these chapters.

The motto was taken from p. 590 of *James Joyce* by R. Ellman (rev. ed., Oxford University Press, 1982).

The contents of chapters III and IV are contained in:

H. Martens (1989): *Phys. Lett. A.* 137, p. 155
H. Martens & W. de Muynck (1990a): *Found. Phys.* 20, p. 257
H. Martens & W. de Muynck (1990b): *Found. Phys.* 20, p. 355
H. Martens & W. de Muynck (1990c): submitted to *Found. Phys.*
H. Martens & W. de Muynck (1990d): submitted to *Phys. Lett. A*

Further elaborations (on neutron interferometry and Kerr QND measurement, respectively) can be found in:

W. de Muynck & H. Martens (1990): *Phys. Rev. A.* 42, p. 5079
H. Martens & W. de Muynck (1990e): paper presented at the International Workshop on *Quantum Aspects of Optical Communications*, Paris (France), proceedings to be published by Springer, Berlin.

This work was supported by the Foundation for Philosophical Research (SWON), which is subsidized by the Netherlands Organization for Scientific Research (NWO)

CHAPTER I

Historical Prelude

Modern quantum mechanics (QM) is usually said to have started with Heisenberg's 1925 paper "Über die quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen" [1]. As already indicated in this title, early views on the interpretation of the new formalism were strongly tainted by the classical background from which it emerged. For this reason it seems worthwhile to explicitly state some of the interpretational presuppositions of classical (statistical) mechanics (CM), as they are commonly (but often implicitly) taken to be. In CM the "state" of the system at time t is given by a point $\omega(t)$ in phase space Ω. An n-particle system is, for example, described by 3n position coordinates and 3n momentum coordinates. This leads to a phase space $\Omega = \mathbb{R}^{6n}$. In general a given history of the system, or *preparation* procedure, will not uniquely determine the system's position in phase space. In such a case it is appropriate to use a probability distribution $P(d\omega,t)$ to describe the system: $P(\Delta\omega,t)$ indicates the probability that the system can at time t be found in the region $\Delta\omega$ of the phase space. In the following we shall reserve the term *state* for this distribution, and speak of a *C-state* when we refer to a point in phase space.

Note that the set of states is *convex* (fig. 1): whenever $P_1(d\omega,t)$ and $P_2(d\omega,t)$ are states, so is the *mixture*

$$(1) \qquad P(d\omega,t) = \lambda P_1(d\omega,t) + (1-\lambda)P_2(d\omega,t) \qquad (0 \leq \lambda \leq 1).$$

The mixed state P can be realized in a situation where we do not always use the same preparation device: we use the preparator that makes P_1 with probability λ, and the preparator that produces P_2 with probability $1-\lambda$. Elements P of a convex set that cannot be decomposed into two other elements P_1 and P_2 as in (1) are called *extreme* (in this context the extreme elements are also called *pure states*). It is not difficult to verify that the δ–distributions $P_{\omega_0}(d\omega,t) = \delta_{\omega_0}(d\omega)$, which are in 1–1 correspondence to C–states, are the pure states. Moreover, every non-extreme state can be written as a mixture of pure states in a *unique* way: in CM the set of states forms a *simplex*[2]. Therefore we may conceive the CM system as being at any time in some definite C–state, which may not be completely known. The non-extreme states are only

---

[1] W. Heisenberg (1925): "On the quantumtheoretical reinterpretation of kinematic and mechanical relations", *Zs. f. Phys.* **33**, p. 879

[2] A. Holevo (1982): *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam), ch. I

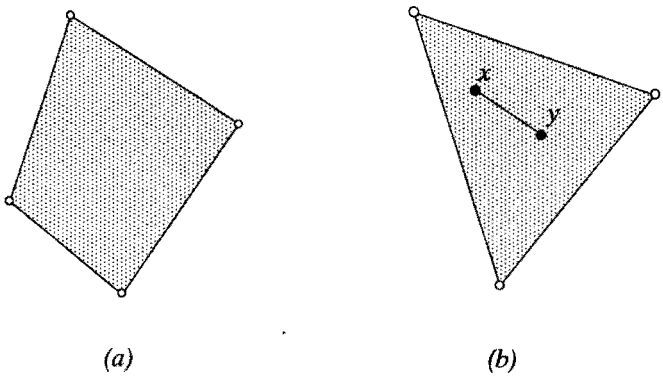


fig. 1 *Two convex sets. Extreme elements are indicated by open circles. The set (b) is a simplex, (a) is not. The line between the points x and y indicates the set of convex combinations of x and y.*

introduced to represent such a lack of knowledge and do not have any *ontic* significance. This is called the *ignorance interpretation* of mixtures.

Classical quantities can be seen as properties independently possessed by the object system: for every quantity $\mathcal{F}$ there is a function $f(\omega)$ determining the value of the quantity, given the C–state of the object system. The quantities supply information on the C–state. In fact, the C–state is no more than the set of values which the quantities assume at a given time. Thus the ultimate quantity *is* the phase point, and *vice versa*: there is no real conceptual difference between 'state' and 'quantities'. Accordingly, measurement of a quantity is ideally intended to see which value the quantity *has*. The nature of the classical measurement ideal follows from the ontological assumption inherent in classical theories that they are about independently possessed object *properties*.

Of course this by no means implies that actual measurements achieve the ideal. On the contrary, real measurements will always be riddled with imperfections. An analysis of the measurement procedure will nevertheless show the precise influence of disturbances, allowing us to interpret our actually performed measurement in terms of the intended one. Thus, while it is not true that in CM all measurements are just

"seeing what value a certain quantity has", it is true that all can be seen as derivatives of such measurements[i]. Furthermore, the character of the classical measurement ideal prompts the view that the property under investigation is well-defined (though perhaps not constant) throughout the measurement process. Consequently a measurement can be used not only to gain information about the object's state just *before* the measurement (the *determinative* aspect of measurement), but also to make predictions about values of the measured quantity in the object's state *after* measurement. Ideally, the post-measurement value of that quantity is equal to the measurement outcome. We shall call the aspect of measurement which deals with the state after measurement, the *preparative* aspect. Accordingly, in CM the preparative and determinative aspects of measurement are quite naturally connected conceptually.

1 THE ADVENT OF QUANTUM MECHANICS

Within the classical conceptual framework QM came into being around 1925. In those days atomic theory was phrased in terms of the "old quantum theory". This eventually evolved into discussing the atom in terms of some symbolic *classical* model ("*Ersatz*"), to which the quantum rules were applied[3]. In this way Bohr's correspondence principle, which started out as the rule that quantum results should become classical results for large quantum numbers, was sharpened into a more quantitative tool. When discussing the problem of radiation and atoms, the *Ersatz* consisted of a set of mechanical oscillators associated with each atom[4]. These *virtual* oscillators had the frequencies of the spectral lines of the atom as eigenfrequencies. The modeling of emission and absorption processes with the aid of these oscillators "solved" the problem of the difference between the mechanical and the electromagnetic frequencies of an atom. An application of the model was the Bohr-Kramers-

---

[3]J. Mehra & H. Rechenberg (1982): *The Historical development of Quantum Theory*, 6 vols. (Springer, NY); see vol. II, p. 199 ff.

[4]This model was due to Slater (Mehra & Rechenberg, *op. cit.*, vol. I, part 2, ch. V and vol. II, p. 125).

Slater theory of radiation[5]. In the theory an atom in a stationary state generates, via the virtual oscillators, a virtual field consisting of components with those frequencies that can be emitted in a transition to lower levels. The *probability* that a given atom actually decays to a certain lower state, depends on the intensity of the virtual field component with the proper frequency at the site of the atom. In this theory there are no photons. Since the occurrence of the transition does not causally depend on whether any other atom makes a transition, energy and momentum are only conserved in the mean. The Bohr-Kramers-Slater theory was soon disproved by experiment[6], but it nevertheless was an important point on the way towards true QM. It in particular formed the starting point for Kramers' theory of dispersion, in which Heisenberg collaborated. Heisenberg was still not satisfied with the status of the correspondence principle, and wanted to further sharpen it. The quantities a(n,m), which denoted the virtual amplitude associated with the transition from level n to level m in the old theory, became *matrices*. These matrices, obeying a non-commutative multiplication rule[ii], were used by Heisenberg to "reinterpret mechanical relations quantum mechanically" (viz. the title of his paper). The theory was, however, still a radiation theory: the "position matrix" q(n,m) corresponded to line intensities in dipole transitions, rather than to electron position. Heisenberg claimed that he had rid the theory of unobservables. For Heisenberg, e.g., electron position was not observable. Instead he referred primarily to line intensities as observable, as opposed to the unobservable mechanical models (such as that of the virtual oscillators) of the old QM. Born and Jordan, with Heisenberg[7], developed Heisenberg's *ansatz* into the consistent formalism of matrix mechanics. A statistical interpretation was added by Born[8]. This interpretation was, however, still to a large extent in line with Heisenberg's original theory as regards its observability notion: Born's interpretation referred to transition probabilities in collision and was intended *only* for momentum and energy, *not* for, e.g., position.

---

[5]N. Bohr, H. Kramers & J. Slater (1924): *Phil. Mag.* **47**, p. 785; see also Mehra & Rechenberg, *op. cit.*, vol. I, part 2, § V.1

[6]By the Compton-Simon and Bothe-Geiger experiments (Mehra & Rechenberg, *op. cit.*, vol. I, part 2, § V.1).

[7]M. Born, W. Heisenberg & P. Jordan (1926): *Zs. f. Phys.* **35**, p. 557

[8]M. Born (1926a): *Zs. f. Phys.* **37**, p. 863; (1926b): *ibid.* **38**, p. 803
For energy this interpretation was already inherent in the Born-Heisenberg-Jordan [Born, Heisenberg & Jordan, *op. cit.*] paper (Mehra & Rechenberg, *op. cit.*, vol. II, p. 138).

Thus we see how QM explicitly originated in classical mechanics, was in fact seen as a reformulation of CM. Dirac puts it in his formulation of the new theory[9] as follows: "In a recent paper[10] Heisenberg puts forward a new theory which suggests that it is not the equations of classical mechanics which are in any way at fault, but that the mathematical operations by which physical results are deduced from them requires modification. All the information supplied by the classical theory can thus be made use of in the new theory". More or less as a by-product[11], new QM used less unobservables than before [i.e. in the old QM]: mechanical models were dispensed with.

2 HILBERT SPACE

In the early papers observability was used in a different sense than the modern one, the latter being characterized by the name *observables* for self-adjoint operators. This latter concept of observability[iii] emerged when the *transformation theory*[12] established the equivalence of all representations of the quantum state vector, and made Born's statistical interpretation available for other quantities than momentum and energy[13]. The new formulation[14] can be roughly summarized in a number of postulates[iv]:

(2a) At a fixed time t the state of a physical system is represented by a positive operator with unit trace $\rho(t)$ on a complex Hilbert space $\mathscr{H}$ (operators are boldfaced).

---

[9]P. Dirac (1925): *Proc. R. Soc. A* **109**, p. 642

[10]Heisenberg, *op. cit.*

[11]Mehra & Rechenberg, *op. cit.*, vol. II, p. 184

[12]P. Dirac (1927): *Proc. R. Soc. A* **113**, p. 621; P. Jordan (1927): *Zs. F. Phys.* **40**, p. 809

[13]Cf. the letter from Pauli to Heisenberg d.d. October 19th, 1926 (W. Pauli (1979): *Wissenschaftlicher Briefwechsel*, vol. I (ed. by A. Hermann, K. von Meyenn and V. Weisskopf; Springer, Berlin), # [143]).

[14]P. Dirac (1930): *The Principles of Quantum Mechanics* (1st ed.; Oxford Univ. Press);
J. von Neumann (1932): *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin)

(2b) Every measurable quantity (*observable*) $\mathcal{A}$ is described by a self-adjoint operator A on $\mathcal{H}$.

(2c) The only possible result of the measurement of an observable $\mathcal{A}$ is an element of the set of the eigenvalues of the corresponding operator A (the *spectrum* $\sigma(A)$ of A).

(2d) When an observable $\mathcal{A}$ is measured, the probability of obtaining a result in the interval ΔA is given by $P(\Delta A) = \mathrm{Tr}[\rho\, E(\Delta A)]$, where $\{E(\Delta A)\}$ is the spectral family associated with the operator A by virtue of the spectral theorem (see (4) below).

(2e) If the measurement of an observable $\mathcal{A}$, corresponding to an operator A with discrete spectrum $\sigma(A)$, gives result $a \in \sigma(A)$, the state of the object system immediately after the measurement is given by[v] $E(\{a\})\rho E(\{a\})/\mathrm{Tr}[\rho\, E(\{a\})]$.

(2f) The time evolution of the density operator is unitary: i.e. there is a family $U(t)$ of unitary operators such that $\rho(t) = U(t)\rho(0)U^{\dagger}(t)$.

Like the classical set of states, the set of quantum states is convex, the one-dimensional projectors $|\psi\rangle\langle\psi|$ [we use Dirac notation] being the extreme states (*pure* states). It is, however, not a simplex: a decomposition

$$\rho = \sum_i w_i\, |\psi_i\rangle\langle\psi_i| \quad (w_i \geq 0\,;\ \sum_i w_i = 1) \tag{3}$$

of a mixed state into pure states is usually not unique[vi]. Therefore the ignorance interpretation of mixtures, viable in CM, runs into difficulties in QM[15].

The spectral theorem[16]

$$A = \int a\, E(\mathrm{d}a) \quad , \tag{4}$$

---

[15]E. Beltrametti & G. Casinelli (1981): *The Logic of Quantum Mechanics* (Addison-Wesley, Reading, Mass.), p. 11; J. Park (1968): *Am. J. Phys.* **36**, p. 211

[16]Holevo, *op. cit.*

which is used in (2d), uniquely associates a spectral family, or *projection valued measure* (PVM), $\{E(\mathrm{d}a)\}_{\sigma(A)}$ with a given self-adjoint operator. Because (2c) and the expectation value rule

$$\langle A \rangle = \int a\, P(\mathrm{d}a) = \int a\, \mathrm{Tr}[\rho\, E(\mathrm{d}a)] = \mathrm{Tr}(\rho\, A) \tag{5}$$

are consequences of (2d), a PVM is a more fundamental object than a self-adjoint operator[17]. Starting from a PVM rather than a self-adjoint operator has additional advantages, such as the removal of the restriction to real eigenvalues[18]. These arguments suggest the use of PVMs instead of self-adjoint operators. Indeed we shall use PVMs in the following, whenever appropriate.

The wording of the postulates (2) is distinctly operationalistic. All classical talk about 'properties' is absent, and replaced by such terms as 'measurement results'. Nevertheless, the classical roots of the new formalism surface in, e.g., (2e): the (natural) characteristics of an ideal measurement in CM are carried over into QM as a postulate. A measurement according to (2e) will give on repetition the same result with certainty. Such a measurement is called a *measurement of the first kind*[vii]. As it is impossible in QM, contrary to CM, to think of the outcome of the measurement as a property of the object *counterfactually*[19] (i.e. one cannot assume that the object would have had the outcome as a property even if the instrument had not been present), the fact that the measurement of the first kind can be interpreted as creating a property to the object, may be seen as an argument in favor of it. After all, if a measurement cannot be thought of as revealing a pre-existing value, it would seem to need at least the preparative attribute (2e) in order to be properly called 'measurement'. Therefore (2e) shifts the emphasis within the concept of 'measurement' from the determinative aspect to the preparative aspect. This new usage of 'measurement' seems,

---

[17]Cf. P. Dirac (1958): *The Principles of Quantum Mechanics* (4th ed.; Oxford Univ. Press), p. 37

[18]Holevo, *op. cit.*; J.-M. Levy-Leblond (1976): *Ann. of Phys.* **101**, p. 319

[19]The troubles which such an ignorance interpretation of the inevitable scatter in quantum measurements runs into, were known soon [M. Jammer (1974): *The Philosophy of Quantum Mechanics* (Wiley, NY), p. 43]

however, hardly in accord with practical definition, which certainly is more about finding out aspects of the object's state before measurement than about fixing a quantity for the future.

Another objection against the 'measurement' definition (2e) is that practical measurements are not required to be even remotely like (2e) (viz. the general usefulness of destructive measurements, especially in the micro-domain, where hardly any others are available)[20]. Therefore it seems more sensible to regard (2e) as a characterization of an ideal measurement, rather than as a definition of the term 'measurement' in full generality[21]. Strict adherence to (2e) in the description of actual measurements would then not be required. But the impossibility in QM of attributing the outcome of the measurement to the object counterfactually, actually means that the ontological arguments that favored the classical measurement ideal are invalid in QM. There is no conceptual basis for (2e) at all. Thus we see that (2e) and, more generally, the observable concept (2b) in fact originate in the analogy with CM rather than in an operational analysis (as is perhaps suggested by the operationalistically sounding nomenclature). Their status is dubious.

3 THE UNCERTAINTY PRINCIPLE

The development stage of "pioneer QM" ended with the discovery in 1927 of the *uncertainty principle*[viii] (UP) by Heisenberg[22]. He was puzzled by the apparent contradiction between on the one hand the impossibility to unite position and momentum representations in one picture ($[Q,P]_- = i1 \neq 0$) and on the other the "particle tracks" seen in a Wilson chamber[23]. Heisenberg first argued that QM is based on

---

[20]Cf. also Jammer, *op. cit.*, p. 487

[21]Beltrametti & Casinelli, *op. cit.*

[22]W. Heisenberg (1927): *Zs. f. Phys.* **43**, p. 172
Cf. also Dirac, *op. cit.* (1927), and the letter from Pauli to Heisenberg quoted earlier (Pauli, *op. cit.*, letter # [143])

[23]W. Heisenberg (1969): *Der Teil und das Ganze* (Piper, Munich), p. 111

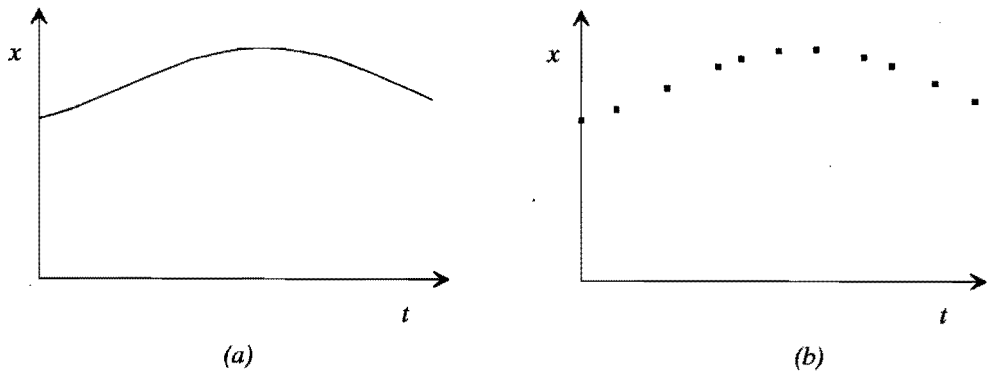


fig. 2 *Classical trajectory (a) vs. the kind of trajectory allowed by the discontinuities in QM (b). Figure taken from Heisenberg,* op. cit. *(1927).*

discontinuities, so that a trajectory of the classical type is no longer possible (fig. 2). In the quantum case[24] "[... ist es] offenbar sinnlos, von der Geschwindigkeit an einem bestimmten Orte zu sprechen, weil ja die Geschwindigkeit erst durch zwei Orte definiert werden kann, und weil folglich zu jedem Punkt je zwei verschiedene Geschwindigkeiten gehören". Thus, the sequence of points formed by the drops in a Wilson chamber does not jointly define position and momentum. Heisenberg next discusses a *γ–microscope* (fig. 3). In the γ–microscope light with wavelength λ is scattered off an electron to determine its position. The light is then collected by a lens with aperture ε onto a photographic plate[25]. The microscope's resolution is $\delta_q = \lambda/2\sin(\frac{1}{2}\varepsilon)$. On the other hand, when the photon reaches the plate, informing us of the electron's position, the direction from which it left the electron is unknown by an amount ε. This leads, *via* the Compton–recoil of the electron, to an uncertainty ("disturbance") $D_p \simeq 2\sin(\frac{1}{2}\varepsilon)/\lambda$ in momentum. Thus we have $D_p\,\delta_q \simeq 1$.

[24]Heisenberg, *op. cit.* (1927): "[...] it is clearly meaningless to speak about one velocity at one position because one velocity can only be defined by two positions, and conversely because any one point is associated with two velocities" [translation from J. Wheeler & W. Zurek (eds.) (1983): *Quantum Theory and Measurement* (Princeton University Press)].

[25]Heisenberg initially forgot to take the aperture into consideration, but was soon set straight by Bohr (see e.g. Jammer, *op. cit.*, p. 64)

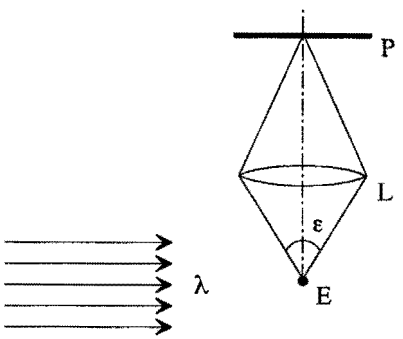


fig. 3 *Heisenberg's γ-ray microscope. Light with wavelength λ is scattered off an electron* E *through a lens* L *onto a photographic plate* P. *The lens has aperture* ε.

Lastly, Heisenberg shows that for states with a Gaussian position representation (viz. the ground state of the harmonic oscillator) the variances in position and momentum satisfy $\langle\Delta^2 Q\rangle\ \langle\Delta^2 P\rangle = \frac{1}{4}$; $\langle\Delta^2 P\rangle$ denotes $\langle(P-\langle P\rangle)^2\rangle$. Later this result was extended to the now familiar *Heisenberg inequality*[26]

$$\langle\Delta^2 P\rangle\ \langle\Delta^2 Q\rangle \geq \tfrac{1}{4} \quad , \tag{6}$$

and *Robertson inequality*[27]:

$$\forall_\rho\ \langle\Delta^2 A\rangle\ \langle\Delta^2 B\rangle \geq \tfrac{1}{4}\left|\textstyle\sum_i w_i\{\langle A\psi_i|B\psi_i\rangle - \langle B\psi_i|A\psi_i\rangle\}\right|^2 \tag{7}$$

$$(\rho = \textstyle\sum_i w_i|\psi_i\rangle\langle\psi_i| \quad ; \quad \langle\psi_i|\psi_j\rangle = \delta_{ij}) \quad .$$

Another common way[28] of introducing the UP is through reference to the *wave particle duality*, which quantum mechanics allegedly entails. One notes that for classical light a wave packet of size Δq must have a wave vector dispersion

$$\Delta k \geq \Delta q^{-1} \quad . \tag{8}$$

---

[26]H. Kennard (1927): *Zs. f. Phys.* **44**, p. 326; H. Weyl (1928): *Gruppentheorie und Quantenmechanik* (Hirzel, Leipzig)

[27]H. Robertson (1929): *Phys. Rev.* **34**, p. 163; K. Kraus & J. Schroeter (1983): *Int. J. Theor. Phys.* **7**, p. 431

[28]Cf. L. Rosenfeld (1971): *Arch. Hist. Exact Sci.* **7**, p. 69 (quoted on p. 59 of Wheeler & Zurek, *op. cit.*); Jammer, *op. cit.*

Since electrons are also supposed to have a wave nature this is then presented as the UP[ix]. This type of "derivation" has even prompted some to doubt whether the UP is specifically quantum mechanical[29]. This doubt is not justified. Such reasoning is based on the similarity in mathematical form of the two inequalities (6) and (8) rather than on any relation in physical content: the size of a classical wave packet (in either direct or reciprocal space) is not related to the uncertainty of the packet's position, just as the finiteness of a chair's size has no consequences for the exactness of its position[30]. As regards light, equations of the type $\Delta q\, \Delta k \geq 1$ are not proper analogs of $\langle \Delta^2 P \rangle \langle \Delta^2 Q \rangle \geq \frac{1}{4}$ for particles[x]. The commutation relations for photons are those of the *field variables*, and for these true uncertainty relations can be derived. Such relations restrict the precise, classical definability of the light field[31], and they are just as ununderstandable from a classical point of view as is the Heisenberg relation (6).

The UP is generally seen as one of the major ingredients of the new theory. Kennard called it[32] "der eigentliche Kern der neuen Theorie". After its discovery QM was essentially finished. From then on most focused their attention on applications of the formalism, and it has indeed proved extremely successful in that respect. As a consequence the presentation of the formalism in textbooks has changed only in minor ways since 1928. This holds especially true for the UP, which is still presented quite like Heisenberg himself did[33]. Popular interpretations, based on Heisenberg's reasoning include the statements that a measurement of some observable disturbs other, incompatible, observables (:= the *disturbance interpretation*: $D_p\, \delta_q \geq 1$; viz. the γ-microscope); that it limits the accuracy achievable in joint measurements of

---

[29]H. Primas (1983): *Chemistry, Quantum Mechanics and Reductionism* (2nd ed., Springer, Berlin), p. 151; M. Vol'kenshtein [M. Vol'kenshtein (1988): *Sov. Phys. Usp.* **31**, p. 140] quotes Mandel'shtam: "[...] the uncertainty principle can be easily explained to people who know radiotelegraphy".

[30]E. McMullin (1954): *The Principle of Uncertainty* (PhD Thesis, Cath. Univ. of Louvain, Belgium), unpublished

[31]Cf. N. Bohr & L. Rosenfeld (1933): *Mat.-Fys. Medd. Dan. Vidensk. Selsk.* **12**, no. 8; N. Bohr & L. Rosenfeld (1950): *Phys. Rev.* **78**, p. 794

[32]Kennard, *op. cit.*

[33]See e.g. A. Messiah (1955): *Quantum Mechanics*, vol. I (North Holland, Amsterdam); C. Cohen-Tannoudji, B. Diu & F. Laloe (1977): *Quantum Mechanics*, vol. I (Wiley, NY); A. Capri (1985): *Non-relativistic Quantum Mechanics* (Benjamin-Cummings, Menlo Park (CA)); T.-Y. Wu (1986): *Quantum Mechanics* (World Scientific, Singapore).

incompatible observables (:= the *inaccuracy interpretation*: $\delta_p \, \delta_q \geq 1$); that it precludes the existence of a trajectory for a quantum particle (viz. Heisenberg's first argument: fig. 2); that is forbids obtaining interference phenomena when the "path" is known[34]. Heisenberg's reasoning is by no means unproblematic, however. The three types of argument (γ–microscope, Wilson chamber and formal calculation) are not explicitly related, and it is not obvious that such a relation exists at all. Where are the "errors" δ_q or D_p in the quantum "track" of fig. 2 ? How can the quantity δ_q in the γ–microscope (purely a property of the measuring instrument) be related to the width $\langle \Delta^2 Q \rangle$ of the wave function, which is calculated without any reference to the measuring process ? How are the quantities $\langle \Delta^2 Q \rangle$ and $\langle \Delta^2 P \rangle$ related to the features of the quantum "track" ?

We saw that Heisenberg gave the *Gedanken* experiments a prominent role in his derivation of the UP. That has the unfortunate consequence of suggesting that the UP has a *physical* origin, that it is a consequence of the physical laws insofar as they govern the interaction between object system and measuring or preparing device[xi]. The derivation of the Heisenberg relation (6), on the other hand, does not refer to the details of the preparation or measurement process at all, let alone to the laws of optics, electrodynamics, etc., used in the description of the thought experiments. Ineq. (6) is *logically* inevitable rather than that it needs physical justification[35]: in the quantum language it is incoherent to talk about systems with states such that both P and Q are sharp. The UP can be considered to have a physical origin only when a new theory has been found from which QM can be derived. Such a sub-quantum theory would then explain this logic, explain incompatibility and explain (6). The (semi–) classical reasoning employed in the description of the imaginary experiments cannot be considered as sufficient for such a task.

---

[34]Cf. the categorisation of interpretations by Jammer, *op. cit.*, and by McMullin, *op. cit.*.

[35]Jammer, *op. cit.*, p. 160

For Heisenberg, in accord with the operationalistic maxim[36] "Wenn man sich darüber klar werden will, was unter dem Worte 'Ort des Gegenstandes', z.B. des Elektrons [...], zu verstehen sei, so muß man bestimmte Experimente angeben, mit deren Hilfe man den 'Ort des Elektrons' zu messen gedenkt", the experiments are important because they give content to the notion of 'position'. But then how can Heisenberg use the word 'momentum' in the discussion of the γ-microscope experiment when momentum is not measured (hence not defined) ? Heisenberg's operationalism is certainly not fully carried through, and Heisenberg can probably not be characterized as an operationalist[37]. From a more general methodological point of view, however, it is unclear whether these *Gedanken* experiments are intended[38] as derivations, explanations or illustrations of the UP. A rigorous derivation of the UP from the formalism should, I think, take precedence over other types of reasoning. That would restrict the use of the thought experiments to illustrations of the failure of classical concepts in quantum mechanics *for pedagogical purposes only*[39].

Therefore such assertions as the popular interpretations of the UP mentioned above (e.g. the disturbance and inaccuracy interpretations) must be formally justified in order to be acceptable. The only base of all of these claims in the formalism consists of the Heisenberg inequality (6) and the Robertson inequality (7). Heisenberg himself denotes the quantities δ_q (measurement accuracy), D_q (disturbance) and $\langle \Delta^2 Q \rangle$ (wave function width) by the same symbol (namely q_1), suggesting a conceptual identification of these notions. He was probably inspired by the classical theory in which, as we saw, preparative and determinative aspects of measurement were merged. From the point of view of the analogy with CM then, the assumption that a measurement's determinative quality (i.e. δ_q) conceptually equals its preparative quality (i.e. D_q , or $\langle \Delta^2 Q \rangle$ after measurement), is indeed tempting.

---

[36]"When one wants to be clear about what is to be understood by the words 'position of the object', for example of the electron [...], then one must specify definite experiments with whose help one plans to measure the 'position of the electron' " (Heisenberg, *op. cit.* (1927); translation taken from Wheeler & Zurek, *op. cit.*).

[37]Jammer, *op. cit.*, p. 58

[38]K. Popper (1972): *The Logic of Scientific Discovery* (6th rev. impr.; Hutchinson, London), app. *xi

[39]For Bohr the thought experiments were more important than this too, as is evident from a letter to Darwin in 1930 (N. Bohr (1985): *Collected Works*, vol. 6 (ed. by J. Kalckar; North Holland, Amsterdam), p. 316). See also ch. II.

Thus the CM analogy played a major role in the genesis of the opinion that the disturbance and inaccuracy interpretation are formally justified in (6) and (7). In fact, however, Born's statistical interpretation of quantum mechanics associates the quantity $|\langle a|\psi\rangle|^2$ with the probability (density) of outcomes in the measurement of a certain quantity. This implies that $\langle\Delta^2 A\rangle$ is the statistical dispersion of the distribution of such measurement outcomes (*scatter*). Consequently the interpretation immediately associated by Born's statistical interpretation with both (6) and (7) is that of a limit to the scatter in *independent* measurements: no joint measurement of A and B need be performed[40] to determine $\langle\Delta^2 A\rangle$ and $\langle\Delta^2 B\rangle$.

The aforementioned popular views on the UP assume a much wider applicability of (6) and (7). The inaccuracy interpretation, in particular, needs a connection between $\langle\Delta^2 A\rangle$ and A measurement inaccuracy. But the self-adjoint operators A and B occur in (7) in complete accordance with von Neumann's axioms: $\langle\Delta^2 A\rangle$ might even be realized as scatter in measurements of the first kind of A. 'Inaccuracy' can hardly be said to be involved in such a measurement. Murdoch[41] suggests that a mere reinterpretation of $\langle\Delta^2 A\rangle$ as the "uncertainty in our knowledge of A" or as the "real indefiniteness of the value of A" may help. But these are just rephrasings of the concept of 'scatter' in the epistemic and ontic interpretation of probability, respectively[xii]. They bring us no closer to a more general relevance of (6) and (7). A more serious attempt to establish the scatter-inaccuracy connection, and thus an inaccuracy interpretation of (6), was made by von Neumann[42]. He uses a C measurement of the first kind to effect a joint A,B measurement. The A c.q. B scatter in the state–after–measurement $|c\rangle\langle c|$, limited because of (7), is in his approach associated with measurement inaccuracy. But this reasoning is not satisfactory either. It uses a notion of 'measurement' that, in accord with the conceptual background of (2e) (see above), focuses exclusively on the preparative side of measuring, so that (7) becomes applicable. As a result this "joint" measurement in general does not even enable us to estimate the (pre-measurement) expectation values of A and B. A true inaccuracy interpretation would limit determinative measurement accuracy. Since (2e) can at best

---

[40]Popper, *op. cit.*

[41]D. Murdoch (1987): *Niels Bohr's Philosophy of Physics* (Cambridge University Press), p. 121

[42]von Neumann, *op. cit.*, § III.4

be regarded as a description of a measurement ideal (see above), its use is singularly inappropriate where the very nature of the problem demands a consideration of more realistic measurements. Even if we drop (2e), however, it can be seen that proposals intended to establish such a determinative inaccuracy interpretation cannot be based on the scatter in the state-after-measurement because there is no fundamental reason[43] why this scatter should be related to the measurement accuracy at all (ch. IV). In particular destructive measurements would not seem to be affected by an inaccuracy interpretation derived along these lines.

An interpretation of (7) as a reciprocal relation between the accuracy of an A measurement and the B disturbance uses, in addition to the interpretation of $\langle\Delta^2 A\rangle$ as an inaccuracy, the association of $\langle\Delta^2 B\rangle$ with 'disturbance'. But $\langle\Delta^2 B\rangle$ refers to the object state before measurement, instead of to the state-after-measurement in which evidence of a B disturbance would be expected to surface. In fact, the state transformation accompanying a measurement is not involved in the derivation of the Robertson relation (7) at all. Even if one accepts the measurement of the first kind postulate (2e), no 'disturbance' interpretation can be justified: after an A measurement the system is in an A eigenstate and its B distribution is in general not related at all to the B distribution before the measurement. Even if we are prepared to call, in the absence of a relation, a mere difference between the two B distributions 'disturbance', $\langle\Delta^2 B\rangle$ can hardly be used as a quantitative measure for it.

4 CONCLUSIONS

The concept of measurement of the first kind [(2e)] is, even when it is only regarded as a template for an ideal measurement, based on the analogy with CM rather than on an operational analysis. Similarly the association observable $\mapsto$ self-adjoint operator [(2b)] is plausible only from the point of view of that analogy. Therefore both are dubious. Indeed we shall show in the following (ch. III and IV) that (2e) and (2b) are

[43] K. Kraus (1987): *Phys. Rev. D* **35**, p. 3070

unnecessarily restrictive, that adopting them bars some interesting problems from being studied adequately.

As regards the UP, we accept the conclusion of Popper and others[44] that (6) and (7) can *only* be interpreted as statistical scatter relations. Therefore Heisenberg's γ-microscope argument (like the Wilson chamber argument, see ch. IV) is *not* related to (6). The general principle which the microscope is to represent is, like the inaccuracy interpretation, yet to be derived. In ch. III we will do precisely this for the inaccuracy interpretation. In ch. IV we will then see that a disturbance principle (γ-microscope) can be derived from these relations. Scatter relations and *inaccuracy relations* are actually independent (ch. III). This suggests that we can speak of a scatter *principle* on the one hand, and an inaccuracy *principle* on the other, constituting a dichotomic UP. This dichotomic UP will be seen to be sufficient to derive other alleged consequences of "the" UP, too.

First, however, we shall go into Bohr's interpretation of the new QM. This is of some importance, since he was the key figure in its development and since his views are still widely held to be authoritative. Thus it may seem that his investigations on the interpretation of QM could be of help for our problems with the UP. Furthermore, new developments with regard to the interpretation (such as those in ch. III) need to be evaluated in the light of Bohr's point of view, and *vice versa*.

NOTES

i Clifford Hooker (C. Hooker (1972): in *Paradigms and Paradoxes* (ed. by R. Colodny; Univ. of Pittsburgh Press), p. 67) describes the classical notion of measurement as follows (p. 72): "Knowledge of the states of physical systems is gained by the making of measurements on the systems. A measurement is a straightforward physical process of interaction between a measuring instrument and a measured system, the outcome of which is directly related to the feature of the system under investigation in a known way. [...] Measurement procedures are such that either they produce no significant disturbance of the measured system, or else such disturbances as are produced are precisely calculable and can be allowed for."

---

[44]Popper, *op. cit.*; Ballentine (1970): *Rev. Mod. Phys.* 42, p. 358; H. Groenewold (1946): *Physica* 12, p. 405

ii The non-commutative multiplication rule was already used implicitly by Kramers and Heisenberg in their last paper on dispersion theory prior to Heisenberg's 1925 paper on quantum theory (Mehra & Rechenberg, *op. cit.*, vol. II, ch. II). The fact that the quantities Heisenberg used were actually matrices was not realized by Heisenberg himself, but by Born.

iii The term 'observable' as a noun, and in the modern sense, was probably coined by Dirac (Dirac, *op. cit.* (1930), p. 25). Von Neumann, *op. cit.* (1932), used the word 'Groeβe' [quantity].

iv These postulates are usually attributed to von Neumann (e.g. Jammer, *op. cit.* (1974), p. 5). In his book (von Neumann, *op. cit.* (1932)) the postulates are introduced (in not precisely above form) on p. 168, p. 104, p. 105, p. 104 [for pure states], p. 113 and p. 186 respectively.

v Postulate (2e) is here given in the Lueders form (G. Lueders (1951): *Ann. der Phys.* **8**, p. 322), which is also suitable for self-adjoint operators with degenerate spectrum. Von Neumann, *op. cit.* (1932), assumed that such a transformation is not possible for operators with a continuous spectrum. The proof of this assertion is perhaps more involved than von Neumann had supposed, and was given only recently by Ozawa (M. Ozawa (1984): *J. Math. Phys.* **25**, p. 79; cf. App. A).
Curiously enough, von Neumann (*op. cit.* (1932), p. 110) quotes the Compton-Simon experiment as empirical evidence for (2e).

vi Von Neumann noted this non-uniqueness (von Neumann, *op. cit.* (1932), p. 175), but only for ρ's with degenerate spectrum. The decomposition is, however, also not unique for other ρ's (this is easily seen when one realizes that the $|\psi_i\rangle$'s in (3) need not be orthogonal). Nevertheless, von Neumann (probably prompted by the analogy with CM) continued to entertain the ignorance interpretation of mixtures in QM.

vii This name was introduced by Pauli (W. Pauli (1933): in *Handbuch der Physik* (2nd ed.; ed. by H. Geiger & K. Scheel; Springer, Berlin), vol. 24, § 9). A *measurement of the second kind* involves a "controlled change of the system".

viii We can only agree with Levy-Leblond [J.-M. Levy-Leblond (1973): *Encart Pedagogique* **1** (suppl. au *Bull. Soc. Fra. Phys.* **14**), p. 15] that what Heisenberg discovered is neither a "principle", nor is it about "uncertainty". We will nevertheless adhere to what has become common usage (but see ch. III).

ix Bohr is often quoted in support of such argumentation. Indeed he originally took wave-particle duality as a starting point for his philosophy [ch. II; Jammer, *op. cit.* (1974)]. In later years, however, he explicitly denied the wave nature of particles and the particulate nature of light any significance beyond mathematical form (ch. II). Thus he cannot rightly be considered as an advocate of this "derivation".

x Possibly conceptions along the lines of Schroedingers original interpretation of the wave function as a field (Jammer, *op. cit.* (1974), p. 24ff.) prompt such an attribution of too much physical significance to a mere mathematical analogy.

xi See e.g. the editorial comment preceding Bohr's 1928 *Nature* paper (*Nature* **121**, p. 579) or von Neumann, *op. cit.* (1932), p. 126. This view on the UP probably also inspired the many attempts to violate the principle by devising e.g a measurement without "disturbance" (cf. Jammer, *op. cit.* (1974), p. 59). The futility of such attempts can be seen when one realizes that there is no theory supporting calculations that falsify the principle (in any of its forms, see ch. III and IV): QM calculations *automatically* satisfy it. Therefore violation claims can only be based on some (semi-)classical intuition, and can be discarded.

xii The name 'scatter' for $\langle \Delta^2 A \rangle$ would probably appear most appropriate within the frequency interpretation of probability. We shall use it here, however, without committing ourselves to one specific interpretation of probability.

CHAPTER II

Bohr

Wovon man nicht sprechen kann, darueber muss man schweigen

Wittgenstein[1]

In order to compare Bohr's views on measurement theory in general, and on the UP in particular, with new developments (ch. III) it is necessary to first study his papers carefully and find out what his views exactly were. This is all the more necessary because these views are often misrepresented (even by his own pupils), probably as a result of Bohr's somewhat idiosyncratic style of writing. Bohr's philosophy goes by the name of *complementarity*[1]. Its first exposition was given by Bohr in his Como-lecture[2], written after having read Heisenberg's 1927 paper on the UP upon his return from a skiing trip[3]. Whereas Heisenberg started from 'discontinuity', Bohr (in 1927) took 'wave-particle duality' as basic. In the following years Bohr sophisticated his views further and further[4]. A consequence of this is that we must be careful with Bohr's earlier work (especially the Como lecture), as it may not adequately reflect complementarity in the form with which Bohr was eventually satisfied. Such care has not always been exercised in the literature.

Keeping this in mind, we can now proceed to a concise (and therefore necessarily schematic) overview of Bohr's philosophy. The first crucial ingredient of complementarity is the necessity to *understand* everything in terms of everyday language, of which the language of classical physics is a refined form. In fact[5], "the language of Newton and Maxwell will remain the language of physicists for all time". If a scientific theory is no longer expressible in such terms, this means that a full understanding [~ visualization] of the processes the theory describes is no longer possible. The applicability of the everyday concepts has become *limited*. It is *a priori*

---

[1]L. Wittgenstein: *Tractatus Logico-Philosophicus*, thesis 7

[2]N. Bohr (1927) **[Como Lecture]**: *Atti del Congresso Internazionale dei Fisici 1927, Como-Pavia-Roma* (Nicola Zanichelli, Bologna), p. 565

[3]See e.g. M. Jammer (1974): *The Philosophy of Quantum Mechanics* (Wiley, NY). The paper referred to is of course W. Heisenberg (1927): *Zs. f. Phys.* **43**, p. 172.

[4]See p. 110ff of E. McKinnon (1985): *Niels Bohr, a centenary volume* (ed. by A. French & P. Kennedy, Harvard University Press), p. 101

[5]N. Bohr (1931a) **[Maxwell Lecture]**: Nature **128**, p. 691

excluded that these concepts become *inapplicable*. The second ingredient of complementarity is the impossibility of the separation of object and observational device [sometimes called "subject" by Bohr]. Any attempt to further analyze this whole will impair the functioning of the observational device.

Thus presented, it is clear that complementarity is not specifically associated with QM, or indeed with physics. It is rather a general methodological framework. Therefore Bohr's suggestions for an application of complementarity in e.g. biology and psychology[6] are not *a priori* absurd. He did, however, fail to show that there is in these disciplines an empirical necessity for such an application, that in these theories there is a *fundamental* (as opposed to *practical*[ii]) restriction on the applicability of the concepts from everyday language ("$\hbar \neq 0$"). Without such a demonstration there is no reason to believe that an application of complementarity in these disciplines is more meaningful than application in, e.g., 18th century physics.

Within physics, complementarity is not restricted to QM. Bohr interprets relativity in terms of complementarity: relativity also limits the applicability of classical concepts, i.c. simultaneity[7]. Complementarity's main application, however, always was QM. There it leads to:

(i) Objectifying description in terms of the quantities from CM: the classical quantities, in as far as they are well-defined [cf. (ii)], are object-properties.

(ii) The measuring instrument must be described completely classically; the UP is not relevant for its working. The unanalyzability (~ indivisibility) of the object-meter system is symbolized by the UP. This principle shows that well-definedness of some classical quantities in the interaction inevitably leads to unanalyzability in others. This unanalyzability may, for instance, appear in the guise of an "uncontrollable momentum exchange".

---

[6]Bohr alludes to such applications of complementarity in many essays. See esp. N. Bohr (1933) [**Light and Life**]: *Nature* **131**, p. 423

[7]N. Bohr (1949) [**Einstein essay**]: p. 201 of the Schilpp volume [P. Schilpp (ed.) (1949): *Albert Einstein, Philosopher-Scientist* (Open Court ,Evanston IL), reprinted on p. 9 of J. Wheeler & W. Zurek (eds.) (1983): *Quantum Theory and Measurement* (Princeton University Press)]. See esp. p. 46 (quotations from the Wheeler & Zurek reprint).

The precise nature of the measuring instrument determines how well-defined the quantities in the interaction and, *as a consequence*, those describing the object [cf (i)] are.

(iii) Wave particle duality plays no role. Electrons are particles and light consists of waves.

(iv) The QM formalism, used for quantitative calculation, is unvisualizable (~ ununderstandable). It is only of symbolic (~ instrumentalistic) value.

(See appendix B for detailed textual evidence.)

Points (i) and (ii) are concretizations of the general ingredients mentioned earlier. The necessity of understanding in classical terms, even at the object level [(i)], is illustrated by Bohr's attitude towards free electron spin. Bohr thought at first that, because its magnitude is directly related to $\hbar$ and therefore not classical, free electron spin is not measurable at all. Only explicit calculations convinced him of the contrary, and even then he argued that there are severe restrictions to its measurement (app. B). Thus a Stern Gerlach device would, according to Bohr, be of no use in an electron spin measurement (but see ch. V).

The third and fourth point have been added because of the many misunderstandings surrounding them. As regards wave-particle duality, anything but (iii) would have made Bohr's point of view inconsistent. Complementarity entails, as we saw, a restriction on the applicability of classical concepts. 'Wave' and 'particle' are already mutually exclusive concepts on a classical level, and complementarity can only make them more so. Thus[8], "[...] the difference between matter and light is as fundamental in quantum theory as it is in the classical one". For Bohr the 'wave nature' of electrons (and similarly the particulate nature of light, epitomized by the 'photon' concept) can only be used in *symbolical* quantitative reasoning, analogous to the quantitative calculations in the Schrödinger formalism itself [cf. (iv)]. It has no realistic or visualizable significance. Therefore even the term 'wave particle duality' (ch. I), with its suggestion of symmetry between the two concepts in QM, is, strictly speaking, at variance with Bohr's point of view.

---

[8]Bohr, *op. cit.* (1939) p. 237

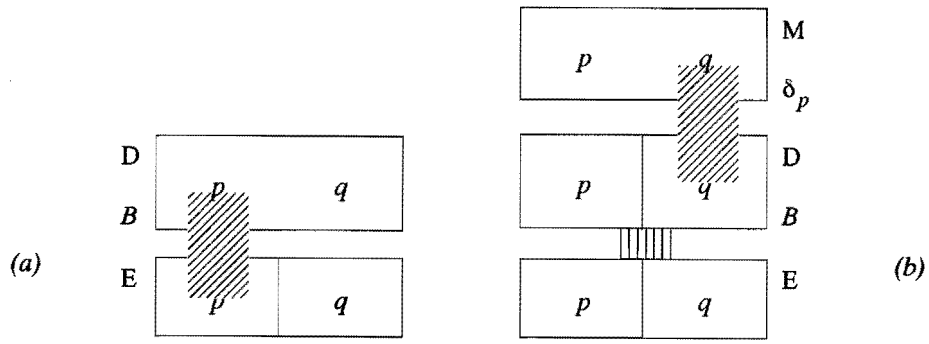


fig. 1 *Single slit thought experiment. The box* E *represents the electron. The uncertainty relations are applicable to it, making joint application of the concepts p and q impossible. This is indicated by a vertical line within the box. The box* D *represents the diaphragm (width B). It is a measuring device [a], so the uncertainty relations are not applicable to it. If we apply a p meter to it [b] (box* M*, inaccuracy* δ_p*), it becomes an object. The shading (▨) indicates unanalyzability of object-subject interaction. (Another type of shading [|||] is used to denote interaction between objects.)*

A first illustration of (i) and (ii) can be given by means of a discussion of an electron passing through a diaphragm (single slit) with width B (fig. 1.a). The diaphragm is classical, as it is a measuring device [(ii)]: both p and q are definable. The interaction with the object, however, is not completely analyzable. The diaphragm is tightly bolted to its support. The amount of momentum going into the support is indeterminate (beyond $\hbar/B$). Momentum is not conserved. Therefore the p picture is not applicable to the interaction. As a consequence the object's p is not defined either; the uncertainty relations are relevant for the object. In fact the momentum is "disturbed": even if the initial object momentum is defined accurately, the indefiniteness of the interaction in p picture causes an indeterminacy of the order $\hbar/B$ in the output momentum. We see here an operationalistic element in Bohr's philosophy: the object properties are defined through their relation to the measuring apparatus, and are well-defined as far as the object-subject interaction is well-defined.

The situation can be improved upon in a number of ways. We can make the slit narrower, thus improving the definition of output position at the cost of momentum

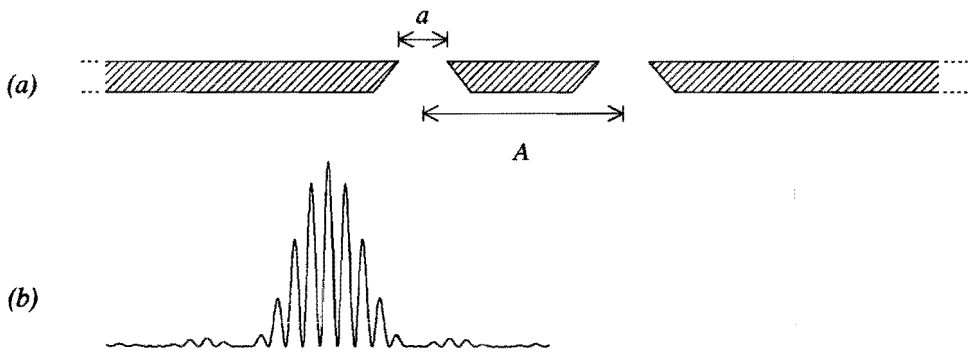


fig. 2 *The double slit [a]: two slits of width a separated by a distance A. If we let an electron in a monochromatic state impinge on this device, an interference pattern results that consists of narrow peaks, modulated by a wider curve [b].*

disturbance. Momentum disturbance may be reduced by widening the slit. We may also think of a cleverer scheme: measure the diaphragm's momentum before and after the passage of the electron and know how much the electron's momentum has changed. Of course there is a catch. In the new situation (fig. 1.b) the diaphragm has itself become an object. The diaphragm-momentum meter interaction is not fully determinate. The smaller the momentum meter's inaccuracy is, the larger the indeterminacy in diaphragm position must become. Knowledge of diaphragm momentum (giving knowledge about the electron momentum change) can only be obtained at the cost of definition of slit position (directly connected to definition of electron output position). We see how indivisibility of the object-meter system contextualizes the objectivism of the description [cf. (i)].

Now consider the double slit (fig. 2). Bohr denotes the two most important complementary pictures by the labels 'space-time coordination' and 'conservation laws'. In extreme cases (Bohr focuses mostly on these), either one is fully applicable at the cost of the other. In the double slit, however, we have neither. On the one hand the slit is firmly fixed to the ground, invalidating momentum conservation. On the other hand, we do not know through which slit the particle went, so that we do not have complete space-time coordination either. Therefore the description of this experiment

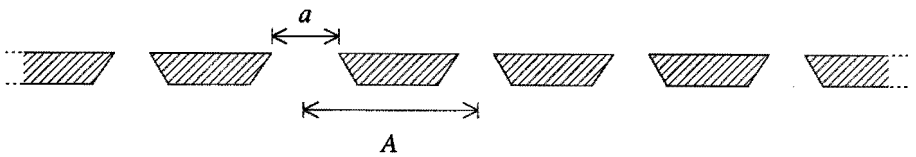


fig. 3 *The double slit (fig. 2) discussion can be improved by extending the two-slit device to a "grating", a system of slits of width a, with period A.*

along the lines of Bohr is somewhat more involved than he himself suggests, especially since he cannot consistently use wave-particle duality [(iii)]. To give an adequate discussion of this situation, we shall first extend the double slit to an infinite periodical (period A) array of slits (width a) (see fig. 3). We let an electron, represented by some quantum state with position width $\langle\Delta^2 Q\rangle$ ($\sim$ quality of input position definition; we assume $\langle\Delta^2 Q\rangle > A$), impinge on this "grating". The double slit then corresponds to a situation where the state spans only two slits. The interference pattern (in p picture) here consists of a number of peaks (width of the order $2\pi\hbar/\langle\Delta^2 Q\rangle$) separated by a distance $2\pi\hbar/A$. This set of peaks is modulated by a curve of width of the order $2\pi\hbar/a$. We can now discern two notions of position: we may want to find out where within a period the particle is without caring in which period it is ("little position", Q mod A), or we may be interested only in the period in which the particle is ("big position", entier$[Q/A]$). The former is conjugated to the number of the interference peak the particle ends up in ("big momentum", entier$[PA/2\pi\hbar]$), the latter to the particle's momentum within a peak, ignoring which peak it is in ("little momentum", P mod $\hbar/2\pi A$). We see that there are in fact two different complementarities in the double slit experiment, the first characterized by the modulation width of the interference pattern (big momentum indeterminacy, of the order $2\pi\hbar/a$) versus slit width (little position indeterminacy, equal to a), the second by interference peak width (little momentum indeterminacy, of the order $2\pi\hbar/\langle\Delta^2 Q\rangle$) versus initial position width (big position indeterminacy, equal to $\langle\Delta^2 Q\rangle$). In accordance with this, we may attempt to shift the balance of definition in either of them. First, we can *either* make the slits narrower *or* apply momentum conservation (cf. the diaphragm discussion: we may measure p before and after with inaccuracy at most of the order $2\pi\hbar/a$). Second, we can *either* narrow down the initial

position width (this corresponds to Bohr's proposal to close one slit) *or* apply momentum conservation (measure p before and after with inaccuracy at most of the order $2\pi\hbar/\langle\Delta^2 Q\rangle < 2\pi\hbar/A < 2\pi\hbar/a$). The fact that momentum conservation can here be used to improve the definition of both big and little momentum is somewhat confusing. Moreover, the rigid connection between support and slit-system does not imply *full* applicability of the space picture, as Bohr sometimes seems to suggest. Nevertheless above reasoning shows that the lack of both momentum conservation and space-time coordination presents no real problem for complementarity.

Similarly, on might conceive of the plan to improve the γ-microscope (ch. I) by measuring its momentum before and after the photon was scattered off the electron[9]. Again the uncertainty relations become applicable to the microscope, blurring the definition of its position, thus reducing its accuracy. (Further *Gedanken* experiments are discussed in appendix B.)

It is important to see that the aforementioned four points only reflect complementarity in its *application* to QM. As a general methodological principle, it will retain its validity when QM has become obsolete[10]. Thus Bohr can claim that[11] "such argumentation does of course not imply that, in atomic physics [i.e. microphysics], we have nothing more to learn as regards experimental evidence and the mathematical tools appropriate for its comprehension. In fact it seems *likely* that the introduction of *still further abstractions* [twice my italics] into the formalism will be required to account for the novel features revealed by the exploration of atomic processes of very high energy. The decisive point, however, is to recognize that in this connection there is no question of reverting to a mode of description which fulfills to higher degree the accustomed demands regarding pictorial representation of the relationship between cause and effect". According to Bohr, new theories will still be interpreted in classical terms [viz. (i) and (ii)], although the classical concepts will become more and more restrictedly applicable. Physical theories can be seen as ever more

---

[9]Bohr, *op. cit.* (1928)

[10]McKinnon, *op. cit.*, p. 119

[11]N. Bohr (1958): *Philosophy in the Mid-Century* (ed. by R. Klibansky, La Nuovo Italia, Florence) p. 308 [reprinted on p. 1 of N. Bohr (1963): *Essays 1958-1962 on Atomic Physics and Human Knowledge* (Wiley, NY)]. See p. 2 (quotations from the 1963 reprint).

encompassing *generalizations* of the classical theories (viz. the correspondence principle). This illustrates nicely an interesting methodological side to the Bohr-Einstein debate. It is commonly held that the advent of a new physical theory has for the previously reigning ones the consequence that they should now be deduced from the new, deeper, theory (*foundationalism*[12]). But the "deduction" is not quite trivial because the concepts occurring in the old theories generally acquire a different meaning when seen from the new theory. So, for instance, 'mass' in relativity theory is something quite different from 'mass' in CM. Feyerabend[13] speaks of an *incommensurability*. This incommensurability prevents the development of science from being seen as a continuous process from a formal methodological point of view. For Bohr, however, there is a continuity in the meaning of the fundamental concepts: despite the fact that *calculations* using the old theory are superseded by calculations using the new theory, all *interpretation* of the new theory must take place in terms of the old one, in classical terms even. For Bohr no shift in meaning of the concepts involved can possibly take place. Bohr's reaction on a proposal by Eddington to change the definitions of certain fundamental concepts, characterizes his attitude. Bohr finds Eddington's proposal unacceptable because[14] "for [a physicist] all these ideas already had a practical significance, simple and well-defined". Bohr, contrary to e.g. Einstein and von Neumann, is not a foundationalist[iii].

As we saw earlier (ch. I), from a classical point of view the distinction between preparation and measurement is not quite obvious. Indeed Bohr was aware of the distinction. He considered both "kinds of measurement" necessary for a complete specification of the context, for a complete phenomenon (app. B). Nevertheless, he does not use a differentiating nomenclature. The word 'determine' can in Bohr's papers mean both 'measure' and 'prepare'. The diaphragm is for Bohr just as much a measuring device as the γ-microscope, although in the former case there is no

---

[12]McKinnon, *op. cit.*, p. 118

[13]P. Feyerabend (1981): *Realism, Rationalism and Scientific Method*, vol. I (Cambridge University Press); p. 45ff

[14]See p. 204 of N. Bohr (1939) **[Warsaw Lecture]**: *New Theories in Physics* (International Institute of Intellectual Co-operation, Paris) p. 11. Cf. Bohr's statement with "Introducing a new theory involves [...] changes in the meaning of even the most 'fundamental' terms of the language employed." (Feyerabend, *op. cit.*, p. 54).

pointer and no outcome. Such language is not inconsistent, but it does tend to obscure the problems concerning the interpretation of the Heisenberg relation noted in ch. I. Indeed, a naive reading of Bohr (without keeping the just named qualification in mind) may well lead to a non-scatter interpretation of the Heisenberg relation (I.6). More generally, Bohr's presentation not only makes the relations "$\langle\Delta^2 Q\rangle\ \langle\Delta^2 P\rangle \geq \frac{1}{4}\hbar$" and "$\delta_p\ \delta_q \geq 1$" look plausible, but also "$\langle\Delta^2 Q\rangle\ \delta_p \geq 1$" and "$\langle\Delta^2 P\rangle\ \delta_q \geq 1$" [viz. ch. III]. For Bohr, however, 'the UP' is not so much the mathematical relation (I.6) [or any other mathematical relation] as it is a conceptual statement about the joint applicability of classical pictures. As such it is an integral part of the *interpretation* of QM, rather than of the *formalism* of QM. The concrete relation (I.6) is a consequence of the formalism that merely *reflects* one aspect of (not *constitutes*) the Bohrian UP. Bohr can consequently not be embarrassed at all by the limitations of the Heisenberg inequality[15] (I.6). His view on the UP is closely connected to his general style of interpreting QM (or other non-classical theories, see above). In contrast to e.g. von Neumann and Einstein, Bohr interprets QM not by associating elements of reality with elements of the theory. His interpretation is not based on properties of the formal mathematical scheme, but rather on simple qualitative considerations only loosely connected to the formalism[16]. This interpretational style is of course the consequence of his insistence on understanding in terms of classical concepts which precludes a more than instrumentalistic significance for the formalism [cf. (iv)], and (more generally) of his non-foundationalist methodology.

Thus Bohr cannot really help us with the interpretational problems surrounding the UP that were noted in ch. I. We are obliged to use formal calculations to check which of the possible forms of the UP can be justified[iv]. In the following two chapters we shall pursue that line of investigation further, before returning to Bohr with the insight then gained (ch. V).

---

[15]Cf. H. Casimir (1986): in *The Lesson of Quantum Theory* (ed. by J. de Boer, E. Dal & O. Ulfbeck, North Holland, Amsterdam), p. 13. Casimir points out that Bohr never committed himself to the particular mathematical form of (I.6).

[16]Feyerabend, *op. cit.*, p. 275. See also W. Heisenberg (1967): in *Niels Bohr: his life and works as seen by his friends and colleagues* (ed. by S. Rozental, North Holland, Amsterdam) p. 94; see esp. p. 98.

NOTES

i Around 1929 there was a brief period where Bohr considered 'reciprocity' to be a more appropriate name. See e.g. N. Bohr (1929c) **[Planck paper]**: *Naturwiss.* **17**, p. 483 and p. 277 of N. Bohr (1985): *Collected Works*, vol. 6 (ed. by J. Kalckar; North Holland, Amsterdam).

ii As regards biology, Bohr later [N. Bohr (1960): Address at the International Congress of Pharmaceutic Sciences in Copenhagen; N. Bohr (1962) **[Light and Life Revisited]**: Address at the inauguration of the Institute for Genetics in Cologne; reprinted on p. 17 and p. 23 of Bohr, *op. cit.* (1963) respectively] seems to have taken back his original statements as to the complementarity of the integrity of a living organism versus the possibility of analyzing its atomic structure (M. Vol'kenshtein (1988): *Sov. Phys. Usp.* **31**, p. 140). In view of the developments in molecular biology, Bohr saw that the mutual exclusiveness of these two concepts is merely a matter of our experimental ingenuity.

iii Because von Neumann is a foundationalist, the "derivation" of CM from QM is sensible to him. So is the measurement problem. For Bohr, on the other hand, this problem does not exist [Bohr, *op. cit.* (1939) p. 97; Cf. Jammer, *op. cit.*, p. 472 and Hooker [C. Hooker (1972): *Paradigms and Paradoxes* (ed. by R. Colodny, Pittsburgh University Press) p. 67], p. 159

iv One can conjecture that Bohr would have found the inaccuracy relation (ch. III) self-evident, and that he therefore was not worried by the fact that it had not been derived. Nevertheless it cannot be excluded that Bohr himself, using his extreme skill in obtaining qualitative results using thought experiments [Feyerabend, *op. cit.*, p. 275], could indeed have investigated the inaccuracy interpretation and other possible concrete forms of the UP separately, to check whether they hold. It seems safer for us to try a formal calculation first, however.

CHAPTER III

The Uncertainty Principle : Formal Aspects

In the previous chapters we have given a short description of the genesis of the UP and its early interpretation, most notably by Bohr. We also saw that the best known representative of the *principle* in the formalism, i.e. the Heisenberg *relation*, can only be interpreted as limiting statistical scatter. We study the *scatter interpretation* of the UP more closely [§ 1]. We shall see that there are many scatter relations, utilizing different notions of 'width' of the probability distribution. Some of these relations are stronger than the Heisenberg relation.

Bohr, however, talked about the 'uncertainty principle' in a much wider sense. Unfortunately, he did not rigorously derive any non-scatter inequalities. From his point of view such a derivation was probably spurious. Since a formal justification seems nevertheless interesting, we shall study one possible non-scatter application of the UP: joint measurement of incompatible observables [§ 2]. Indeed we will be able to rigorously derive an inaccuracy inequality for such joint measurements, thus proving an *inaccuracy interpretation* of the UP.

Finally, the relations between the scatter and the inaccuracy interpretation are studied [§ 3]: they are found to be quite independent. This justifies the names 'scatter *principle*' and 'inaccuracy *principle*'. As other interpretations of the UP can be seen to be consequences of either of these two principles [ch. IV], we can speak of a dichotomic UP.

1 THE SCATTER PRINCIPLE

The scatter principle is the notion that statistical dispersions of probability distributions of measurement outcomes of incompatible observables (self-adjoint operators) on similarly prepared systems, cannot both be arbitrarily small. Concrete lower bounds for the scatter *principle* are given by scatter *relations*. The most popular of these are of course the Heisenberg relation

(1) $\langle \Delta^2 P \rangle \, \langle \Delta^2 Q \rangle \geq \frac{1}{4}$

and the Robertson relation

$$(2) \qquad \forall_{\rho = \Sigma_i w_i |\psi_i\rangle\langle\psi_i|,\ \langle\psi_i|\psi_j\rangle = \delta_{ij}} \langle\Delta^2 A\rangle\ \langle\Delta^2 B\rangle \geq \tfrac{1}{4}\left|\sum_i w_i\{\langle A\psi_i|B\psi_i\rangle - \langle B\psi_i|A\psi_i\rangle\}\right|^2 .$$

The scatter principle is the best known expression of the UP and, until recently, the only one formally justifiable. For this reason it is perhaps surprising that there are no standard *Gedanken* experiments connected with it (although the diaphragm is related [ch. IV]). Still, an informal illustration of ,e.g., ineq. (1) is not difficult to find. Note first that the quantum state (density operator) can be thought of as an abstract representation of a source of quantum systems. A beam of non–interacting particles emerging from such a *preparator* can be used as a suitable "ensemble" for testing probabilistic predictions. For such a beam of particles of known energy, the Heisenberg relation (1) relates minimum diameter to divergence: it is impossible to construct the source such that we get a beam that is arbitrarily narrow if its divergence must not to exceed a certain limit. The quantities $\langle\Delta^2 Q\rangle$ and $\langle\Delta^2 P\rangle$ can be seen as a characterization of the quality of the preparator: they are associated with width and divergence of a particle beam, respectively. In that sense, they are properties of the particle source (preparator). Relation (1) clearly is of interest to builders of cyclotrons and the like, rather than to manufacturers of particle detectors. We conclude that scatter inequalities limit preparation possibilities.

1.1 Non–standard Scatter Relations

Relations (1) and (2) are not the only scatter relations available. There are many others[1]. That is fortunate because (2) is not suitable for all cases. For example, for discrete spectra, such as occur on finite dimensional Hilbert spaces, the Robertson

[1]See for a review for instance V. Dodonov & V. Man'ko (1989): *Invariants and the Evolution of Nonstationary Quantum Systems* (ed. by M. Markov, Nova Science Publishers, Commack, NY), p. 3

inequality (2) is quite useless[2], even if the unsatisfactory nature of variances[3] is ignored (but cf. the remarks concerning single and double slit experiments further below). Consider the spin ½ case, where $A = \sigma_1$; $B = \sigma_3$ (Pauli operators). Then one would expect a meaningful scatter relation to assert at least that, if the state vector is such that σ_1 is sharp, σ_3 is completely undetermined. Instead, (2) gives in this case no bound for the scatter in σ_3 at all. More generally, the bound in (2) becomes trivial for any state for which the sesquilinear form on the right hand side is zero.

The recently derived Maassen–Uffink relation[4] (actually it is one of a whole class of scatter relations derived by them) is an entropic scatter relation for the finite dimensional case $\mathcal{H} = \mathbb{C}^n$ without such drawbacks:

$$(3) \qquad H_X[\rho] + H_Y[\rho] \geq -\log(F) \quad .$$

Here X and Y are self-adjoint operators associated with the orthonormal bases $(|x_k\rangle)_K$ and $(|y_k\rangle)_K$ on $\mathcal{H}$, respectively:

$$X = \sum_{k\in K} k|x_k\rangle\langle x_k| \; ; \; Y = \sum_{k\in K} k|y_k\rangle\langle y_k| \; ; \; K = \{0, \ldots, n-1\},$$

and:

$$H_X[\rho] := -\sum_{k\in K} \langle x_k|\rho|x_k\rangle \log \langle x_k|\rho|x_k\rangle$$

$$F := \max_{k,k'\in K} |\langle x_k|y_{k'}\rangle|^2 \quad .$$

An entropic scatter relation also exists for the position–momentum case[5]:

$$(4) \qquad H_P[\rho] + H_Q[\rho] \geq \log(e\pi)$$

---

[2]E. Beltrametti & G. Casinelli (1981): *The Logic of Quantum Mechanics* (Addison-Wesley, Reading, Mass.), § 3.4

[3]J. Uffink (1990): *Measures of Uncertainty and the Uncertainty Principle* (PhD Thesis, University of Utrecht, Netherlands), unpublished; J. Uffink & J. Hilgevoord (1985): *Found. Phys.* **15**, p. 925

[4]H. Maassen & J. Uffink (1988): *Phys. Rev. Lett.* **60**, p. 1103

[5]W. Beckner (1975): *Ann. Math.* **102**, p. 159; Maassen & Uffink, *op. cit.*

with:

$$H_P[\rho] := -\int dp\, \langle p|\rho|p\rangle \log \langle p|\rho|p\rangle$$

(analogously for Q).

Scatter relations of a different kind result from the application of the inequality[6]

$$\arccos\left[\sqrt{\langle A\rangle}\right] + \arccos\left[\sqrt{\langle B\rangle}\right] \geq \arccos\left[\sqrt{\|AB\|}\right] \quad , \tag{5}$$

which holds for any two projectors A and B, to two PVMs. We can, for example, apply (5) to the position-momentum case

$$A := \int_{p_0-\frac{1}{2}Dp}^{p_0+\frac{1}{2}Dp} |p\rangle\langle p|\, dp \quad , \quad B := \int_{q_0-\frac{1}{2}Dq}^{q_0+\frac{1}{2}Dq} |q\rangle\langle q|\, dq \quad . \tag{6}$$

There we have[7]

$$\|AB\| = \lambda(\tfrac{1}{4}\, Dq\, Dp) \tag{7}$$

with

$$\lambda(x) = \frac{2x}{\pi}\left(R_{00}^{(1)}(x,1)\right)^2 \quad (R_{00}^{(1)} \text{ is a } \textit{prolate radial wave function}).$$

Thus relations (5) and (7) give rise to the Slepian-Landau-Pollak (SLP) inequality for position and momentum[8]:

$$W_{Q,\alpha}[\rho]\, W_{P,\beta}[\rho] \geq \begin{cases} 0 & \text{if } \alpha+\beta \leq 0 \\ 4\lambda^{-1}\left(\left[\tfrac{1}{2}\sqrt{(1+\alpha)(1+\beta)} - \tfrac{1}{2}\sqrt{(1-\alpha)(1-\beta)}\right]^2\right) & \text{otherwise.} \end{cases} \tag{8}$$

---

[6]H. Landau & H. Pollak (1961): *Bell Syst. Techn. J.* **40**, p. 63

[7]D. Slepian & H. Pollak (1961): *Bell Syst. Techn. J.* **40**, p. 43

[8]Ineq. (8) implies the "support property" [P. Busch & P. Lahti (1985): *Phil. Sc.* **52**, p. 64]; see Uffink, *op. cit.*.

The *overall width* W is defined as

$$(9) \qquad W_{Q,\alpha}[\rho] \;:=\; \min\Big\{\, |q_1 - q_2| \;\Big|\; \int_{q_1}^{q_2} \langle q|\rho|q\rangle \, dq \;\geq\; \tfrac{1}{2} + \tfrac{1}{2}\alpha \Big\} \qquad (-1 \leq \alpha \leq 1).$$

The scatter relation (8) can be used for a discussion of the single slit thought experiment. We assume the particle's initial state to be monochromatic, and further assume that the slit acts as a perfect filter. The particle's state $|\phi\rangle_f\langle\phi|$ after the slit is then

$$\langle q|\phi\rangle_f \;\sim\; \begin{cases} \exp(ip_0 q) & \text{if } -\tfrac{1}{2}A \leq \tfrac{1}{2}A \;; \\ 0 & \text{otherwise.} \end{cases}$$

In this situation $W_{Q,1}[|\phi\rangle_f\langle\phi|] = A$ and $W_P[|\phi\rangle_f\langle\phi|] = O(\pi/A)$, as is easily verified[9]. Note that the standard Heisenberg relation is completely irrelevant for this case because $\langle\Delta^2 P\rangle$ does not exist for the state $|\phi\rangle_f\langle\phi|$.

Define also the the *fine structure width* of the position probability distribution:

$$(10a) \qquad \tilde{w}_{Q,\alpha}[\rho] \;:=\; \min\Big\{\, |\theta| \;\Big|\; \int_{-\infty}^{\infty} \sqrt{\langle q|\rho|q\rangle\langle q+\theta|\rho|q+\theta\rangle}\, dq \;\leq\; 1-\alpha \Big\} \;.$$

It can be shown[10] that:

$$(11) \qquad \tilde{w}_{Q,\alpha}[\rho]\; W_{P,\beta}[\rho] \;\geq\; C(\alpha,\beta) \;;$$

with:

$$C(\alpha,\beta) \;=\; 4 \arccos\sqrt{\frac{2-\alpha}{1+\beta}} \qquad (\alpha + \beta \geq 1).$$

---

[9] Uffink & Hilgevoord, *op. cit.*

[10] J. Hilgevoord & J. Uffink (1988): in *Proceedings of the International Conference on Microphysical Reality and Quantum Description, Urbino, Italy* (ed. by F. Selleri, A. van der Merwe & G. Tarozzi, Reidel, Dordrecht), p. 91; Uffink & Hilgevoord, *op. cit.*.

Just as (8) was needed for the diaphragm, the scatter relation (11) is useful for a discussion of certain aspects of the double slit *Gedanken* experiments[11] [cf. also § IV.1]. Consider again a monochromatic state impinging on the device of fig. II.1 ($A \gg a$). Then[12], on the one hand

$$W_{Q,1}[|\phi\rangle_f\langle\phi|] = A + a \quad \text{and} \quad \tilde{w}_P[|\phi\rangle_f\langle\phi|] = \mathcal{O}(\pi/A) \quad ,$$

on the other

$$\tilde{w}_Q[|\phi\rangle_f\langle\phi|] = \mathcal{O}(a) \quad \text{and} \quad W_P[|\phi\rangle_f\langle\phi|] = \mathcal{O}(\pi/a) \quad .$$

Note that, again, the standard Heisenberg inequality is wholly unsuitable to exhibit either of these reciprocities.

Based on the fine structure width concept, we can easily derive a further scatter relation. Define first the *interior width* of the Fourier transform of the momentum probability distribution:

$$(10b) \qquad w^F_{P,\alpha}[\rho] := \min\left\{|\theta| \;\middle|\; \left|\int_{-\infty}^{\infty} \langle p|\rho|p\rangle \exp(i\theta p)\, dp\right| \leq \alpha^2\right\} \quad .$$

Now, the quantity $1/w^F_P$ can be used as a measure for momentum scatter. The use of such a scatter measure may appear somewhat strange at first sight, but this use of parameters of the Fourier-transform of the momentum probability distribution is not uncommon, e.g. for angle variables[13]. Thus, the relation

$$(12) \qquad \tilde{w}_{Q,\alpha}[\rho] \geq w^F_{P,\alpha}[\rho] \quad ,$$

which is particularly easy to derive for pure states, is in fact a scatter relation !

---

[11]Hilgevoord & Uffink, *op. cit.*

[12]Uffink & Hilgevoord, *op. cit.*

[13]A. Holevo (1982): *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam); J.-M. Levy-Leblond (1976): *Ann. of Phys.* **101**, p. 319

The different scatter relations we have just presented are interrelated. The entropic inequality (4) is conceptually stronger than (1), as we have[14]

$$\exp[2\, H_{P}(\rho)] \;\leq\; \langle \Delta^2 P \rangle\, 2\pi e \quad . \tag{13}$$

The SLP inequality (8) is also conceptually stronger than (1): application of Chebyshev's inequality to (8) leads to (1), though with a non-optimal bound. In a similar vein, the relation (11) is stronger than (8) in the sense that (8) can be derived from (11) using theory independent mathematics (we have[15] $\tilde{w}_{Q,\alpha} \leq W_{Q,\beta}$ if $(1-\alpha)^2 + (\frac{1}{2}+\frac{1}{2}\beta)^2 \geq 1$; the lower bound is again not optimal). But, reasoning along these lines, (12) is even stronger than (11), as in fact the *theory-independent* relation

$$w^{\mathrm{F}}_{P,\alpha}[\rho]\; W_{P,\beta}[\rho] \;\geq\; C(\alpha,\beta) \quad , \tag{14}$$

derivable along similar lines as (11), can be used to establish (12)⟹(11).

1.2 Shift–scatter Relations

Consider the case where (at least) one of the self-adjoint operators involved in a scatter relation can be related to the group action generated by the other. In the position-momentum case, for example, the unitary operator $\exp(i\theta P)$ induces a position shift. Then we can derive relations, which cannot properly be called 'scatter relations', but are nevertheless closely connected to the scatter principle. Define

$$w_{q,\alpha}[|\phi\rangle\langle\phi|] \;:=\; \min\Big\{|\theta| \;\Big|\; \Big|\int_{-\infty}^{\infty} \langle q+\theta|\phi\rangle\langle\phi|q\rangle\, dq\Big| \leq \alpha^2\Big\} \tag{15}$$
$$(0 \leq \alpha \leq 1).$$

---

[14] See e.g. Maassen & Uffink, *op. cit.*; Uffink, *op. cit.*.

[15] Uffink & Hilgevoord, *op. cit.*

It can be shown[16] that, analogously to (11) and (14):

$$(16) \qquad w_{q,\alpha}[|\phi\rangle\langle\phi|]\, W_{P,\beta}[|\phi\rangle\langle\phi|] \geq C(\alpha,\beta) \quad .$$

Relation (16) links an overall width W to a quantity w. The meaning of the latter is not immediately clear. At first, it may be thought to represent some kind of fine structure width for the position probability distribution, perhaps analogously to (10a). That it fails to do this is apparent if one considers the following example:

$$(17) \qquad \langle\psi|q\rangle = c \exp(i\left[\frac{q}{a_1}\right]^2 - \left[\frac{q}{a_3}\right]^2) \cos\left[\frac{q}{a_2}\right] \quad (a_1 < a_2 < a_3).$$

For the state $|\psi\rangle\langle\psi|$, the position fine structure $\tilde{w}$ has a characteristic size a_2, independent of a_1. On the other hand, $w_q \leq \mathcal{O}(a_1)$, which is not related to the position distribution at all.

In order to give a proper interpretation to w, we first define a family of position-shifted versions of one state ρ:

$$(18) \qquad \{\rho_\theta\}_{\mathbb{R}} \;\; ; \;\; \rho_\theta = S_q(\theta)\, \rho\, S_q^\dagger(\theta) \quad , \quad \rho \text{ a density operator.}$$

Here S_q is the position shift operator, defined by

$$(19) \qquad S_q(\theta) := \int_{-\infty}^{\infty} |q\rangle\langle q+\theta|\, dq = \exp(i\theta P) \quad .$$

We next look for a quantity that characterizes the distinguishability of these states, that characterizes the quality with which measurements on a member of this family allow an estimation[17] of the shift parameter θ. Suppose we have a PVM $\mathfrak{e} = \{E(d\eta)\}_H = \{E(\eta)d\eta\}_H$. Define as a number characterizing the quality with which a measurement of this PVM discerns[18] different values of θ:

$$(20) \qquad \mathfrak{R}_{\rho,\alpha}[\mathfrak{e}] := \min\left\{|\theta-\theta'| \;\middle|\; \int_H \sqrt{\mathrm{Tr}[\rho_\theta\, E(\eta)]\; \mathrm{Tr}[\rho_{\theta'}\, E(\eta)]}\, d\eta \leq \alpha^2\right\} \quad .$$

---

[16] Uffink & Hilgevoord, *op. cit.*

[17] C. Helstrom (1976): *Quantum Detection and Estimation Theory*; Academic, NY; Holevo, *op. cit.*.

[18] W. Wootters (1979): *Phys. Rev. D* **19**, p. 473

Then we redefine w_q as a measure of the distinguishability of the shifted states, a *translation width*:

$$(10c) \qquad w_{q,\alpha}[\rho] := \inf_{\epsilon}\left[\mathfrak{R}_{\rho,\alpha}[\epsilon]\right] \ .$$

This is consistent with (15), as it can be shown that

$$(21) \qquad w_{q,\alpha}[\rho] \geq \min\left\{|\theta| \;\middle|\; |\mathrm{Tr}[\rho\, S_q(\theta)]| \leq \alpha^2\right\} \ ,$$

with equality for pure states[19].

This new w is related to the ones of the previous sub-section:

$$(22a) \qquad \tilde{w}_{Q,\alpha}[\rho] \geq w_{q,\alpha}[\rho] \ ;$$

$$(22b) \qquad w_{q,\alpha}[|\phi\rangle\langle\phi|] = w^{\mathrm{F}}_{P,\alpha}[|\phi\rangle\langle\phi|] \ ; \ w_{q,\alpha}[\rho] \geq w^{\mathrm{F}}_{P}[\rho] \ .$$

For pure states these relations are trivial to derive [using (21)]. Relation (22a), perhaps despite appearances, is theory independent (like (13) and (14)), because it follows from the reasoning

$$(23) \qquad w_{q,\alpha}[\rho] = \inf_{\epsilon}\left[\mathfrak{R}_{\rho,\alpha}[\epsilon]\right] \leq \mathfrak{R}_{\rho,\alpha}[\mathrm{q}] = \tilde{w}_{Q,\alpha}[\rho] \ .$$

(Here q denotes the position PVM $\{|q\rangle\langle q|\mathrm{d}q\}_{\mathbb{R}}$.) Thus, we have (22b)⇒(16)⇒(11) and (22b)⇒(12). Relations of the type (22b) and (16) are, strictly speaking, *not* scatter relations because w is not a measure of statistical scatter. They may be called *shift-scatter* relations. They also present limits to preparation possibilities. Therefore, and because shift-scatter relations are conceptually so strongly related to scatter relations (see also above chain of derivations), we will also group them under the heading 'scatter principle'.

Shift-scatter relations can also be derived for cases where there is no PVM covariant with respect to the shifts involved. An example that immediately comes to mind is the time-energy case, where the Hamiltonian operator generates time-evolution

[19] J. Uffink & J. Hilgevoord (1988): *Physica B* **151**, p. 309

(time-shift) without there being a time-PVM. For that case an analog of the shift-scatter relation (22b) is straightforwardly derived, as it is quite similar to another, more usual, energy-time "uncertainty relation": the Mandel'shtam-Tamm relation[20]

$$(24) \qquad \tau_A^2 \, \langle \Delta^2 H \rangle \geq \tfrac{1}{4} \; ;$$

where H is the Hamiltonian operator and

$$\tau_A := \inf_t \left| \frac{\langle \Delta^2 A(t) \rangle^{\frac{1}{2}}}{\partial \langle A(t) \rangle / \partial t} \right| \quad , \quad A \text{ is any self-adjoint operator.}$$

Relation (24) can, of course, also be formulated for the position-momentum case using position-shifts instead of time-shifts[21]. In w we used the distinguishability measure $\mathfrak{K}$ between the e *probability distributions* of two shifted states as a starting point. In (24) we have taken the difference of the A *expectation values* in these two states as basic. Since (24) holds for all t and A,

$$(24') \qquad \tau^2 \, \langle \Delta^2 H \rangle \geq \tfrac{1}{4} \; ;$$

where

$$\tau := \inf_A \inf_t \left| \frac{\langle \Delta^2 A(t) \rangle^{\frac{1}{2}}}{\partial \langle A(t) \rangle / \partial t} \right| \; .$$

In case of a pure state $\rho = |\varphi\rangle\langle\varphi|$, optimal distinguishability is effected by measuring[22] $A = |\varphi\rangle\langle\varphi|$. In that case, as is not difficult to show,

$$(25) \qquad \tau = \tau_{|\varphi\rangle\langle\varphi|} = \inf_t \left| \frac{\partial \arcsin(2\langle A(t) \rangle - 1)}{\partial t} \right|^{-1} \leq \tfrac{1}{2} \, w_{t,\alpha} / \arccos(\alpha^2)$$

(where $w_{t,\alpha}$ is defined analogously to $w_{q,\alpha}$ in (15)).

---

[20]L. Mandel'shtam & I. Tamm (1945): *Bull. Acad. Sci. USSR, Phys. Ser.* **9**, p. 249; Holevo, *op. cit.*.

[21]Holevo, *op. cit.*, p. 105; J. Hilgevoord & J. Uffink (1989): in *Sixty-two years of Uncertainty: Historical Philosophical and Physical Inquiries into the Foundations of Quantum Mechanics* (ed. by A. Miller, Plenum, NY), p. 121

[22]Uffink & Hilgevoord, *op. cit.*

2 THE INACCURACY PRINCIPLE

As we saw, one of the possible forms of the UP is associated with the joint measurability of incompatible observables. This issue has been much debated: because relation (2) is itself not directly applicable to this situation (as we have seen), the inaccuracy interpretation, although frequently suggested[23], is dubious. Discussion centers around two questions: are incompatible observables jointly measurable at all and, if so, is there a relation (typographically) similar to Robertson's relation (2) restricting the accuracy of joint measurements ? An answer to these questions is of some interest because in such measurements (if possible at all) one might hope to find typically quantum mechanical inaccuracies limiting the possibilities of (future) high-precision measurement devices. We shall first tackle the problem within the von Neumann-Dirac formalism. In accord with the conclusion of ch. I, we find that formalism to be inadequate, and we give a more general one, due to Davies and Ludwig, with which we can work (see app. A). A widely used approach to the joint measurement problem in this generalized frame is based on reproduction of expectation values. It is examined in § 2.1, and found unsatisfactory because of its very weak 'measurement' notion. We then proceed to present our alternative, and investigate its properties [§ 2.2 to § 2.4]. Joint measurement is also feasible in the latter approach, although limited by an inaccuracy inequality [§ 2.5]. Examples are given in § 2.6, after which the results are evaluated [§ 2.7].

If we study existing approaches to the joint measurement issue, we see that a major problem in the discussion is the divergence of opinions about the meaning of the words 'joint', 'measurement' and 'accuracy'. It is clear that the answers to the relevant questions depend to a large extent on the content of these words. The proposal by von Neumann for a joint measurement interpretation of the scatter principle, which we discussed in ch. I, illustrates this. In this proposal 'measurement' was

---

[23] J. Jauch (1968): *Foundations of Quantum Mechanics* (Addison-Wesley, Reading, Mass.), p. 162; M. Jammer (1974): *The Philosophy of Quantum Mechanics* (Wiley, NY), p. 81; D. Bohm (1951): *Quantum Theory* (Prentice-Hall, Englewood Cliffs (NJ)), p. 99; A. Boehm (1979): *Quantum Mechanics* (Springer, Berlin), p. 46; R. Feynman, R. Leighton & M. Sands (1965): *Feynman lectures on physics*, vol. 3 (Addison-Wesley, Reading, Mass.), p. I-11

reinterpreted unsatisfactorily. Another is due to Park & Margenau[24]. They present several procedures intended to measure position and momentum jointly. A typical example is the 'time of flight' method. Here we have a one-dimensional situation, and a particle that is (to begin with) in a state concentrated in a small region Δq. We let this particle evolve freely for a long time T. Then we measure its position q, and attribute to the particle a momentum $p = mq/T$. Intuitively it is clear that there is no "jointness" in this measurement in any physically meaningful sense. It works because the preparation procedure (state concentrated in small region + long lasting free evolution) creates a state ρ where the position distribution $\langle q|\rho|q\rangle$ is equal to the momentum distribution $\langle p|\rho|p\rangle$ (up to a scale transformation, and in arbitrarily good approximation)[25]. Therefore this algorithm is very restrictedly applicable. Moreover, due to the fact that the class of ρ's achievable by the 'time of flight' method is so narrow that within this class no variations in "correlation" occur, the "joint" distribution is completely determined by its marginals. The other procedures given by Park and Margenau have similar bad properties, so that their conclusions as regards the possibility of joint accurate measurement are unfounded[26].

Very strong demands on 'measurement' lead other authors[27] to take the other extreme point of view, to conclude that joint measurements of incompatible observables are entirely impossible. The von Neumann-Dirac formalism, as has been realized for some time, was devised to cope only with idealized measurement devices [ch. I]. Accordingly, the impossibility conclusion is unavoidable *within* the von Neumann-Dirac formalism. But, more generally, this conclusion can only be obtained by ignoring or underestimating the possibility of allowing non-ideality in the measurement (as the results in the following sections will show).

---

[24]J. Park & H. Margenau (1968): *Int. J. Theor. Phys.* **1**, p. 211; J. Park, W. Band & W. Yourgrau (1980): *Ann. der Phys.* **37**, p. 189

[25]Park, Band & Yourgrau, *op. cit.*

[26]The Park & Margenau proposal was already criticized by de Muynck *et al* [W. de Muynck, P. Janssen & A. Santman (1979): *Found. Phys.* **9**, p. 71].

[27]P. Suppes (1961): *Phil. of Sc.* **28**, p. 378; E. McMullin (1954): *The Principle of Uncertainty* (PhD Thesis, Catholic University of Louvain, Belgium), unpublished.

If one analyzes a concrete measurement setup quantum mechanically, one generally sees that the outcome probabilities are not produced by a PVM at all. The choice for self-adjoint operators (or PVMs) as representatives of "perfect" observables that is made in the von Neumann-Dirac formalism is also somewhat unsatisfactory, because it amounts to limiting oneself *by postulate* to a class of observables that are, in some "self-evident" sense, "optimal" [ch. I]. Therefore this formalism, although it still has some authority among physicists, is losing ground in its approach to measurement in favor of more modern developments, which generalize it. Within such a wider frame, it is possible to take a better approach, in which a class of "optimal" measurements is selected from the class of all measurements by using *operational* arguments. Then we can be sure that the limitation to "optimal" observables is possible without loss of generality.

The most elementary measurements are yes–no measurements, conventionally represented by projectors[28]. We first generalize the representation of the *determinative* aspect of these yes–no measurements to *effects*[29]. An effect is a self-adjoint operator M satisfying

(27) $\qquad 0 \leq M \leq 1 \ .$

The probability of obtaining the result 'yes' is then given by $\langle M \rangle = \mathrm{Tr}(\rho M)$. An effect corresponds to a generalized yes-no measurement: we no longer require $M^2 = M$. In this chapter we shall not go into the *preparative* aspects of measurement (i.e. into the question of what happens to the object system after the measurement) since these are not crucial to the definition of measurement as such [ch. I]. Hence issues concerning, e.g., 'measurement of the first kind' will not be treated here [but see ch. IV].

Busch and Lahti[30] base a joint measurement theory on the effect notion. They argue that the thesis that quantum mechanics admits arbitrarily accurate joint measurements of incompatible observables can be substantiated if only 'measurement' is given a

---

[28]Jauch, *op. cit.*

[29]K. Kraus (1983): *States, Effects and Operations* (Lecture Notes in Physics **190**, Springer, Berlin)

[30]P. Busch & P. Lahti (1984): *Phys. Rev. D* **29**, p. 1634

suitable content. The proposal depends on an interpretation of the natural partial ordering on the set of effects. If $M \leq N$, they claim this means that a procedure realizing the effect M "gives information" about N. A joint measurement of two effects M and N should therefore realize an effect $O \neq 0$, $O \leq M$, $O \leq N$. That this interpretation leads to difficulties[31] is clear from the following example:

> Let $M = \frac{1}{2}1$; $N = \frac{1}{2}1 + \frac{1}{2}O$; O an arbitrary effect. Then obviously $M \leq N$, so that a realization of M "gives information" about N. But the probability of a positive result of an M realization does not depend on the state of the object system, so that such a procedure gives no information at all, let alone information about N.

These difficulties surface in a description of joint measurement along the lines of Busch & Lahti. Consider the following case:

> We have two effects $M = aE + (1-a)\bar{E}$ and $N = bF + (1-b)\bar{F}$ $(0 \leq a \leq b \leq \frac{1}{2})$. Here $\{E,\bar{E}\}$ and $\{F,\bar{F}\}$ are two *non-degenerate* PVMs (i.e. PVMs consisting of one-dimensional projectors) on $\mathscr{H} = \mathbb{C}^2$. In their terminology M and N are *fuzzy* versions of $\{E,\bar{E}\}$ and $\{F,\bar{F}\}$, respectively [cf. § 2.2]. Then it can be seen that the greatest lower bound of M and N is equal to
>
> $$O = a1 + 2\frac{(b-a)(1-b-a)}{(1-2a)}\bar{E} \quad .$$
>
> Obviously $O \neq 0$ for arbitrary small $a,b > 0$, for arbitrarily small fuzziness. But O tells us no more about the N probabilities than did M (or $\{E,\bar{E}\}$ itself). Thus a measurement based on O can be bad, even though the fuzziness in M and N is very small.

Hence calling an M–realization an N–measurement if only $M \leq N$ seems to imply an unreasonably weak definition of 'measurement'.

---

[31] Cf. E. Davies & J. Lewis (1969): *Comm. Math. Phys.* **17**, p. 239

More general measurements involve scales consisting of more than two elements. Such a measurement is in the general framework[32] represented by an *effect-valued measure* (EVM; also known as a positive operator-valued measure) rather than by a PVM. Thus, an EVM m is (for a discrete outcome set K) a family of bounded linear operators $\{M_k\}_K$ satisfying the relations (cf. (I.4)):

(28a) $\quad \forall_{k\in K}\, M_k \geq 0$

(28b) $\quad \sum_{k\in K} M_k = 1$

If the object system is in a state represented by the density operator ρ, an m-measurement will yield outcome k with probability $\mathrm{Tr}(\rho\, M_k)$. Of course probabilities are positive [(29a)] and normalized [(28b)].

The use of this extra sophistication is necessitated by the fact that a restriction to yes-no observables is not possible without loss of generality, because the structure of the set of effects (~ *binary* EVMs) is much simpler than that of the EVMs [§ 2.2]. Textbook observables [ch. I], referred to as *simple observables* in the following, fit into the EVM-framework as special cases: any PVM is an EVM.

Define further:

DEFINITION 1 The EVMs $\mathrm{m} = \{M_k\}_K$ and $\mathrm{n} = \{N_l\}_L$ are *coexistent*

$$:=$$

There exists an EVM $\mathrm{o} = \{O_{kl}\}_{K\times L}$ such that

(30)
$$\begin{cases} \mathrm{o}^{(1)} = \mathrm{m} \\ \mathrm{o}^{(2)} = \mathrm{n} \end{cases};$$

[32]Helstrom, *op. cit.*; Holevo, *op. cit.*; E. Davies (1976): *Quantum Theory of Open Systems* (Academic, NY); E. Davies & J. Lewis, *op. cit.*; G. Ludwig (1983): *Foundations of Quantum Mechanics*, vol I (Springer, Berlin)

where the *marginal* EVMs are defined by:

$$(31) \qquad \begin{cases} o^{(1)} = \{\sum_{\ell \in L} O_{k\ell}\}_K \\ o^{(2)} = \{\sum_{k \in K} O_{k\ell}\}_L \end{cases} .$$

The EVM o in def. 1 is said to represent a *joint (ideal) measurement* of m and n .

Two observables are termed *compatible* whenever their EVMs are coexistent[33]. As is well-known, in quantum mechanics not all observables are compatible. Our definition of compatibility coincides with the usual one for simple observables, as the following theorem shows:

THEOREM 1 If two EVMs m and n (as in def. 1) satisfy $\forall_{k \in K, \ell \in L} [M_k, N_\ell]_- = 0$ they are coexistent. If they are PVMs the converse is also true.

The second part of th. 1 is well known[34]. The first part is proved by the explicit construction of the EVM $o = \{O_{k\ell}\}_{K \times L}$; $O_{k\ell} := M_k N_\ell$.
We end this section with an important theorem in the theory of EVMs, due to Naĭmark (for the proof of which see e.g. the book by Holevo[35]):

THEOREM 2 For every EVM $\{M_k\}_K$ on $\mathcal{H}$ there is a Hilbert space $\mathcal{H}'$, a density operator ρ' on $\mathcal{H}'$ and a PVM $\{E_k\}_K$ on $\mathcal{H} \otimes \mathcal{H}'$ such that:

$$(33) \qquad M_k = \mathrm{Tr}_{\mathcal{H}'} (\rho' E_k) .$$

[33]Holevo, *op. cit.*, § II.6

[34]Holevo, *op. cit.*, prop. II.6.1; W. de Muynck & J. van den Eijnde (1984): *Found. Phys.* **14**, p. 111

[35]Holevo, *op. cit.*, § II.5; Helstrom, *op. cit.*, § III.3. See also App. A.

2.1 Expectation–value Based Approaches

In the following we shall study the possibilities of joint measurement of incompatible simple observables. Therefore the condition (30) of correct marginals, which forbids such joint measurements, has to be dropped. Nevertheless, some connection between the marginals of the joint EVM and the PVMs to be measured jointly must exist, in order to warrant the interpretation of the measurement result as "inaccurate" results of simple observables. Perhaps the first such criterion to come to mind is *unbiasedness*. Consider the joint "inaccurate" measurement of the EVMs corresponding to the self-adjoint operators R and T. Then we weaken (30) into demanding that the marginals only yield the correct expectation value (instead of the correct probability distribution):

$$\text{(32)} \qquad \begin{cases} \sum_{k\in K} \sum_{\ell\in L} k\, O_{k\ell} = R \\ \\ \sum_{k\in K} \sum_{\ell\in L} \ell\, O_{k\ell} = T \end{cases} .$$

The expectation-value criterion *(32) has been widely used. In, e.g., de Muynck et al.*[36] it was required that the joint measurement procedure should satisfy (32) using the original outcome sets (i.e. spectra). Analogous requirements were urged by Busch[37] and by Schroeck[38]. In this section we shall discuss the results obtained within such an expectation-value based approach in more detail.

Naĭmark's theorem leads us to consider a model for "inaccurate" measurement consisting of two quantum systems, an object and an ancillary system. They correspond to the Hilbert spaces $\mathscr{H}$ and $\mathscr{H}'$, respectively. Ancilla observables are primed; composite system observables are denoted by double primes. Consider the situation where we want to measure an object observable represented by $R\otimes 1'$. We let ancilla

---

[36]de Muynck, Janssen & Santman (1979); W. de Muynck & J. Koelman (1983): *Phys. Lett. A* 98, p. 1; P. Kruszynski & W. de Muynck (1987): *J. Math. Phys.* 28, p. 1761

[37]P. Busch (1985): *Int. J. Theor. Phys.* 24, p. 63

[38]F. Schroeck (1982): *Found. Phys.* 12, p. 825

and object interact for a time τ. The joint state of object and ancilla (on $\mathcal{H}\otimes\mathcal{H}'$) just prior to the beginning of the measurement is given by:

$$\rho'' = \rho\otimes\rho' \quad . \tag{34}$$

Here ρ' is the initial state of the ancilla. After the interaction is completed an ancilla observable A'_R ("read-out observable") is measured[39], which in the Heisenberg picture can be written as:

$$A'_R(\tau) := R\otimes 1' + G''_R \quad . \tag{35}$$

The operator G''_R represents the deviations of the measurement results from the "true" values. It represents *noise*. At first sight, it seems natural to assume the operator G''_R to be an observable of the ancilla:

$$G''_R = 1\otimes G'_R \quad . \tag{36}$$

As said above, the A'_R–measurement is intended as an "inaccurate" measurement of R. Therefore we require, as in (32), unbiasedness:

$$\begin{aligned} &\forall_\rho \ \mathrm{Tr}[\rho\otimes\rho' \, A'_R(\tau)] = \mathrm{Tr}[\rho R] \\ &\qquad\qquad \Updownarrow \\ &\forall_\rho \ \mathrm{Tr}[\rho\otimes\rho' \, G''_R] = 0 \quad . \end{aligned} \tag{37}$$

This can, using (36), be rewritten as a condition for ρ' , viz.

$$\mathrm{Tr}(\rho' \, G'_R) = 0 \quad . \tag{38}$$

The outcome probability distribution of the measurement is given by

$$P_{A'_R(\tau)}(\mathrm{d}r') = \mathrm{Tr}[\rho\otimes\rho' \, E_{A'_R(\tau)}(\Delta r')] \quad , \tag{39}$$

$\{E_{A'_R(\tau)}(\mathrm{d}r')\}$ denoting the PVM of $A'_R(\tau)$. We can rewrite (39) by partial tracing over the ancillary variables, viz.

$$P_{A'_R(\tau)}(\mathrm{d}r') = \mathrm{Tr}_{\mathcal{H}}[\rho \, M_{A'_R(\tau)}(\mathrm{d}r')] \quad , \tag{39'}$$

[39] Alternatively, the model may involve a measuring device interacting with both ancilla and object.

with:

(40) $$M_{A'_{\mathbb{R}}(\tau)}(\Delta r') = \mathrm{Tr}_{\mathscr{H}'}[\rho' E_{A'_{\mathbb{R}}(\tau)}(\Delta r')] \quad .$$

The EVM $\{M_{A'_{\mathbb{R}}(\tau)}(\mathrm{d}r')\}$ forms a representation, alternative to $A'_R(\tau)$, of this noisy measurement [cf. th. 2].

From (37) it follows directly[40] that the "inaccuracy" of the measurement leads to excess scatter in measurement results, as is evidenced by an increased variance:

(41) $$\forall_\rho \quad \langle \Delta^2 A'_R(\tau) \rangle = \langle \Delta^2 R \rangle + \langle \Delta^2 G''_R \rangle \quad .$$

This accounts for an excess spreading obtained already by Arthurs and Kelly[41] in a model joint measurement of position and momentum.

Moreover, (35) and (36) imply that the final A'_R outcome distribution can be derived from the R outcome distribution by *convolution* with the G'_R–distribution:

(42) $$\forall_\rho \quad P_{A'_{\mathbb{R}}(\tau)}(\Delta r') = \int_{-\infty}^{\infty} P_R(\mathrm{d}r)\, g(\Delta r', r)$$

with:

(43) $$g(\mathrm{d}r', r) = \tilde{g}(r' - r)\mathrm{d}r' = p_{G'_{\mathbb{R}}}(r' - r)\, \mathrm{d}r'$$

(P's denote probabilities, p's probability densities.) Using (40), we may equivalently write

(42′) $$M_{A'_{\mathbb{R}}(\tau)}(\Delta r') = \int_{-\infty}^{\infty} E_R(\mathrm{d}r)\, g(\Delta r', r) \quad .$$

---

[40] de Muynck & Koelman, *op. cit.*; Kruszynski & de Muynck, *op. cit.*.

[41] E. Arthurs & J. Kelly (1965): *Bell Syst. Techn. J.* **44**, p. 725; C. She & H. Heffner (1966): *Phys. Rev.* **152**, p. 1103

The function g in (42) and (42′) characterizes the "smearing" in the final A'_R-distribution with respect to R. The function g is the G'_R probability distribution, and therefore a property of the measuring device (i.e. the ancilla) *alone*. The function g depends only on ρ', *not* on the initial object state ρ. Thus condition (36) defines *object-independent* noise.

The approach (35)–(37) can be used as a starting point for a joint measurement of two observables[42] represented by the non-commuting Hermitian operators R and T. The aim is then to find two noise operators G'_R and G'_T on $\mathscr{H}'$ such that $A'_R(\tau) = R + G''_R$ and $A'_T(\tau) = T + G''_T$ commute. Next, a non-trivial lower bound to the amounts of excess noise, as defined by (41), is to be derived. But this approach faces one major handicap[43]. Since

$$(44) \qquad [A'_R(\tau)\,,\, A'_T(\tau)]_- = 0$$

$$\Updownarrow$$

$$[R\,,T]_- \otimes 1' + 1 \otimes [G'_R\,,\, G'_T]_- = 0\ ,$$

relation (44) can only be valid if $[R\,,\,T]_- = ic1$, c a constant. This is only the case for position-momentum like pairs. For such pairs[44], as follows immediately from equations (44), (41) and the Heisenberg inequality (1) for the noise operators, the excess variances are bounded by

$$(45) \qquad \langle \Delta^2 G''_R \rangle \, \langle \Delta^2 G''_T \rangle \geq \tfrac{1}{4} c^2\ .$$

---

[42]C. Helstrom (1974): *Found. Phys.* **4**, p. 453; S. Personick (1971): *Bell Syst. Techn. J.* **50**, p. 213; also E. Davies (1970): *J. Funct. Anal.* **6**, p. 318 and Helstrom, *op. cit.* (1976); Davies, *op. cit.* (1976); Holevo, *op. cit.*.

[43]Cf. H. Yuen (1982): *Phys. Lett. A* **91**, p. 101

[44]Helstrom, *op. cit.* (1976); S. Personick, *op. cit.*; Holevo, *op. cit.* (1982); A. Holevo (1986): *Theor. & Math. Phys.* **65**, p. 1250

If we want to generalize (45), we must first find a less restrictive characterization of 'noisy' measurements. The set of conditions *additivity* (35), object-independence (36) and unbiasedness (37), which formed that characterization, has to be weakened. We saw above that there exist two ways of viewing the noise in the formalism treated thus far. Considering object-independence (36) as primary, the smearing (42) described the noise. On the other hand, from the point of view of unbiasedness (37), the excess noise (41) formed such a description. Thus, depending on one's attitude towards 'measurement', different possibilities of generalizing (45). If a measurement is characterized by its expectation value, indeed often an important result, (36) should be dropped as a general requirement. Arthurs & Goodman[45], working along these lines, generalized (45) to

$$\langle \Delta^2 G_R'' \rangle \, \langle \Delta^2 G_T'' \rangle \geq \tfrac{1}{4} |\langle [R,T]_- \rangle|^2 \quad . \tag{46}$$

Yuen[46] also intends to generalize inequality (45) using an expectation-value based approach. He starts from a particular joint measurement POVM for position and momentum, viz.

$$M_{P''\&Q''}(\mathrm{d}p'',\mathrm{d}q'') = |\alpha\rangle\langle\alpha| \frac{\mathrm{d}^2\alpha}{\pi} \quad ; \quad \alpha = \tfrac{1}{2}\sqrt{2}(q'' + \mathrm{i}p'') \quad . \tag{47}$$

Here $|\alpha\rangle$ are the harmonic oscillator coherent states. They are the eigenstates of the (non-Hermitian) annihilation operator $a = \frac{1}{2}\sqrt{2}(Q + \mathrm{i}P)$. As indicated by (47), these states satisfy a closure relation. But they are not orthogonal. The set of states $\{|\alpha\rangle\}$ is *overcomplete*: every vector can be expanded in $|\alpha\rangle$'s in many ways. The non-negative definite and normalized distribution $\langle\alpha|\rho|\alpha\rangle/\pi$ is called the *Husimi* or Q-distribution[47]. The EVM (47) that generates the Husimi distribution results from the joint PVM of $P'' = P + P' = A'_P(\tau)$ and $Q'' = Q - Q' = A'_Q(\tau)$ on $\mathcal{H}\otimes\mathcal{H}'$ by taking $\rho' = |0\rangle\langle 0|$ (the harmonic oscillator ground state).

---

[45]E. Arthurs & M. Goodman (1988): *Phys. Rev. Lett.* **60**, p. 2447

[46]Yuen, *op. cit.*

[47]K. Husimi (1940): *Proc. Phys. Math. Soc. Japan* **22**, p. 264; M. Hillery, R. O'Connell, M. Scully & E. Wigner (1984): *Phys. Rep.* **106**, p. 123

Yuen then proposes to generalize this by looking for a POVM generated by the eigenstates of the operator $B = R + iT$. But note that the joint PVM of Q'' and P'' on $\mathcal{H} \otimes \mathcal{H}'$ has *two* arguments, reflecting the fact that in a *joint* measurement every measurement outcome consists of two numbers. Consequently, if the POVM on $\mathcal{H}$ is to be derived from a set of states, this set must be *bivariate*. It is unclear to what extent the operator B can have a bivariate set of eigenstates if R and T are not position and momentum operators. This makes it doubtful whether the generalization Yuen achieves in this way is indeed substantial.

But even in the Arthurs & Goodman version, the expectation-value based approach has a number of serious drawbacks. Noise processes will in general violate eq. (37). If they do, it may appear that this can be remedied by simply subtracting the bias from the measurement result. But by dropping (36), the noise operators are allowed to act on the full product Hilbert space. Therefore the bias is generally not a constant: it may differ for different initial object states ρ. Then bias subtraction is impossible. Thus, from a physical point of view, unbiasedness constitutes a substantial idealization.

Moreover, also the excess variance $\langle \Delta^2 G''_R \rangle$ and, more generally, the connection between R and $A'_R(\tau)$ probability distributions, will now in general depend on ρ. Consider as an example a two-dimensional particle, with position operators Q_1 and Q_2. In order to measure Q_1, we couple it with an ancilla such that

$$A'_{Q_1}(\tau) = Q_1 \otimes 1' + Q_2 \otimes Y' \quad ,$$

for some ancilla operator Y'. The ancilla state ρ' is such that $\mathrm{Tr}(\rho' Y') = 0$. Then the measurement is indeed unbiased with respect to Q_1. But it is clear that the difference between measurement outcome and true value, the noise, will also depend on ρ through Q_2's presence in $Q'_1(\tau)$: $G''_{Q_1} = Q_2 \otimes Y'$. As the example shows, the excess variance $\langle \Delta^2 G'' \rangle$ is an inaccuracy characterization that, contrary to common practice, cannot be seen as a property of the measurement device. In the excess variance bound (46), the fact that the noise now *in general* depends on ρ, results in a bound that depends on ρ, too (unless of course $[R, T]_- \sim 1$; cf. (45)).

Consider as an example illustrating these approaches the spin ½ case of σ_1 and σ_2. Yuen's method is not applicable here: $\sigma_1 + i\sigma_2$ has only one eigenstate, instead of the four we need. The unbiasedness condition (32) can therefore certainly not be met. Arthurs & Goodman's approach is applicable, and leads to the inequality

$$(48) \qquad \left(\langle \Delta^2 A'_{\sigma_1}(\tau)\rangle - \langle \Delta^2 \sigma_1\rangle\right)\left(\langle \Delta^2 A'_{\sigma_2}(\tau)\rangle - \langle \Delta^2 \sigma_2\rangle\right) \geq \tfrac{1}{4}|\langle \sigma_3\rangle|^2 \quad .$$

As a consequence of the possible dependence of the noise on the initial object state ρ, the bound in above inequality also depends on ρ: for certain object states it can even be zero. Therefore the above inequality cannot be interpreted as a (non-trivial) limit to the accuracy with which σ_1 and σ_2 are jointly measurable.

Thus the notion of inaccuracy as purely a property of the measuring instrument has to be abandoned on behalf of general applicability within an expectation-value based approach. Its reliance on the labeling of the measurement scale forms a second drawback. The outcome set of the POVM (~ spectrum of the self-adjoint operator on the product Hilbert space), which contains this labeling, is merely a matter of convenience[48], however. Nothing physical in the measurement device is changed if we alter its scale. Therefore an acceptable 'inaccuracy' notion should be insensitive to the labeling of the observable measured. A third disadvantage of this approach is that the expectation value is not always the only parameter of a probability distribution we are interested in. In dropping (36) we, however, in general cut ourselves off from obtaining more information about the R–distribution. For this reason we now want to consider a different way of generalizing the notion of inaccurate measurement. First note that relation (42) is invertible (for suitable G'_R–distribution). We may write this according to[i]

$$(49) \qquad \forall_\rho \; P_R(\Delta r) = \int_{-\infty}^{\infty} P_{A'_R(\tau)}(dr')\, f(\Delta r, r') \quad .$$

The "function" $f(dr,r')$ in (49) is (unlike $g(dr,r')$ in (42)) not necessarily non-negative definite. Since the results of an inaccurate measurement are in themselves not interesting, but are used to make deductions about what a measurement of the

[48] Ludwig, *op. cit.*, p. 97

desired observable would have given, this invertibility is a very appealing property. If the full invertibility (49) is indeed feasible, not only the expectation value, but *all* moments of the R–distribution can be derived from the $A'_R(\tau)$–measurement. This holds true *independent of* (37). Therefore we can interpret $A'_R(\tau)$ results in terms of R using (42), without referring to the unbiasedness condition (37) at all. If (42) is not invertible, it is not possible to derive from the $A'_R(\tau)$ results the expectation values $\langle f(R) \rangle$ for all functions f. Nevertheless (42) guarantees that it is possible for some functions [see further § 2.3].

Since (36) implies (42), there are better arguments for holding on to (36) then there are for upholding (37). In other words, an approach based on (36) rather than on (37) and (35) looks most promising from the point of view of characterizing measurement by probability distributions rather than by expectation values. 'Inaccuracy' would be seen as a definite relation between the *distribution* of the *realized* observable and that of the *desired* observable. This relation should be, like the function g in (42), purely a property of the measuring device, and the 'inaccuracy' concept should not involve the labeling of the observables.

2.2 Non–ideality: Definitions & Properties

A study of 'inaccuracy' is not only relevant to the joint measurement problem. As noted above, within the general framework of EVMs certain observables may be defined as "optimal" through operational arguments. An 'inaccuracy' notion could provide such an argument. It would allow to distinguish between "optimal" observables and "bad" observables that mix information about the object system with non–information coming from the measurement device. One may think of[49] "randomization" or noise (as in the previous section) affecting our meter. If this noise

[49]Holevo, *op. cit.* (1982) p. 19

totally dominates our device's operation, it measures an *uninformative observable*. Such an observable is represented by an EVM $\{M_k\}_K$ of the form:

$$(50) \qquad \forall_{k\in K}\ M_k = f_k 1 \ ,\ \forall_{k\in K}\ f_k \geq 0 \ ,\ \sum_{k\in K} f_k = 1$$

(f_k a scalar). The outcome probability distribution of such an uninformative observable does not depend on the state of the object system at all. 'Inaccuracy' may provide a structure in the class of observables that gives us a subclass maximally free from this non-information, observables whose outcomes are maximally associated with the object system alone. We could indeed call such observables 'optimal'. If it were known what observables are optimal, we could restrict ourselves to this class, since the properties of all measurements are characterized by those of only the optimal ones *via* the structure.

One proposal for optimal observables is based on the structure on the class of EVMs that has been most systematically investigated so far: *convexity*. If two EVMs $\{M_k\}_K$ and $\{N_k\}_K$ with the same outcome set K are given, the set of operators $\{O_k\}_K$ defined by:

$$(51) \qquad O_k := \lambda M_k + (1-\lambda) N_k \ ;$$

is also an EVM for all $\lambda \in (0,1)$. $\{O_k\}_K$ corresponds physically to a situation where we let chance decide wether we use an $\{M_k\}_K$ or an $\{N_k\}_K$ device (with probabilities λ and $(1-\lambda)$, respectively). The set of EVMs is convex, like the set of states [ch. I]. EVMs $\{O_k\}_K$ that cannot be decomposed into other EVMs as in (51), are called *extreme*. The observables they correspond to will be called *pure*. This class contains all simple observables. But, unless the outcome set K consists of only two elements (yes-no observables), there exist non-simple pure observables[50]. It has been suggested[51] that pure observables constitute the subclass of optimal observables referred to above.

Relation (51), however, does not imply that the EVMs on its right-hand side are related: they may for instance be position and momentum. In such a case it would be

[50]Holevo, *op. cit.* (1982) p. 30; Ludwig, *op. cit.*, p. 138

[51]E.g. by Ludwig, *op. cit.*, p. 135.

hardly appropriate to call either $\{M_k\}_K$ or $\{N_k\}_K$ "better" than $\{O_k\}_K$, since the latter's outcomes give information provided by those of neither of the former. Therefore an EVM that is non-extreme is not necessarily non-optimal. Convexity seems to be less suitable for the task that is to be performed here than an "inaccuracy" notion would be: if one cannot speak of "better", one certainly cannot speak of "optimal". Another disadvantage of the convexity structure is its inability to relate EVMs in a label-independent way. As a consequence there appears to be no natural way to connect two EVMs with different outcome sets *via* convexity.

In this section we shall study an inaccuracy notion[ii], which can be seen as an elaboration of (42). We shall work on a finite dimensional complex Hilbert space $\mathscr{H}$, and ignore superselection rules. In this way we can get an idea of the properties of the structure such a concept generates without having to deal with too many mathematical technicalities. Moreover, many infinite dimensional situations can be seen as limiting cases of finite dimensional ones. We define for two EVMs $\mathfrak{m} = \{M_k\}_K$ and $\mathfrak{n} = \{N_\ell\}_L$ the following relation:

DEFINITION 2 $$\mathfrak{n} \longrightarrow \mathfrak{m} \;:=\; \exists_{\{\lambda_{k\ell}\}\in\mathbb{R}^{K\times L}} \begin{cases} \sum_{k\in K} \lambda_{k\ell} = 1 \\ \lambda_{k\ell} \geq 0 \\ M_k = \sum_{\ell\in L} \lambda_{k\ell} N_\ell \end{cases} .$$

The matrix $\{\lambda_{k\ell}\}$ is a stochastic matrix[52]. It is a property of an $\mathfrak{m}$–device (*not* of the object, since it has no relation to the density operator), characterizing its relation to the observable corresponding to $\mathfrak{n}$, in as far as determinative aspects are concerned.

A more restrictive version of the relation $\longrightarrow$ is defined (for two EVMs $\mathfrak{m}$ and $\mathfrak{n}$ as above) by:

DEFINITION 3 $$\mathfrak{n} \xrightarrow{i} \mathfrak{m} \;:=\; \mathfrak{n} \longrightarrow \mathfrak{m} \wedge \exists_{\{\mu_{\ell k}\}\in\mathbb{R}^{L\times K}} N_\ell = \sum_{k\in K} \mu_{\ell k} M_k \quad .$$

[52] J. Ortega (1987): *Matrix Theory* (Plenum, NY); App. A

The matrix summations in both def. 2 and def. 3, as well as later series, are required to be elementwise absolutely convergent. This is necessary because the order of summation should not be relevant for the result. If the summation concerns only positive elements (such as in def. 2, but *not* in def. 3) absolute convergence follows from convergence. In def. 2 the EVM m is to represent a "smeared" (*non-ideal*) version of n. If it is the case that $n \xrightarrow{i} m$, the "smearing" can be undone in a certain sense. We will go into these physical aspects of the structure more closely in § 2.3.

Define the equivalence relation:

DEFINITION 4 $\quad m \longleftrightarrow n \;:=\; m \longrightarrow n \;\wedge\; n \longrightarrow m \;.$

If we define $\xleftrightarrow{i}$ analogously, it is trivial to show that:

THEOREM 3 $\quad m \longleftrightarrow n \;\Leftrightarrow\; n \xleftrightarrow{i} m \;.$

Thus both $\longrightarrow$ and $\xrightarrow{i}$ define a partial order relation between the equivalence classes defined by $\longleftrightarrow$. In a partial order structure it is natural to define:

DEFINITION 5 $\quad m \text{ is } \textit{maximal} \;:=\; \forall_{\text{EVM } n} \left[n \longrightarrow m \;\Rightarrow\; n \longleftrightarrow m \right] \;.$

DEFINITION 6 $\quad m \text{ is } \textit{minimal} \;:=\; \forall_{\text{EVM } n} \left[m \longrightarrow n \;\Rightarrow\; n \longleftrightarrow m \right] \;.$

Using $\xrightarrow{i}$ we can define *i-maximality* and *i-minimality* in a similar way.

We further introduce the following notations:

(52) $\quad \boldsymbol{A} \sim \boldsymbol{B} \;:=\; \exists_{c \in \mathbb{R} \setminus \{0\}} \; \boldsymbol{A} = c\boldsymbol{B} \;;$

(53) $\quad L(m) \;:=\; \left\{ \boldsymbol{X} \;\middle|\; \exists_{(\alpha_k) \in \mathbb{R}^K} \; \boldsymbol{X} = \sum_{k \in K} \alpha_k \boldsymbol{M}_k \right\}$ [53];

(54) $\quad K(m) \;:=\; \left\{ \boldsymbol{X} \;\middle|\; \exists_{(\alpha_k) \in \mathbb{R}^K;\, \alpha_k \geq 0} \; \boldsymbol{X} = \sum_{k \in K} \alpha_k \boldsymbol{M}_k \right\} \;.$

---

[53] Cf. Davies, *op. cit.* (1970) and F. Schroeck (1989): *Int. J. Theor. Phys.* 28, p. 247

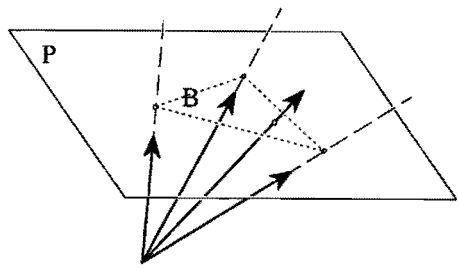


fig. 1 *Cone generated by four vectors. Three generate extremal rays (dashed). When the cone is intersected with a plane* P *(not going through the origin), the base* B *(dotted) results.*

It can be verified that both L and the *cone* K are closed.
Obviously, if $n \longrightarrow m$ then $K(m) \subseteq K(n)$. If $n \xrightarrow{i} m$ we have in addition to this: $L(n) = L(m)$.

(55) $B(m) := \{X \in K(m) \mid \mathrm{Tr}(X) = 1\}$.

The functional $f(X) = \mathrm{Tr}(X)$ is a *strictly monotonic* linear functional on the cone K (:= $f(X) > 0$ for all non-zero X in K). Hence B is a *base*[54] for the cone K (:= there is an $\alpha > 0$ such that $\alpha X \in B$ for all non-zero X in K) [cf. fig. 1].

(56) $K_{max}(m) := L(m) \cap \mathscr{B}^+$.

The set $\mathscr{B}$ is the set of all bounded operators on $\mathscr{H}$, $\mathscr{B}^+$ consists of all bounded positive operators on $\mathscr{H}$.

(57) $B_{max}(m) := \{X \in K_{max}(m) \mid \mathrm{Tr}(X) = 1\}$.

B_{max} is the base of K_{max} .

[54] G. Jameson (1970): *Ordered Linear Spaces* (Lecture Notes in Mathematics **141**, Springer, Berlin)

The extreme elements of a convex set C are denoted by $\partial_E C$; the elements of the *extremal rays*[55] of a cone K are denoted by ∂K:

(58) $$\partial K := \{X \in K \mid \forall_{Y \in K}\ X - Y \in K \Rightarrow Y \sim X\} \ .$$

Also useful are:

DEFINITION 7 The EVM m is *pairwise linearly independent*

$$:=$$

$$\forall_{k,\ell \in K} \left(M_k \sim M_\ell \Rightarrow k = \ell \right) \ .$$

In a pairwise linearly independent EVM no two non-zero elements lie on the same ray. The support of a pairwise linearly independent EVM is not unnecessarily large.

DEFINITION 8 The EVM $m = \{M_k\}_{k \in K}$ is *self-extremal* $:= \forall_{k \in K}\ M_k \in \partial K(m)$.

In a self-extremal EVM no element can be written as a non-trivial convex sum of other elements.

Our main results regarding structural properties are (the EVMs n and m as above; proofs can be found in the appendix):

THEOREM 4 Any two equivalent EVMs $\{M_k\}_K$ and $\{N_\ell\}_L$ satisfy:

$$\forall_{k \in K}\ \sum_{k'} | M_{k'} \sim M_k \ M_{k'} = \sum_{\ell'} | N_{\ell'} \sim M_k \ N_{\ell'}$$

This theorem characterizes the content of the equivalence classes. Maximal EVMs are characterized by:

THEOREM 5 m is maximal $\Leftrightarrow \forall_{k \in K}\ M_k \in \partial \mathcal{B}^+$.

---

[55] Jameson, *op. cit.*

The set $\partial\mathscr{L}^+$, the extremal rays of $\mathscr{L}^+$, consists precisely of operators that are up to a scalar factor one-dimensional projectors. Note that th. 5 implies that there is more than one equivalence class of maximal EVMs. It also implies that our definition reduces to the usual one (= non-degeneracy) for PVMs.

THEOREM 6 $\quad \mathsf{m}$ is i–maximal $\Leftrightarrow \forall_{k\in K}\ \boldsymbol{M}_k \in \partial K_{max}(\mathsf{m})$.

Thus all i–maximal EVMs are self-extremal (def. 8).

THEOREM 7 $\quad \mathsf{m}$ is minimal $\Leftrightarrow \forall_{k\in K}\ \boldsymbol{M}_k \sim \mathbf{1} \Leftrightarrow \mathsf{m} \longleftrightarrow \mathsf{i}$.

Hence uninformative observables are represented precisely by minimal EVMs. The trivial EVM $\{1\}$ is denoted by i.

Straightforward is (def. 3):

THEOREM 8
- m is maximal $\Rightarrow$ m is i–maximal ;
- m is minimal $\Rightarrow$ m is i–minimal .

Less obvious are perhaps:

THEOREM 9 If $\dim(\mathscr{H}) = 2$, an EVM is i–maximal iff it is either minimal or maximal. This is not true if $\dim(\mathscr{H}) > 2$;

THEOREM 10
- m is i–minimal $\Leftrightarrow$ m is minimal ;
- m is i–minimal $\Rightarrow$ m is i–maximal .

The structure is closed in the sense that:

THEOREM 11
- For every EVM m there is a maximal EVM $\mathsf{n} \rightarrow \mathsf{m}$;
- For every EVM m there is an i–maximal EVM $\mathsf{n} \xrightarrow{\mathsf{i}} \mathsf{m}$;
- For every EVM m there is a minimal EVM $\mathsf{n} \leftarrow \mathsf{m}$;
- This last assertion is not true for i–minimality and $\xleftarrow{\mathsf{i}}$.

The connection between the structure induced by $\rightarrow$ and the convexity structure is indicated by the following theorems:

THEOREM 12 If the EVM $\mathfrak{m}$ is extreme, then it is i–maximal and pairwise linearly independent and $B(\mathfrak{m})$ is a simplex. The converse is true only if $\dim(\mathscr{H}) = 2$.

THEOREM 13 If an EVM $\mathfrak{m}$ is maximal or minimal and it is pairwise linearly independent and $B(\mathfrak{m})$ is a simplex, then it is extreme. The converse is true only if $\dim(\mathscr{H}) = 2$.

Since PVMs are always extreme[56], PVMs are i–maximal and pairwise linearly independent. A maximal PVM is the PVM associated with a complete orthonormal basis. We can therefore conclude that a maximal EVM on $\mathbb{C}^n$ must have at least n elements.

In special cases EVMs can be related to PVMs via the $\rightarrow$ structure[57]:

THEOREM 14

- For every binary EVM $\mathfrak{m} = \{M, \bar{M}\}$ there is a PVM $\mathfrak{e}$ such that $\mathfrak{e} \rightarrow \mathfrak{m}$;
- Suppose $\mathfrak{m} = \{M_k\}_{k\in K}$ is an EVM. Then:
  $\forall_{k,\ell\in K}\ [M_k, M_\ell]_- = 0 \ \Leftrightarrow\ \exists_{\text{PVM } \mathfrak{e}}\ \mathfrak{e} \rightarrow \mathfrak{m}$.

This theorem is easily proven (the first assertion follows from the fact that $[M, \bar{M}]_- = [M, 1\text{–}M]_- = 0$).

Our notion $\rightarrow$ can also be used to give an alternative definition of coexistence: we have

THEOREM 15 EVMs $\mathfrak{m}$ and $\mathfrak{n}$ are coexistent $\Leftrightarrow\ \exists_{\text{EVM } \mathfrak{o}}\ \mathfrak{o} \rightarrow \mathfrak{m} \wedge \mathfrak{o} \rightarrow \mathfrak{n}$.

Proof of the theorem follows straightforwardly from def. 1 and def. 2.

---

[56]Holevo, *op. cit.* (1982) p. 30; Ludwig, *op. cit.*, p. 138

[57]A. Holevo (1972): *Trans. Moscow Math. Soc.* **26**, p. 133; Helstrom, *op. cit.* (1976) p. 87

The importance of this theorem is that it shows that the typical feature of quantum measurements, viz. incompatibility, can be defined completely in terms of the relation $\rightarrow$. The fact that there are incompatible observables in quantum mechanics can thus be seen immediately from the fact that there is no unique equivalence class of maximal observables (th. 5). For the classical case the situation is quite different. A classical analog[58] [ch. I] of $\mathbb{C}^n$ is a finite phase space $\Omega := \{\omega_1, \ldots, \omega_n\}$. The states are given by probability distributions $(p_\omega)_\Omega$; $\forall_{\omega\in\Omega}\, p_\omega \geq 0$; $\sum_{\omega\in\Omega} p_\omega = 1$. An observable is represented in such a classical model by a set $\{f_k(\omega)\}_K$ of functions (a *positive function valued measure*, PFVM) that satisfies:

$$(59) \qquad \forall_{k\in K,\omega\in\Omega}\, f_k(\omega) \geq 0 \;\; ; \;\; \forall_{\omega\in\Omega} \sum_{k\in K} f_k(\omega) = 1 \;\; .$$

The probability of outcome k is given by $\sum_{\omega\in\Omega} p_\omega f_k(\omega)$.

The most characteristic difference with the quantum case is that here the equivalence class of maximal PFVMs is *unique*, reflecting to the compatibility of classical observables (th. 15). All maximal PFVMs are equivalent to the PFVM $\{g_k(\omega)\}_K$; $g_k(\omega) := \delta_{k\omega}$; $K = \Omega$. This is in agreement with the reasoning of ch. I: it reflects the fact that in classical models any measurement can be related to a determination of a property of the object system. The ultimate property ($\sim$ maximal PFVM) is the system's C–state ω.

The structure induced by $\xrightarrow{i}$ for the classical case is simpler in similar ways. This is a consequence of the fact that the set of PFVMs on Ω is isomorphic with respect to $\xrightarrow{i}$ and $\rightarrow$ (and convexity) to the set of EVMs $m \leftarrow n$; $n = \{N_{k_1}, \ldots, N_{k_n}\}$ a fixed EVM with $\forall_k \, \|N_k\| = 1$ (e.g. a PVM; cf. the second part of th. 10).

We can therefore conclude that some of the intricacies the structure has in quantum theory are indeed unclassical, much as this is the case for the convexity structure[59] [ch. I]. Moreover, the above suggests that the properties of the $\rightarrow$ structure on the class of observables for a given statistical theory can be used to characterize the

[58]Holevo, *op. cit.* (1982)

[59]Holevo, *op. cit.* (1982)

theory *in abstracto*, much as such an abstract characterization is the aim of quantum logic. Further investigation of the properties of such $\longrightarrow$ structures would therefore be interesting.

2.3 Non-ideality: Physical Interpretation

As said above, $\longrightarrow$ is intended as a concretization of "inaccuracy". We shall in the following reserve the term 'non-ideal' to refer to $\longrightarrow$:

if $\mathfrak{n} \longrightarrow \mathfrak{m}$ we say that $\mathfrak{m}$ is associated with a *non-ideal* $\mathfrak{n}$-meas*urement*;

if $\mathfrak{n} \xrightarrow{i} \mathfrak{m}$ we say that $\mathfrak{m}$ is associated with an *invertibly* non-ideal $\mathfrak{n}$-measurement

($\mathfrak{n}$ and $\mathfrak{m}$ two EVMs). A meter realizing $\mathfrak{m}$ is called a "non-ideal $\mathfrak{n}$-meter". In order to clarify the physical meaning of $\longrightarrow$, on order to justify this nomenclature, we first consider a possible description on a product Hilbert space [cf. th. 2 and § 2.1]:

THEOREM 16 Suppose $\mathfrak{a} = \{M_a\}_{\sigma(A)}$ is given as the PVM of $A = \sum_{a\in\sigma(A)} aM_a$ on $\mathscr{H}$. Here $\sigma(A)$ denotes the spectrum of A. Let $\{E_z\}_{\sigma(Z)}$ on $\mathscr{H}\otimes\mathscr{H}'$ be the PVM of $Z = \sum_{z\in\sigma(Z)} zE_z$. If Z can be written as a function $z(A,T)$ of A and a self-adjoint operator T on $\mathscr{H}'$, then the EVM $\mathfrak{o} = \{O_z\}_{\sigma(Z)}$ defined by:

(60) $$O_z := \mathrm{Tr}_{\mathscr{H}'}(\rho' E_z)$$

has the property $\mathfrak{a} \longrightarrow \mathfrak{o}$ for every density operator ρ' on $\mathscr{H}'$.

Proof of this theorem is straightforward, using the spectral theorem on $\mathscr{H}'$.

Since for certain cases (e.g. finite outcome sets) the converse of th. 16 can also be proved, the definition of "noisy" measurement as $\longrightarrow$ (def. 2) boils down to generalizing the additive object independent noise of (35)–(36) not to additive object-

dependent noise, as Arthurs & Goodman do, but to *object independent non-additive* noise. It is tantamount to generalizing (35) to

$$(61) \qquad A'_R(\tau) = h(R\otimes 1', G''_R)$$

for an arbitrary function h, and keeping (36).

More precisely, $\rightarrow$ can for certain cases be shown to imply that there is a Hilbert space $\mathscr{H}'$ with state ρ' and noise operator G''_R such that (61) generates the non-ideal EVM. But this is a purely mathematical construction, and neither ρ' nor G''_R need to match the actual physics of the meter. Conversely, the mathematical description of the device actually realized need not follow eq. (61). Since it is also usually not very convenient to use a noise operator explicitly, we shall work with the matrix $\{\lambda_{k\ell}\}$ or, more generally, with an analog of this for the continuous case (e.g. the function $g(\Delta r', r)$ in (40)). From that point of view, non-ideality relates two measurement procedures (observables): it is a *relative* inaccuracy notion rather than an absolute one. It depends in no way on the state of the object (c.q. the preparation). The matrix $\{\lambda_{k\ell}\}$ in def. 2 is a property of the m–meter. In above representation of a non-ideal measurement, however, the amount of "smearing" is related to the amount of noise in the ancillary system, to the scatter associated with the auxiliary Hilbert space. But this scatter is not scatter in the object system. Moreover, such a connection of smearing with scatter is only available within this particular type of model for the measurement process. Therefore our earlier conclusion [ch. I] that "reinterpretation of (the symbols in) a scatter relation" cannot widen the relevance of scatter relations to incorporate inaccuracy limits, is not jeopardized.

The matrix $\{\lambda_{k\ell}\}$ in def. 2 characterizes the inaccuracy of the m–meter, when seen as an n–meter, in a mathematically precise way. This characterization allows us to deduce from the m results certain information about what the n results would have been like. This information will, however, in general be not as good as when one would have been able to realize n directly. This shows, for example, when the measurement is used for state separation: the EVM n separates the states ρ at least as well[iii] as m (th. 17). The separation is equally good in the case of equivalence (th. 18).

THEOREM 17 $\quad \mathfrak{n} \rightarrow \mathfrak{m} \Rightarrow \forall_{\rho_1,\rho_2} \sum_{\ell \in \mathbb{L}} |\mathrm{Tr}\left[(\rho_1-\rho_2)N_\ell\right]| \geq \sum_{k \in \mathbb{K}} |\mathrm{Tr}\left[(\rho_1-\rho_2)M_k\right]| .$

THEOREM 18 $\quad \mathfrak{n} \longleftrightarrow \mathfrak{m} \Rightarrow \forall_{\rho_1,\rho_2} \sum_{\ell \in \mathbb{L}} |\mathrm{Tr}\left[(\rho_1-\rho_2)N_\ell\right]| = \sum_{k \in \mathbb{K}} |\mathrm{Tr}\left[(\rho_1-\rho_2)M_k\right]| .$

(Proofs in the appendix; ρ_1, ρ_2 density operators.)

As noted above, non-ideality can be inverted in the sense that a non-ideal measurement can be used to estimate one or several parameters (linear functionals) of the probability distribution of the observable one wanted to measure. For a given labeling one might think of the mean value (in general the only parameter that can be estimated *via* the approach of § 2.1), or a higher moment, or the probability that the outcome lies in some interval. In general this job can be summarized as follows:

> Estimate the expectation value of a given operator $F \in \mathrm{L}(\mathfrak{n})$, using a measurement of $\mathfrak{m}$, $\mathfrak{n} \rightarrow \mathfrak{m}$.

A measure for the number of repetitions of the experiment needed to estimate this parameter with a given reliability is the variance: the larger the variance, the larger the number of repetitions needed[60] (we shall assume the variance to be finite). We shall show that, if the estimate is possible at all using an $\mathfrak{m}$–measurement, the number of repetitions is necessarily at least as large as the number one would have needed, had one been able to realize $\mathfrak{n}$ (th. 19). This theorem may be considered an analog within the non-ideality framework of the earlier result of de Muynck *et al.* [see § 2.1]. In the case of the relation $\overset{i}{\rightarrow}$ we speak of "invertibility" because it is possible to estimate $\langle F \rangle$ for *any* $F \in \mathrm{L}(\mathfrak{n})$ (th. 20). This means that it is possible to calculate the entire $\mathfrak{n}$–distribution, if the $\mathfrak{m}$–distribution is given. Then any question that can be answered using an $\mathfrak{n}$–measurement, can also be answered using an $\mathfrak{m}$–measurement, although in the latter case it may take more repetitions of the experiment. In the case of equivalence, no reason exists to prefer either measurement over the other (th. 21).

---

[60] E.g. Helstrom, *op. cit.* (1976) § I.4.

THEOREM 19 Suppose that $n \longrightarrow m$ and that $F \in L(n)$ is given. Then one of the following alternatives is true:

i) $F \notin L(m)$;

ii) For every sequence $(g_k)_K$ such that $\sum_{k\in K} g_k M_k = F$ there is a sequence $(f_\ell)_L$ such that $\sum_{\ell\in L} f_\ell N_\ell = F$ and

$$\sum_{k\in K} g_k^2 M_k \geq \sum_{\ell\in L} f_\ell^2 N_\ell \ .$$

THEOREM 20 Suppose that $n \xrightarrow{i} m$ and that $F \in L(n)$ is given. Then $F \in L(m)$. For every sequence $(g_k)_K$ such that $\sum_{k\in K} g_k M_k = F$ there is a sequence $(f_\ell)_L$ such that $\sum_{\ell\in L} f_\ell N_\ell = F$ and

$$\sum_{k\in K} g_k^2 M_k \geq \sum_{\ell\in L} f_\ell^2 N_\ell \ .$$

THEOREM 21 Suppose that $n \longleftrightarrow m$ and that $F \in L(n)$ is given. Then $F \in L(m)$. For every sequence $(f_\ell)_L$ such that $F = \sum_{\ell\in L} f_\ell N_\ell$ one of the following alternatives is true:

i) There is a sequence $(f'_\ell)_L$ such that $\sum_{\ell\in L} f'_\ell N_\ell = F$ and

$$\sum_{\ell\in L} f'^2_\ell N_\ell \begin{Bmatrix} \leq \\ \neq \end{Bmatrix} \sum_{\ell\in L} f_\ell^2 N_\ell \ ;$$

ii) There is a sequence $(g_k)_K$ such that $\sum_{k\in K} g_k M_k = F$ and

$$\sum_{k\in K} g_k^2 M_k = \sum_{\ell\in L} f_\ell^2 N_\ell \ .$$

(Proofs can be found in the appendix.)

On a more conceptual level we may say that, if $n \longrightarrow m$, an m–measurement result is to be interpreted as a *fuzzy* n–measurement result. The matrix $\{\lambda_{k\ell}\}$ has in this connection been called a *confidence function* by Prugovečki[61], meaning that a particular

[61] E. Prugovecki (1977): *J. Phys. A* **10**, p. 543

m result corresponds to a n–result with a confidence proportional to $\lambda_{k\ell}$. Such an interpretation is close to a likelihood interpretation of the non-ideality matrix. It might be tempting to substantiate this by claiming that the *probability* that the real (in the naive sense of the word) result (of the n–measurement, that could unfortunately not be performed) was ℓ, given that our non-ideal measurement (i.e. of m) gave k, is proportional to $\lambda_{k\ell}$. The other way round, $\lambda_{k\ell}$ could represent the probability that the non-ideal measurement gives result k where an ideal measurement would have given ℓ. But of course all such statements cannot be taken literally (let alone that an interpretation of the structure can be based on them), since there is generally no event corresponding to a "real value", so that any talk of "probability" is in this connection at best a figure of speech.

The requirement of labeling insensitivity urged in § 2.1, is conveniently represented in this structure by the fact that two EVMs differing only in labeling are members of the same equivalence class (th. 4). An equivalence class consists of EVMs representing devices that measure physically identical quantities equally well. Another consequence of label-independence is that, as is readily seen, defs. 2 and 3 apply to cases where m and n have very different outcome sets. It is for example perfectly possible to have $n \rightarrow m$, where m has outcome set {red, yellow, blue} and n has outcome set $\mathbb{N}$ (or *vice versa*). Hence, in contrast to convexity, the non-ideality relation structures the class of EVMs as a whole, irrespective of outcome set.

These considerations show that i–maximal (or even maximal) pairwise linearly independent EVMs are likely candidates to represent the "optimal" observables referred to in the § 2.2. The observables of most textbooks (~ PVMs) are optimal in this sense, just like pure observables (which were the observables that appeared optimal from the convexity point of view). There are many others, however[62]. Moreover, calling observables corresponding to i–maximal EVMs optimal has some operational justification: they cannot be improved upon in the sense of th.'s 17 through 21. Such justification is not available for the other two classes mentioned.

[62]An interesting example occurs on p. 74 of Helstrom, *op. cit.* (1976).

2.4 Non-ideality Measures

If $\mathfrak{n} \longrightarrow \mathfrak{m}$, it is for certain purposes (e.g for an inaccuracy relation !) it is convenient to express *how* non-ideal the non-ideal $\mathfrak{n}$–meter is, as a real number. As we saw in the previous section, $\lambda_{k\ell}$ represents, roughly speaking, the likelihood that the $\mathfrak{m}$ result would have been ℓ, given that the $\mathfrak{n}$ result was k. A non-ideality measure should obviously quantify the width of the ℓ–distribution $\lambda_{k\ell}$ for given k. More precisely, we need a mapping from the set of non-ideality matrices to $\mathbb{R}^+$, consistent with $\longrightarrow$ (see th. 22 and th. 23).

Matrices like $\{\lambda_{k\ell}\}$ should be well known to readers familiar with information theory. There they are used to represent discrete memoryless channels. If we restrict ourselves to the case of finite outcome sets, information theory also supplies us with a measure for the non-ideality represented by a matrix $\{\lambda_{k\ell}\}$: Shannon's channel capacity[63]. Despite the fact that the meaning of $\{\lambda_{k\ell}\}$ is very different in information theory, the channel capacity (and related quantities) can be used here, too. Of course there are many other measures than the ones we derive here[iv]. Just as the choice of scatter measure is not crucial for the expression of the scatter principle [§ 1.2], the choice for one particular non-ideality measure is not of fundamental importance.

Suppose two EVMs $\mathfrak{n} = \{N_\ell\}_L$ and $\mathfrak{m} = \{M_k\}_K$ are given such that $\mathfrak{n} \longrightarrow \mathfrak{m}$ with non-ideality matrix $\{\lambda_{k\ell}\}$; $L = \{\ell_1, \ldots, \ell_n\}$. For a given probability distribution $(p_\ell)_L$, define the *conditional entropy*:

$$J\left(\{\lambda_{k\ell}\};(p_\ell)\right) := -\sum_{k\in K}\sum_{\ell\in L} q_{k\ell}\log(q_{k\ell}/r_k) \quad ; \tag{62}$$

with:

$$q_{k\ell} := \lambda_{k\ell}p_\ell \quad ; \quad r_k := \sum_{\ell\in L}\lambda_{k\ell}p_\ell \quad (k \in K;\ \ell \in L).$$

---

[63] C. Shannon (1948): *Bell Syst. Techn. J.* **27**, p. 379; R. McEliece (1977): *The Theory of Information and Coding* (Addison-Wesley, London), ch. I. Cf. Shannon's ordering of communication channels [C. Shannon (1958): *Inform. Control* **1**, p. 390], which is similar to $\longrightarrow$.

Define further the *mutual information*:

$$(63) \qquad I\left(\{\lambda_{k\ell}\};(p_\ell)\right) := \sum_{k\in K}\sum_{\ell\in L} q_{k\ell}\log(q_{k\ell}/(r_k p_\ell)) \quad ,$$

and the *channel capacity*:

$$(64) \qquad C\left(\{\lambda_{k\ell}\}\right) := \sup_{(p_\ell)_L}\left\{I\left(\{\lambda_{k\ell}\};(p_\ell)\right)\right\} \quad .$$

The following properties of the capacity are easily verified[64]:

$$(65) \qquad 0 \le C\left(\{\lambda_{k\ell}\}\right) \le \log(n) \quad ;$$

$$(66) \qquad C\left(\{\lambda_{k\ell}\}\right) = 0 \;\Leftrightarrow\; \lambda_{k\ell} = \tilde{\lambda}_k \quad (k \in K;\, \ell \in L);$$

$$(67) \qquad C\left(\{\lambda_{k\ell}\}\right) = \log(n) \text{ iff there is for every } k \in K \text{ at most one } \ell \in L \text{ such that } \lambda_{k\ell} > 0 \,.$$

The conditional entropy J can be seen to satisfy similar properties. The interpretation of C in this context is straightforward: the larger $C[\{\lambda_{k\ell}\}]$ is, the smaller the amount of non-ideality in $\{\lambda_{k\ell}\}$ is. For J the reverse holds.

Up to this point we have, by associating C and J, as defined above, with a given $\rightarrow$, implicitly assumed that the matrix $\{\lambda_{k\ell}\}$ is unique. This condition is in general fulfilled only if $B(n)$ is a simplex and n is both self-extremal and pairwise linearly independent. PVMs satisfy this condition, but not all EVMs do. Hence C and J are not guaranteed to be compatible with $\rightarrow$. This is especially clear if we take m minimal. In that case we would expect from a reasonable non-ideality measure that it assumes its largest value for all n. If $\{\lambda_{k\ell}\}$ is not unique, $-C$ and J do not necessarily have this property. Therefore we define a capacity tailored to fit a given relation $n \rightarrow m$. Define the set Λ of all matrices $\{\lambda_{k\ell}\}$ connecting the EVMs n and m:

$$(68) \qquad \Lambda_{n\rightarrow m} := \left\{\{\lambda_{k\ell}\} \mid \lambda_{k\ell} \ge 0 \;\wedge\; \sum_{k\in K}\lambda_{k\ell} = 1 \;\wedge\; \sum_{\ell\in L}\lambda_{k\ell} M_\ell = N_k\right\} \quad .$$

[64] McEliece, *op. cit.*

In view of $n \longrightarrow m$, the set Λ can be seen as a property of the m–meter characterizing its functioning as a non-ideal n–meter. Define also:

$$\tilde{J}_{n\longrightarrow m}[(p_\ell)] := \sup_{\{\lambda_{k\ell}\}\in\Lambda_{n\longrightarrow m}} \left\{J\left(\{\lambda_{k\ell}\};(p_\ell)\right)\right\} \quad ; \tag{69}$$

$$\tilde{I}_{n\longrightarrow m}[(p_\ell)] := \inf_{\{\lambda_{k\ell}\}\in\Lambda_{n\longrightarrow m}} \left\{I\left(\{\lambda_{k\ell}\};(p_\ell)\right)\right\} \quad ; \tag{70}$$

$$\tilde{C}_{n\longrightarrow m} := \sup_{(p_\ell)} \left\{\tilde{I}_{n\longrightarrow m}[(p_\ell)]\right\} \quad . \tag{71}$$

The probability distribution (p_ℓ) in (69) is a parameter in the mapping $\tilde{J}$ represents. It should in particular not be thought of as associated with some physical object state. For different choices of (p_ℓ) different non-ideality measures result. This arbitrariness can be useful when one, for instance, wants to emphasize that accuracy in some part of the n outcome set is more important than in the rest.

To check whether these measures are indeed a consistent quantification of non-ideality, we introduce the following notions:

DEFINITION 9 Suppose $n \longrightarrow m \longrightarrow o$. Then:

- a non-ideality measure Q is *right order preserving*
  $:=$
  $Q_{n\longrightarrow o} \geq Q_{n\longrightarrow m}$;
- a non-ideality measure Q is *left order preserving*
  $:=$
  $Q_{n\longrightarrow o} \geq Q_{m\longrightarrow o}$.

If Q is right order preserving, it satisfies $Q_{n\longrightarrow o} = Q_{n\longrightarrow m}$ for all $n \longrightarrow m \longleftrightarrow o$. Thus a right order preserving $Q_{n\longrightarrow o}$ is insensitive to the labeling of o. Left order preservation has an analogous consequence. The measure based on channel capacity is completely insensitive to labeling:

THEOREM 22 The measure $-\tilde{C}$ is both left and right order preserving.

(Proof in appendix.) Now look at a continuous example: the non-ideal measurement of position q (we denote the position PVM $\{|q\rangle\langle q|\ dq\}_{\mathbb{R}}$ by q) *via* some EVM m. Non-ideality measures $\delta_{q\to m}$ for this case are usually attributed the same dimension as position: they are expressed in some *unit*. But that means that, from a mathematical point of view, $10\delta_{q\to m} = \delta_{10q\to m} \neq \delta_{q\to m}$ even though $10q \longleftrightarrow q$. Consequently, in view of possible generalizations of the theory to continuous outcome sets, it is not necessarily wise to insist on left order preservation. Right order preservation, on the other hand, is connected with the labeling of the non-ideal measurement device itself. This labeling can, as argued above, be altered arbitrarily without changing the physics of the device. Right order preservation *is* essential. Thus J is more suitable for a generalization to the continuous case, because:

THEOREM 23 The measure $\tilde{J}$ is only right order preserving.

Since it is also easier to calculate, we shall use J rather than C in the following.

In the previous sub-section we saw that a non-ideal measurement can be used to estimate linear functionals of the probability distribution of the desired observable. This, however, resulted in a lower reliability. Suppose $n \to m$ and $F = \sum_\ell f_\ell N_\ell = \sum_k g_k M_k$. Then

DEFINITION 10
$$\epsilon^{(F)}_{n\to m} := \left[\inf_{(g_k)|F=\sum_k g_k M_k}\left[\sum_k g_k^2\left(\sum_\ell \lambda_{k\ell} p_\ell\right)\right] + - \inf_{(f_\ell)|F=\sum_\ell f_\ell N_\ell}\left[\sum_\ell f_\ell^2 p_\ell\right]\right]^{\frac{1}{2}}$$

is an entirely different non-ideality measure, based on this estimation possibility. The larger ϵ is, the worse the measurement. The fixed distribution (p_ℓ) is to be chosen as realizable via $\mathrm{Tr}(\rho N_\ell)$. Therefore the choice of $\{\lambda_{k\ell}\}$ from Λ is arbitrary. The distribution (p_ℓ) plays the same role in ϵ as it did in J. The measure ϵ is both left and right order preserving, as can be easily seen from the results in § 2.3. The problem we noted above concerning the generalization to continuous outcome sets does not occur here because ϵ is, unlike J, not based directly on the non-ideality matrix. The measure ϵ has the dimension of f (or g). A change in n's parameterization (such as $q \to 10q$) will be precisely compensated for by a change in f necessitated by the condition that the expectation value of the linear functional it represents remain equal to F.

2.5 Joint Non-ideal Measurement

Given the above, the natural definition of joint (non-ideal) measurement is:

DEFINITION 11 The EVM $\mathfrak{o} = \{O_{mm'}\}_{M\times M'}$ is associated with a *joint non-ideal measurement* of the EVMs $\mathfrak{m}$ and $\mathfrak{n}$

$$:=$$

$$\mathfrak{n} \longrightarrow \mathfrak{o}^{(2)} \ \wedge \ \mathfrak{m} \longrightarrow \mathfrak{o}^{(1)} \ ,$$

with:

$$\begin{cases} \mathfrak{o}^{(1)} := \{\sum_{m'\in M'} O_{mm'}\}_{M} \\ \mathfrak{o}^{(2)} := \{\sum_{m\in M} O_{mm'}\}_{M'} \end{cases} . \tag{31}$$

A *joint invertibly non-ideal measurement* is defined analogously[v], using $\xrightarrow{i}$.

A particularly interesting joint measurement problem arises when $\mathfrak{a}$ and $\mathfrak{b}$ are PVMs of the self-adjoint operators A and B respectively. For this case, th. 16 suggests that we find two operators C and D on an auxiliary Hilbert space $\mathscr{H}'$ and two functions f and g such that $f(A,C)$ and $g(B,D)$ commute. In § 2.1 we already met such a solution to the joint measurement problem. We can solve this problem generically in a simple way by using projectors E and $\bar{E} = 1-E$ on $\mathscr{H}'$ to define two operators on $\mathscr{H}\otimes\mathscr{H}'$:

$$F := A\otimes E \ ; \ G := B\otimes\bar{E} \ . \tag{72}$$

The operators F and G commute, and therefore have a joint PVM. For an arbitrary ρ' on $\mathscr{H}'$ this gives an EVM on $\mathscr{H}$ in the following way:

> Obviously $\sigma(F) = \sigma(A) \cup \{0\}$. If 0 is not already an eigenvalue of A (i.e. $0 \notin \sigma(A)$), introduce $M_0 = 0$ into the PVM. Thus we have a PVM $\mathfrak{a} = \{M_a\}_{\sigma(F)}$ on $\mathscr{H}$. Analogously we get a PVM $\mathfrak{b} = \{N_b\}_{\sigma(G)}$ on $\mathscr{H}$.

Then the joint PVM $\{P_{ab}\}_{\sigma(F)\times\sigma(G)}$ on $\mathscr{H}\otimes\mathscr{H}'$ is given by:

$$(73)\qquad P_{ab} = \delta_{a0}\delta_{b0}(N_0\otimes E + M_0\otimes\bar{E}) + (1-\delta_{a0})\delta_{b0}\,M_a\otimes\bar{E} + \\ + (1-\delta_{b0})\delta_{a0}N_b\otimes E = \\ = \delta_{b0}\,M_a\otimes\bar{E} + \delta_{a0}\,N_b\otimes E \ .$$

Taking the partial trace over $\mathscr{H}'$ gives the EVM[65] $o = \{O_{ab}\}_{\sigma(F)\times\sigma(G)}$,

$$(74)\qquad O_{ab} = \mathrm{Tr}_{\mathscr{H}'}(\rho'\,P_{ab}) = \lambda\,\delta_{b0}M_a + (1-\lambda)\,\delta_{a0}N_b \ ,$$

with $\lambda = \mathrm{Tr}(\rho'\,E)$.

Indeed o can be seen to satisfy def. 11 for $\mathfrak{a}$ and $\mathfrak{b}$. Relation (74) can evidently be extended to satisfy def. 11 for two arbitrary incompatible EVMs. To clarify its interpretation, we introduce the following notions:

DEFINITION 12 – A joint measurement of two EVMs as in def. 11 is *trivial*[66]

$$:=$$

$$\forall_{(m,m')\in M\times M'}\ O_{mm'} \in L\left[\{M_k\}_K \cup \{N_l\}_L\right]$$

– A joint measurement of two EVMs as in def. 11 is an *either/or measurement*

$$:=$$

$$\exists \left\{ \begin{array}{l} \text{EVMs } \{\tilde{M}_{mm'}\}_{M\times M'\longleftarrow m} \\ \qquad\quad \{\tilde{N}_{mm'}\}_{M\times M'\longleftarrow n} \\ \lambda \in (0,1) \end{array} \right. O_{mm'} = \lambda\tilde{M}_{mm'} + (1-\lambda)\tilde{N}_{mm'}$$

One of the most important reasons to do a joint measurement of two observables is the possibility of determining some kind of correlation between the values these two

---

[65] Abu-Zeid [O. Abu-Zeid (1987): *Phys. Lett. A* **125**, p. 167] proposed an EVM essentially the same as eq. 80, as does Schroeck [F. Schroeck (1982): *Found. Phys.* **12**, p. 479] for $\mathbb{C}^2$. Neither notes its physical interpretation as an either/or measurement, however.

[66] As a joint measurement of A and $f(A)$ is always trivial, our triviality notion is a generalization of that of Park & Margenau, *op. cit.*.

observables "assume". In a trivial joint measurement, the joint probability distribution is fixed if the m and n distributions are given. Hence the correlation in the joint probability distribution depends on the particular procedure used, but it does not depend on the state of the object system at all. Trivial joint measurements do not provide information about ρ beyond the information obtainable with separate measurements of the observables involved[67]. An either/or measurement is always trivial because the joint EVM is a convex sum of two EVMs equivalent to the EVMs to be measured jointly. An either/or measurement can be realized using a procedure where sometimes (probability λ) m is measured, and sometimes (probability $1-\lambda$) n . The outcomes are then suitably relabeled to disguise them as joint results. The EVM of (74) represents an either/or measurement. There the relabeling consists of attributing the value 0 to the observable that is not measured.

Thus we shall have to search a little harder to find a meaningful new application of th. 16 [see § 2.6] and, more generally, a true joint non-ideal measurement of two incompatible observables. As the following results show, such a measurement can be found. We take $\mathscr{H} = \mathbb{C}^n$. Two PVMs $\mathfrak{e} = \{E_m\}_M$ and $\mathfrak{f} = \{F_\ell\}_L$ are given. Since they are completely arbitrary, apart from the finite dimensionality of the Hilbert space, our results are (nevertheless) quite general. We will assume that there is a third PVM $\{G_i\}_I$ satisfying

$$(75) \qquad \forall_i (\forall_m \ [G_i,E_m]_- = 0 \ \wedge \ \forall_\ell \ [G_i,F_\ell]_- = 0) \ .$$

This PVM is introduced to resolve cases where the PVMs $\mathfrak{e}$ and $\mathfrak{f}$ have eigenspaces in common. It may be minimal (i.e. $\{G_i\}_I = \{1\}$), so that this assumption does not result in loss of generality. Then[vi]:

THEOREM 24 The PVMs $\mathfrak{e}$ and $\mathfrak{f}$ (defined above) are jointly non-ideally measurable. The joint measurement can be both non-trivial and invertible.

[67] In this sense also the proposals of Park & Margenau, *op. cit.*, are trivial.

THEOREM 25 Suppose an EVM $\mathrm{m} = \{M_{jj'}\}_{J\times J'}$ represents a joint non-ideal measurement (as in def. 11) of the PVMs e and f (defined above). Then the quality of this joint measurement is limited by:

(76) $$J_{\mathrm{e}\to\mathrm{m}(1)} + J_{\mathrm{f}\to\mathrm{m}(2)} \geq \sum_{i\in I} \mathrm{Tr}(G_i)\, c_i$$

with:

(77) $$c_i := -2\log[\tfrac{1}{4} + \tfrac{1}{4}\max_{k,\ell}\|E_k F_\ell G_i\|^{\frac{1}{2}}] \ ;$$

(78) $$J_{\mathrm{e}\to\mathrm{m}} := \tilde{J}[\Lambda_{\mathrm{e}\to\mathrm{m}}; (r_m)] = J[\{\lambda_{jm}\}; (r_m)] =$$

$$= -\sum_{j\in J}\sum_{m\in M} \lambda_{jm}\, r_m \log\left[\lambda_{jm}\, r_m / (\textstyle\sum_k \lambda_{jk}\, r_k)\right] \ ;$$

(79) $$r_m = \tfrac{1}{n}\mathrm{Tr}(E_m) \ .$$

For maximal PVMs this bound can be improved to

(77′) $$c_i := -\log[\max_{m,\ell}\|E_m F_\ell G_i\|] \ ;$$

(Proofs in the appendix.) Because, as we noted above, $0 \leq J$ for non-ideal measurements of e and f, inequality (76) gives a non-trivial bound iff the PVMs are incompatible (for an optimal choice of $\{G_i\}_I$). The bound of (76) is in particular non-trivial if the PVMs have (some, not all) eigenvectors in common. As an example we can consider a three-dimensional situation, with maximal PVMs e and f with inner product matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ 0 & \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{bmatrix} \ .$$

In this case we can take $\{G_i\}_I = \{G_1, G_2\} = \{E_1, E_2 + E_3\}$. We get the bound log(2) for (76)+(77′) [log(24-16$\sqrt{2}$) for (76)+(77)], well above 0.

2.6 Examples

In order to illustrate the application of non-ideality and the resulting approach to the joint measurement problem further, we shall study a number of examples. The first of these involves maximal PVMs on an n-dimensional Hilbert space. Define the PVM $\mathfrak{x} = \{E_k\}_K$,

(80) $$E_k := |x_k\rangle\langle x_k|$$

$(|x_k\rangle)_K$ is an orthonormal base for $\mathscr{H} = \mathbb{C}^n$; $K = \{0,\ldots,n-1\}$;

and:

(81) $$S_x := \sum_{k\in K} |x_{[[k+1]]}\rangle\langle x_k| \quad ;$$

$$[[k]] := k \bmod n \quad ;$$

$$X := \sum_{k\in K} kE_k \quad .$$

Apart from *unbiasedness* [§ 2.1], *covariance* is an important characteristic for measurements:

DEFINITION 13 – The EVM $\mathfrak{n} = \{N_\ell\}_\ell$ with the labeling $(f_\ell)_L$ is unbiased (with respect to X)

$$:=$$

$$\sum_{\ell\in L} f_\ell N_\ell = X \quad ;$$

– The EVM $\mathfrak{o} = \{O_k\}_K$ is covariant (with respect to S_x)

$$:=$$

$$\forall_{k\in K}\ S_x O_k S_x^\dagger = O_{[[k+1]]} \quad ;$$

(S_x and K as above; L arbitrary).

The following theorem connects these notions to $\longrightarrow$:

THEOREM 26 For $\mathfrak{x}$ as above and $o = \{O_k\}_K$ an EVM (K as above):

$$\left.\begin{array}{l} \forall_{k\in K}\ S_x O_k S_x^\dagger = O_{[[k+1]]} \\ \exists_{(f_k)_K}\ \sum_k f_k O_k = X \end{array}\right\} \Longrightarrow \mathfrak{x} \xrightarrow{i} o \ .$$

Neither covariance nor unbiasedness (for some labeling) is by itself sufficient for $\longrightarrow$ or $\xrightarrow{i}$.

We now introduce a second orthonormal base $(|y_k\rangle)_K$ on $\mathbb{C}^n$, such that:

$$\langle x_k | y_{k'} \rangle = n^{-\frac{1}{2}} \exp(i\tfrac{2\pi}{n} kk') \ , \tag{82}$$

with shift operator S_y, self-adjoint operator Y and PVM $\mathfrak{y}$, analogous to S_x, X and $\mathfrak{x}$ respectively. These two PVMs are in fact finite-dimensional analogs of position and momentum[68,vii]. They satisfy

$$\forall_{c,d\in \mathbb{Z}}\ S_y^c S_x^d = S_x^d S_y^c \exp(i\tfrac{2\pi}{n} cd) \ ,$$

analogous to the Weyl commutation relations. For these PVMs the bound (77') gives

$$J_{\mathfrak{x}\longrightarrow \mathfrak{m}(1)} + J_{\mathfrak{y}\longrightarrow \mathfrak{m}(2)} \geq \log(n) \ . \tag{83}$$

The demand of covariance, discussed above, leads[69] to the following criterion (more stringent than def. 11) for a joint measurement of the PVMs $\mathfrak{x}$ and $\mathfrak{y}$:

DEFINITION 14 An EVM $\mathfrak{m} = \{M_{k\ell}\}_{K\times K}$ is associated with a covariant joint non-ideal measurement of the PVMs $\mathfrak{x}$ and $\mathfrak{y}$

$$:=$$

$$\forall_{k,\ell\in K}\ S_x M_{k\ell} S_x^\dagger = M_{[[k+1]]\ell} \wedge S_y M_{k\ell} S_y^\dagger = M_{k[[\ell+1]]}$$

(S_x, S_y and K as defined above, $\mathfrak{x}$ and $\mathfrak{y}$ defined as above including (82)).

---

[68] J. Schwinger (1960): *Proc. Nat. Acad. Sci.* **46**, p. 570

[69] Helstrom, *op. cit.* (1976); Holevo, *op. cit.* (1982) p. 122

We can see that this nomenclature is consistent, that the covariance of def. 14 indeed implies association with a joint non-ideal measurement, if we realize that the covariance as in def. 14 leads to invariance of the marginals:

THEOREM 27 For $\mathfrak{x}$ and S_y as above and $\mathfrak{o} = \{O_\ell\}_L$ an EVM:

$$\mathfrak{x} \to \mathfrak{o} \;\Leftrightarrow\; \forall_{\ell \in L}\; S_y O_\ell S_y^\dagger = O_\ell \quad (\textit{invariance}) \quad .$$

Indeed an explicit construction of a covariant joint measurement EVM through th. 16 establishes the usability of def. 14. The analogy of this example with position-momentum is further illustrated by this application of th. 16 [cf. the way the position-momentum was treated in § 2.1]:

THEOREM 28 Take $\mathscr{H}' = \mathbb{C}^n$, and define operators X' and Y' on $\mathscr{H}'$ just as X and Y were defined on $\mathscr{H}$ (incl. (82)). Then the operators $X'' := [[X \otimes 1' + 1 \otimes X']]$ and $Y'' := [[Y \otimes 1' - 1 \otimes Y']]$ on $\mathscr{H} \otimes \mathscr{H}'$ commute.

Now, using a density operator ρ_0 on $\mathscr{H}'$, we get our covariant joint measurement EVM from the joint PVM of X'' and Y'' by taking the partial trace, as in th. 16. The $\mathbb{C}^2$ case of this example is formed by the pair of spin $\frac{1}{2}$ PVMs $\mathfrak{s}_3$ and $\mathfrak{s}_1$, corresponding to the self-adjoint (Pauli) operators σ_3 and σ_1, respectively. Our requirements lead to an unconditionally non-trivial lower bound to the non-ideality achievable, as substitution of $n = 2$ into (83) shows. Th. 28 here amounts to construction of the self-adjoint operators $\sigma_1 \otimes \sigma_3'$ and $\sigma_1 \otimes \sigma_3'$, which can easily be seen to commute (σ_1' and σ_3' are spin operators defined on an auxiliary Hilbert space $\mathscr{H}'$, analogously to σ_1 and σ_3). It remains doubtful, however, whether th. 16 can solve joint measurement problems that are not (analogs of) position-momentum.

Since for the EVM $\mathfrak{m}$ of def. 14 $M_{00} \geq 0$, and $M_{k\ell} = (S_x^k S_y^\ell) M_{00} (S_x^k S_y^\ell)^\dagger$, the condition $\sum_k \sum_\ell M_{k\ell} = 1$ implies that $\mathrm{Tr}(M_{00}) = n^{-1}$. This means there is a density

operator ρ_0 on $\mathscr{H}$ such that[70]

$$M_{k\ell} = \tfrac{1}{n}(S_x^k S_y^\ell)\rho_0(S_x^k S_y^\ell)^\dagger \quad . \tag{84}$$

The EVM m of def. 14 is maximal iff the density operator ρ_0 is pure. It is not difficult to verify that the non-ideality matrices for $\mathfrak{x} \longrightarrow \mathfrak{m}^{(1)}$ and $\mathfrak{y} \longrightarrow \mathfrak{m}^{(2)}$ respectively, are given by

$$\begin{cases} \lambda_{k\ell}^{(1)} = \langle x_{[[\ell-k]]}|\rho_0|x_{[[\ell-k]]}\rangle \\ \\ \lambda_{k\ell}^{(2)} = \langle y_{[[\ell-k]]}|\rho_0|y_{[[\ell-k]]}\rangle \end{cases} . \tag{85}$$

These non-ideality matrices are *symmetric*[71]: every row is a permutation of every other row and every column is a permutation of every other column. Then we can easily evaluate the amount of non-ideality (using the measure based on channel capacity or that based on conditional entropy) in the matrices (85). We see that it is related to the scatter in ρ_0:

$$\begin{cases} J_{\mathfrak{x} \longrightarrow \mathfrak{m}^{(1)}} = H_X[\rho_0] = \log(n) - C_{\mathfrak{x} \longrightarrow \mathfrak{m}^{(1)}} \\ \\ J_{\mathfrak{y} \longrightarrow \mathfrak{m}^{(2)}} = H_Y[\rho_0] = \log(n) - C_{\mathfrak{y} \longrightarrow \mathfrak{m}^{(2)}} \end{cases} , \tag{86}$$

where

$$H_Y[\rho_0] := -\sum_{k\in K} \langle y_k|\rho_0|y_k\rangle \log(\langle y_k|\rho_0|y_k\rangle) \quad \text{(analogously for } H_X\text{)}.$$

Note that the ℓ width of $\lambda_{k\ell}^{(i)}$ for given k does nor depend on k because both matrices are covariant. Thus no weighting distribution is necessary. Indeed J is independent of the (p_ℓ)–distribution of (62) or (69). In this n dimensional case $0 \leq J \leq \log(n)$. Minimal J is achieved iff the associated probability distribution is dispersion free. That is not achievable for both distributions at the same time, since the observables are incompatible. This is most conveniently represented by the Maassen-Uffink scatter relation (3), which reads for this case:

$$H_X[\rho] + H_Y[\rho] \geq \log(n) \quad . \tag{87}$$

---

[70]Holevo, *op. cit.* (1982) § III.6

[71]McEliece, *op. cit.*, § I.2.1

Combination of (86) and (87) gives[viii] a special case of (83), viz.

$$J_{\xi \to \mathfrak{m}^{(1)}} + J_{\eta \to \mathfrak{m}^{(2)}} \geq \log(n) \quad , \tag{88}$$

which is valid *only* for covariant observables.

The covariant EVM of def. 14 has one more interesting property. It can separate the states[72,ix] (*informational completeness*). The PVMs ξ and η do not have this property, not even together[73]. Trivial joint measurement EVMs therefore cannot achieve informational completeness either. The following theorem gives the demands ρ_0 is to satisfy, in order to get informational completeness:

THEOREM 29 There is a 1–1 relation between ρ and the probability distribution $(p_{k\ell})$ $(p_{k\ell} := \mathrm{Tr}(\rho \boldsymbol{M}_{k\ell})$, the EVM $\{\boldsymbol{M}_{k\ell}\}_{K\times K}$ as in def. 14) iff the density operator $\rho_0 := n\boldsymbol{M}_{00}$ satisfies:

$$\forall_{a,b\in K} \quad \mathrm{Tr}(\rho_0 \boldsymbol{S}_x^a \boldsymbol{S}_y^b) \neq 0$$

The covariant joint ξ,η measurement we just studied is, however, only of limited relevance because def. 14 depends for its consistency on the special properties of the operators $\boldsymbol{X}$ and $\boldsymbol{Y}$ in a highly non-trivial way. Moreover, covariance explicitly confines the labeling of the non-ideal observable: not only does it require L = K (M = K), but covariance also fixes the order of the elements in L (M) through the restrictions it places on the non-ideality matrices (i.e. $\lambda_{k\ell}^{(i)} = \tilde{\lambda}_{[[k-\ell]]}^{(i)}$). When applied to the analogous position case, covariance would exclude, for instance, the EVM $\{\boldsymbol{M}_n\}_{\mathbb{Z}}$, $\boldsymbol{M}_n := \int_n^{n+1} |x\rangle\langle x| dx$, as a non-ideal position measurement, whereas it is perfectly acceptable as such from a pragmatic point of view[74]. Hence the demand

---

[72]E. Prugovecki (1984): *Stochastic Quantum Mechanics and Quantum Space–Time* (Reidel, Dordrecht); P. Busch (1987): *Found. Phys.* **17**, p. 905; Busch, *op. cit.* (1985)

[73]V. Bargmann, as quoted by Reichenbach (H. Reichenbach (1944): *Philosophical Foundations of Quantum Mechanics* (Univ. of California press, Berkeley), p. 92); A. Vogt (1978): *Mathematical Foundations of Quantum Theory* (ed. by A. Marlow, Academic, NY), p. 368

[74]Cf. the inaccurate measurement of self-adjoint operators with continuous spectrum in von Neumann, *op. cit.* (1932).

of covariance conflicts with the criterion, urged in § 2.1, that the labeling, however convenient, should not play a *fundamental* role[x].

A further example is the joint measurement of number and position. The number operator N and the position operator Q, corresponding to PVMs $\mathfrak{k} = \{|n\rangle\langle n|\}_{\mathbb{N}}$ and $\mathfrak{q} = \{|q\rangle\langle q|\, dq\}_{\mathbb{R}}$, respecively, are defined in terms of boson annihilation and creation operators:

$$N = a^\dagger a \quad , \quad Q = \tfrac{1}{2}\sqrt{2}(a^\dagger + a) \quad ([a,a^\dagger]_- = 1). \tag{89}$$

For this case an inaccuracy relation is not available. Nevertheless, we can show a joint measurement to be feasible. Consider the following operators on $\mathscr{H}\otimes\mathscr{H}'$:

$$\begin{aligned} a_b &:= \sqrt{\eta}\, a\otimes 1' + \sqrt{1-\eta}\, 1\otimes a' \quad ; \\ a_c &:= \sqrt{1-\eta}\, a\otimes 1' - \sqrt{\eta}\, 1\otimes a' \quad ; \\ N_b &= a_b^\dagger a_b \quad ; \\ Q_c &= \tfrac{1}{2}\sqrt{2}(a_c^\dagger + a_c) \quad . \end{aligned} \tag{90}$$

The operator a' ($a'^\dagger$) is a boson annihiliation (creation) operator, defined on $\mathscr{H}'$. Therefore the operators a_b and $a_b^\dagger$ (as well as a_c and $a_c^\dagger$) also have all the properties of boson annihilation and creation operators. Accordingly, the operators N_b and Q_c have the properties of a number operator and a position operator, respectively. Moreover, since $[a_b,a_c]_- = 0$ and $[a_b,a_c^\dagger]_- = 0$, they commute. Now we perform a joint measurement of the observables corresponding to the operators N_b and Q_c on $\mathscr{H}\otimes\mathscr{H}'$.

We take $\rho' = |0\rangle\langle 0|$, where $|0\rangle$ is the harmonic oscillator ground state. Then it is not difficult to show that:

$$P_{Q_c}(dx) = \int_{-\infty}^{\infty} P_Q(dy)\, g\,(dx,y) \quad ; \tag{91a}$$

(91b) $$P_{N_b}(n) = \sum_{m=0}^{\infty} \mu_{nm} P_N(m)$$

with:

$$g(dq,r) = \frac{dq}{\sigma\sqrt{2\pi}} \exp[-\frac{1}{2\sigma^2}(q - \sqrt{1-\eta}\, r)^2] \quad ; \quad \sigma = \sqrt{\tfrac{1}{4}\eta} \; ;$$

$$\mu_{nm} = \begin{cases} 0 & \text{if } m < n \\ \binom{m}{n} \eta^n (1-\eta)^{m-n} & \text{otherwise} \end{cases} .$$

We have a joint non-ideal $\mathfrak{k}$–$\mathfrak{q}$ measurement. The non-ideality in (91b) can be seen as noise in the $\mathfrak{k}$–measurement. That it can be seen as independent of the object is the reason for the possibility of representing it by a fixed matrix $\{\mu_{nm}\}$. That μ_{nm} is not only a function of (n–m) implies that it is non-additive. As $\eta \longrightarrow 0$ the $\mathfrak{k}$–measurement becomes totally dominated by noise, whereas the $\mathfrak{q}$–measurement approaches ideality. If $\eta \longrightarrow 1$ the opposite happens. Both (91a) and (91b) are invertible, so that the measurement is a joint *invertibly* non-ideal measurement. In the case of (91a) this inversion takes place through deconvolution. In fact (91a) has already been treated in § 2.1, as

(92) $$\boldsymbol{X}_c := \sqrt{1-\eta}\, \boldsymbol{X}_o - \sqrt{\eta}\, \boldsymbol{X}_a \quad .$$

Relation (91b) has the inverse

(93) $$P_N(n) = \sum_{k=0}^{\infty} \lambda_{nk} P_{N_b}(k) \quad ; \quad \lambda_{nk} = \begin{cases} 0 & \text{if } k < n \\ \binom{k}{n} \eta^{-k} (\eta-1)^{k-n} & \text{otherwise} \end{cases} .$$

A possible way of realizing the EVM of (91) will be discussed in the next section.

For the important $\mathfrak{q},\mathfrak{p}$ case (we denote the momentum PVM $\{|p\rangle\langle p|\, dp\}_{\mathbb{R}}$ by $\mathfrak{p}$) an inequality as general as (76) has not been derived as yet, due to the complications brought about by the infinite dimensionality of the Hilbert space involved. However, as noted in § 2.1, important partial results have been achieved. The most far reaching

of these arises out of an application of the Arthurs & Goodman relation (46) in our formalism.

As an extension of def. 2 to $\mathfrak{m} \longrightarrow \mathfrak{n}$ for a possibly continuous EVM $\mathfrak{m} = \{M(\mathrm{d}x)\}_X$ over the field $\mathscr{F}_X$ on the outcome set X, we use [cf. (43)]

$$N(\Delta y) = \int M(\mathrm{d}x)\ g(\Delta y,x) \quad , \tag{94}$$

where $\{N(\mathrm{d}y)\}_Y$ is an EVM over the field $\mathscr{F}_Y$ on the outcome set Y, and $g(\mathrm{d}y,x)$ is a measure that satisfies

$$\forall_{\Delta y \in \mathscr{F}_Y} \forall_x \quad g(\Delta y,x) \geq 0 \quad ;$$

$$\forall_x \quad g(Y,x) = 1 \quad .$$

Now, suppose that $\mathfrak{n}$ has, in addition to $\mathfrak{q} \longrightarrow \mathfrak{n}$, the property that there exists a function f satisfying

$$0 = \int\int r(\mathrm{d}q)\ g(\mathrm{d}y,q)\ [f(y) - q] \tag{95}$$

for all absolutely continuous (w.r.t. Lebesque measure) probability distributions $r(\mathrm{d}q)$ (unbiasedness, cf. def. 13). Then (def. 10) a suitable non-ideality measure is

$$\epsilon^{(Q)}_{\mathfrak{q}\longrightarrow\mathfrak{n}} = \left[\int\int r(\mathrm{d}q)\ g(\mathrm{d}y,q)\ [f^2(y) - q^2]\right]^{\frac{1}{2}} \quad , \tag{96}$$

for a fixed probability distribution $r(\mathrm{d}q)$. Note that for the special case that we have a covariant $\mathfrak{n}$, ϵ does not depend on r. Now, if an EVM $\mathfrak{m} = \{M(dy,dz)\}_{Y\times Z}$ represents a joint non-ideal measurement of $\mathfrak{p}$ and $\mathfrak{q}$ following def. 11 and (94), such that the unbiasedness condition (95) is fulfilled by both marginals $\mathfrak{m}^{(1)}$ and $\mathfrak{m}^{(2)}$, it can be seen[75] to satisfy

$$\epsilon^{(Q)}_{\mathfrak{q}\longrightarrow\mathfrak{m}^{(1)}}\ \epsilon^{(P)}_{\mathfrak{p}\longrightarrow\mathfrak{m}^{(2)}} \geq \tfrac{1}{2} \quad , \tag{97}$$

[75] Arthurs & Goodman, *op. cit.*; § 2.1

for distributions $r(\mathrm{d}q)$ and $s(\mathrm{d}p)$, occurring in $\epsilon^{(Q)}_{\mathsf{q}\longrightarrow\mathsf{m}(1)}$ and $\epsilon^{(P)}_{\mathsf{p}\longrightarrow\mathsf{m}(2)}$, respectively, that are associated with the q– and p–distributions of some fixed density operator ρ_ϵ. The weighting distributions r and s are parameters in the mappings ϵ, and need not correspond to distributions in any physical system [§ 2.4]; both ϵ's are purely properties of the m device.

The inaccuracy relation (76) is valid for *all* PVMs on Hilbert spaces of *arbitrary* finite dimension. Finite dimensional spaces do not differ essentially from infinite dimensional ones as regards uncertainty relations (cf. the first example of this section). Moreover, as noted earlier, infinite dimensional results can be approximated arbitrarily closely by finite dimensional calculation. Therefore we shall in the following, pending a rigorous derivation assume than an inaccuracy inequality[xi]

(98) $$\delta_{\mathsf{q}\longrightarrow\mathsf{m}(1)}\,\delta_{\mathsf{p}\longrightarrow\mathsf{m}(2)} \geq 1$$

exists also for measurements that do not satisfy (95), its precise form depending on the definition of the inaccuracy measure δ.

2.7 Evaluation

We have derived for the first time an inaccuracy inequality (th. 25) that is applicable to a quite general class of pairs of incompatible observables, and that does not involve unnecessary restrictive assumptions like covariance and unbiasedness. This inequality shows that quantum mechanics indeed entails an inaccuracy principle. It shows of what nature the bounds are that quantum mechanics sets to our ability to measure. We shall end with a further evaluation of our approach to inaccuracy and to joint measurement, in the light of the other approaches mentioned and in the light of practical realizability.

Despite differences in emphasis, our approach is not *a priori* incompatible with Arthurs & Goodman's [§ 2.1] since it is often possible to derive the expectation value $\langle R \rangle$ from the $A'_R(\tau)$-results if only def. 2 is required. On the other hand, it is also quite possible that (an analog of) def. 2 is satisfied if only unbiasedness is required. But the expectation value $\langle R \rangle$ no longer takes absolute precedence among all $\langle f(R) \rangle$ in our approach, and usually many such parameters (linear functionals of the probability distribution) can be recovered, instead of only one [§ 2.3]. Hence dropping unbiasedness does not really imply any concessions in this area; on the contrary the definition of noisy measurement along our lines is a more restrictive one than that used in approaches based exclusively on the expectation value, such as Arthurs & Goodman's. This difference in outlook on the content of 'measurement' may be expressed by saying that we aim at an R–measurement, whereas Arthurs & Goodman [and others, see § 2.1] aim at an $\langle R \rangle$-measurement. As a result of this, our approach is *in principle* of more restricted applicability.

Within quantum estimation theory[76] (QET) the joint measurement of position and momentum has been treated, analogously to the derivation of (88) [viz. (45)]. But this derivation presupposed certain important properties of the observables to be measured, and is therefore not generally applicable [§ 2.6]. Indeed the QET literature offers no applications of their approach to joint measurement, other than position-momentum. Nevertheless, the formal joint measurement procedure described by (45) is experimantally relevant: both (balanced) heterodyning[77] and parametric amplification of "position" and "momentum" [78] [§ 4.1.3] give rise to excess noise describable by (42). Since Arthurs and Goodman generalize the above formalism, these are also realizations of their approach (heterodyning and amplification are in fact the examples given by Arthurs and Goodman). But since our approach generalizes the above formalism, too, these experiments cannot be used to differentiate between the two views.

---

[76]Helstrom, *op. cit.* (1976); Holevo, *op. cit.* (1982)

[77]H. Yuen & J. Shapiro (1978): *IEEE J. Inf. Th.* **IT-24**, p. 657; (1979): *ibid.* **IT-25**, p. 179; (1980): *ibid.* **IT-26**, p. 78

[78]Y. Yamamoto & H. Haus (1986): *Rev. Mod. Phys.* **58**, p. 1001

Non-ideality appears in a number of other more or less realistic situations. An example is photo-detection for a single-mode optical field[79]. Introducing the number states $|n\rangle$, eigenstates of the operator N, it is well known that a detector with quantum efficiency $\eta = 1$ realizes the PVM $\mathfrak{k} = \{E_n\}_{\mathbb{N}}$; $E_n = |n\rangle\langle n|$. In a more realistic situation, however, we have $\eta < 1$. In that case we measure the EVM[80] $\mathfrak{m} = \{M_k\}_{\mathbb{N}} \leftarrow \mathfrak{k}$,

$$(99) \qquad M_k = \sum_{n=k}^{\infty} E_n \binom{n}{k} \eta^k (1-\eta)^{n-k} = \sum_{n\in\mathbb{N}} \lambda_{kn} E_n$$

with:

$$\lambda_{kn} = \begin{cases} 0 & \text{if } n < k \\ \binom{n}{k} \eta^k (1-\eta)^{n-k} & \text{otherwise} \end{cases} .$$

In fact (99) is the same as (91b): the $\mathfrak{k}$ marginal of the formal scheme for joint measurement of q and $\mathfrak{k}$ is equivalent to inefficient photo-detection. Since (91b) had the inverse (93), we also have $\mathfrak{k} \xrightarrow{i} \mathfrak{m}$. In this example the "inaccuracy" is apparent: $\eta < 1$. Nevertheless we can recover the exact probability distribution: non-ideality does not imply that any information is irretrievably lost. We can in particular recover the correct number expectation value: the operator $A''_N = N_b/\eta$ on $\mathscr{H}\otimes\mathscr{H}'$ satisfies the unbiasedness criterion[81] (37). The noise operator $G''_N = A''_N - N$ on $\mathscr{H}\otimes\mathscr{H}'$ can be seen as an additive object dependent representation of the N–measurement noise, alternative to a non-ideality matrix. This further illustrates the relation between our approach and Arthurs & Goodman's.

Inspired by the equivalence of one marginal to inefficient detection, a realization of the joint $\mathfrak{k}$,q–measurement procedure is easily envisaged: let $\mathscr{H}$ represent one mode of the EM field, incident on a beam-splitter with transmittivity η. The outgoing beams are labeled b (transmitted) and c (reflected). We measure in the b beam photon number: N_b , and in the c beam "position": Q_c (e.g. through homodyning[82]).

---

[79]R. Loudon (1983): *Quantum Theory of Light* (2nd ed., Clarendon, Oxford)

[80]Loudon, *op. cit.*, p. 240

[81]Note, however, that the spectrum of $A''_{\mathbb{N}}$ differs from that of N, barring any naive identification of $A''_{\mathbb{N}}$ outcomes and N outcomes.

[82]Yuen & Shapiro, *op. cit.* (1978, 1979, 1980)

The spin ½ example was intractable using Yuen's approach, whereas Arthurs & Goodman's led to a bound which could be trivial [§ 2.1]. That is somewhat counter-intuitive: as two spin ½ variables are maximally incompatible[83], one would expect the limit on joint measurability to be especially stringent for this case. Accordingly, our approach results in an unconditionally non-trivial inaccuracy bound, to the highest achievable bound on $\mathbb{C}^2$ [§ 2.6].

Joint spin measurement along our lines is feasible[84] [§ 4.5.2]. Several experiments have been proposed by Busch for the analogous case of photon polarization[85]. A realization of a simplified version of one of these by Mittelstaedt *et al.*[86] to obtain mixed wave-particle behavior, shows that Busch's proposals are indeed doable.

Summarizing, we see that our approach leads to a notion of inaccuracy that can be seen as an apparatus property. Despite the fact that it is more stringent than the expectation value based approach, its practical applicability for joint measurements does not seem more limited at all. Moreover, it leads in some cases to interesting bounds where the expectation value based approach, due to its weaker notion of inaccurate measurement, leads to trivial bounds.In ch. IV we shall investigate some applications of non-ideality further.

3 INDEPENDENCE

From a certain point of view, QM can be seen as a theory of two types of devices: preparators and detectors (fig. 2). Thus it is natural to expect that the UP has consequences for both types of devices, that there are two forms of the UP. Indeed that is what we have found: scatter relations limit preparation and inaccuracy

[83]Kraus, *op. cit.* (1987)

[84]Schroeck, *op. cit.* (1982)

[85]Busch, *op. cit.* (1987)

[86]P. Mittelstaedt, A. Prieur & R. Schieder (1987): *Found. Phys.* 17, p. 891

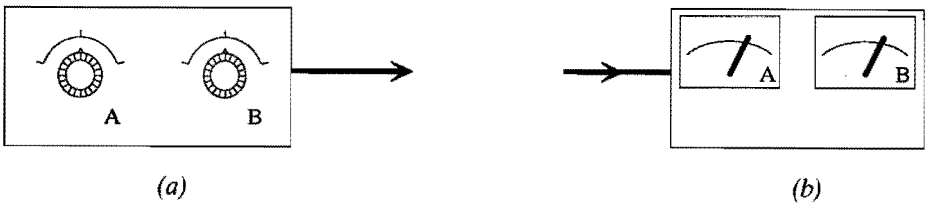


fig. 2 *Symbolic representation of the two types of devices for which the uncertainty principle has consequences: preparators (a) and meters (b). A preparator is symbolized by a (set of) knob(s) having a well-defined position adjusted by the experimenter. A meter is symbolized by a (set of) pointer(s) indicating measurement results.*

relations limit measurement. Nevertheless, there may seem to be some connection between the inaccuracy and scatter relations. That is, however, not so. A strong argument for this independence thesis is that it is not difficult to construct a theory with a scatter principle, but without an inaccuracy principle (or *vice versa*), e.g. through suitable *ad hoc* restrictions on classical statistical mechanics. Of course such a theory would lack a certain harmony in that it allows measurement to be more precise than preparation (or contains preparations beyond measurement sensitivity). From that point of view the fact that (3) and (76)+(77′) both derive from the same *mathematical* inequality (see appendix), is satisfactory.

At first sight, however, this common mathematical origin of an inaccuracy and a scatter relation appears problematic. Indeed, when we restrict ourselves to covariant joint measurement, it appears that inaccuracy follows from scatter in the measurement device, and that therefore the inaccuracy principle follows from the scatter principle. But the latter reasoning in fact is only valid for joint measurement models working *via* th. 16. Such models have *only* been shown to exist for covariant joint measurements, and these only exist for (analogs of) position-momentum [see § 2.1 and § 2.6]. Similarly the common mathematical origin of the two relations is not fatal for the independence thesis. Relation (C.25), uninterpreted, is of a purely mathematical nature. We can derive both a scatter relation ((3)+interpretation) and an inaccuracy relation ((76)+interpretation) from it, because on finite dimensional spaces there is no difference between trace class and bounded operators. But, as soon as we have

interpreted C in (C.25) as representing a state, i.e. as soon as we have associated it with a preparator, there is no way back. Then we can never get to (76), since to do that we would have to get from a preparator property to a detector property.

The non-ideality inequality (76) does not require anything special for the state of the object system after the measurement: it may even be destroyed. Conversely the scatter relation (3) does not restrict the state of the object system before the preparation: it may be created out of nothing and the preparation need not be associated with any kind of measurement.

Another argument for (76)'s independence of preparation is the fact that the numbers on the left hand side of (76) are the result of a comparison of the determinative aspects of two measurement devices; they characterize the difference between two measuring instruments. This means that (76) refers exclusively to measurement. Preparation, or the state of the object system, does not have anything to do with (76). On the other hand, observables *are* present in scatter relations. Although, as we saw, this does not jeopardize their interpretation as preparational limits, observables do form undesirable elements of scatter relations. It can be shown, for instance, that one can have a statistical theory without limitations on preparation, where certain "scatter" relations still hold:

> Imagine a model consisting of the usual classical statistical mechanics model for 1 particle. As we saw in § 2.3, all measurements of the classical model are non-ideal versions of the observable $\mathfrak{d} = \{\delta_\omega(\Delta\omega')\}_{\mathbb{R}^2}$, corresponding to measuring the phase point itself. We can now map every PFVM $\mathfrak{m}$ into a PFVM $\mathfrak{m}'$ such that $\mathfrak{d}' \longrightarrow \mathfrak{m}'$ has the same non-ideality matrix as $\mathfrak{d} \longrightarrow \mathfrak{m}$. Here $\mathfrak{d}'$ is a non-maximal observable, a joint non-ideal measurement of p and q, such that the scatter in its marginals satisfies a Heisenberg-type relation. Note that the modified model is without incompatible observables. Then, although we have imposed no limitation on preparational possibilities whatsoever, a scatter relation holds: there is no state that is dispersionless with respect to the available measurements.

(Such a model may correspond physically to an interpretation of the scatter principle in terms of the impossibility of perfect measurements in the microdomain, to claiming that particles actually have both p and q values, but we can't accurately measure them.) If we were to construct a similar modified classical model by restricting preparation, we would find that we can in such a way never get a relation like (76) to hold. Whereas scatter relations rely for their definition explicitly on measurement, making above construction possible[xii], the inaccuracy relation does not involve preparation. Consequently the scatter relations are less pure than the inaccuracy relation.

Tentatively, we could reformulate the scatter principle as follows:

(100) It is impossible to prepare a system such that two incompatible degrees of freedom are jointly controllable with arbitrary quality.

Thus stated, it does not require from the outset the presence of observables in the relations representing it. The usual scatter relations, such as the Heisenberg relation (1), are not quite satisfactory representations of the scatter principle in the new formulation. It might be expected that a scatter relation where the quality-numbers arise directly out of a comparison between two preparators does not have the drawback of (implicitly) referring to measurement. For (1) this would mean that the two numbers should characterize how much a given preparator differs from an ideal $\boldsymbol{Q}$ preparator (~ narrow beam) and an ideal $\boldsymbol{P}$ preparator (~ mono-chromatic beam), respectively, without assuming an optimal *measurement* of these observables to be practicable.

As a first step in this direction a refinement of the notion of preparator is necessary. The representation of such a device by a density operator is too unstructured to make a scatter relation of the above type possible. One can think of the analogous case of measurements: there we have effects as unstructured objects and EVMs as the structures built from them [§ 2.2]. Meters are represented by EVMs, and the restriction to yes–no measurements cannot be performed without loss of generality. What we are looking for is therefore something that is related to ρ in a way that is similar to the

way in which the EVM $\{M_k\}_K$ is related to a single effect M_k . One proposal for such a structured notion of preparator is due to Ludwig[87]. He introduces objects of the type

(101) $\{\rho_k\}_K \quad ; \rho_k = \sqrt{\rho}\, M_k \sqrt{\rho}$; ρ a density operator, $\{M_k\}_K$ an EVM;

to represent 'preparator'. We might call this a *trace class operator valued measure*: the parameter k can be integrated over. There does, however, not seem to exist any reason to want this. In an EVM the index represents an *outcome*, something the experimenter does not have any control over. Hence this index is associated with a probability distribution and it must be possible to integrate over it. For a preparator the most natural interpretation for an index is that it represents a "knob", a parameter under the experimenter's control (cf. fig. 2). In this sense the "degree of freedom" of (100) should be interpreted. There is then no reason to associate a probability distribution with this index. This reasoning leads to the conclusion that for our purposes a representation of preparator by *density operator valued function* (DOVF), which is a mapping of $k \in K$ into the set of positive trace class operators with trace one, is appropriate. Examples of such preparators are the parameterized sets of states that have been used in QET, such as (18), which corresponds to a family of q shifted versions of the state ρ. Such a family can, as we have seen, be used quite naturally in the interpretation of 'translation width' [§ 1.2].

It would then be necessary to devise a non-ideality notion for preparators that would allow a partial ordering of these objects and a definition of (in)compatible preparators. Through this notion a conceptually more pure kind of scatter relation in terms of amounts of this 'non-ideality' should be derivable.

NOTES

i The "function" $f(dr,r')$ in (49) is (unlike $g(dr,r')$ in (42)) not necessarily non-negative definite. The inversion (49) symbolizes *deconvolution*. See e.g. D. Champerney (1973): *Fourier Transforms and their Physical Application* (Academic, NY). The inverse mapping is not continuous in any very natural sense [Davies, *op. cit.* (1970)].

---

[87]Ludwig, *op. cit.*, p. 160

ii Relations amounting to (special cases of) $\longrightarrow$ have been used before incidentally. See, e.g., G. Allcock (1969): *Ann. of Phys.* **53**, p. 311; E. Prugovecki, *op. cit.* (1984); Schroeck, *op. cit.* (1989); and further references in § 2.1. Allcock notes the possibility of invertibility (def. 3) explicitly, as does Davies, *op. cit.* (1970).

iii We have here used the $\mathscr{L}^1$-norm to quantify the distinguishability of two probability distributions. Similar results may hold for other distinguishability measures. For instance, for the "statistical distance" [used in (10a)], advocated by Uffink and Hilgevoord [Uffink & Hilgevoord, *op. cit.* (1988); Wootters, *op. cit.*], theorems like the above can be derived from Minkowski's inequality.

iv The diverse aspects of imprecision in physics have been investigated by Ludwig [G. Ludwig (1981): *Structure and approximation in physical theories* (ed. by A. Hartkaemper & H.-J. Schmidt, Plenum, NY), p. 7]. To characterize the imprecision of a measurement he introduces the concept of *imprecision set*. An imprecision set consists of the pairs (ℓ,ℓ') of $\mathfrak{n}$ outcomes that cannot be told apart by the measurement procedure used, i.e. the $\mathfrak{m}$–procedure ($\mathfrak{m}$ and $\mathfrak{n}$ as in the beginning of this section, $\mathfrak{n} \longrightarrow \mathfrak{m}$).

The non-ideality measure we have just introduced is very different from this concept. In order to show that our non-ideality is indeed compatible with Ludwig's we introduce the *discernibility matrix* $\{A_{\ell\ell'}\}$:

$$A_{\ell\ell'} := \inf_{\{\lambda_{k\ell}\}\in\Lambda_{\mathfrak{n}\longrightarrow\mathfrak{m}}} \left\{ \sum_{k\in K} \tfrac{1}{2} \left| \lambda_{k\ell} - \lambda_{k\ell'} \right| \right\}$$

(Λ is defined in (68).) It is derived from the distinguishability of the "probability" distributions the columns of the non-ideality matrix represent. [Instead of the $\mathscr{L}^1$-norm, we might also have taken Wootters' distinguishability measure: Wootters, *op. cit.*; Uffink & Hilgevoord, *op. cit.* (1988)]. The discernibility matrix has the properties:

$$1 \geq A_{\ell\ell'} \geq 0 \quad ; \quad A_{\ell\ell'} = A_{\ell'\ell} \quad ; \quad A_{\ell\ell} = 0 \quad .$$

Note that it includes the labeling of $\mathfrak{n}$, but not that of $\mathfrak{m}$ (cf th. 23). The quantity $A_{\ell\ell'}$ shows how good the non-ideal measurement can discern the $\mathfrak{n}$ outcomes ℓ and ℓ'. Ideally, $A_{\ell\ell'} = 1 - \delta_{\ell\ell'}$.

The discernibility matrix may be seen as a quantification of the concept of imprecision set. Imprecision sets may be derived from the discernibility matrix as the set of pairs (ℓ,ℓ') such that $A_{\ell\ell'} \leq \epsilon$. The threshold ϵ should be a small positive number. Alternatively, $1 - A_{\ell\ell'}$ can be seen as the membership function of the pair (ℓ,ℓ'), defining a *fuzzy* imprecision set.

v As we noted earlier, the matrix $\{\mu_{k\ell}\}$ in $\xrightarrow{\text{i}}$ (def. 3) can be used to calculate the outcome probability distribution of the ideal, if the distribution of our non-ideal measurement is known. If we have a joint invertibly non-ideal measurement of two incompatible observables, this also applies to the marginals of the outcome distribution of a measurement of such an EVM. It is interesting, however, to apply the inverse non-ideality matrices to the joint distribution itself, instead of to the marginals. Suppose we have an EVM O associated with a joint invertibly non-ideal measurement of two non-coexistent EVMs M and N (O, M and N defined as above), such that:

$$\begin{cases} M_\ell = \sum_{m\in M} \mu_{\ell m} \sum_{m'\in M'} O_{mm'} \\ N_k = \sum_{m'\in M'} \nu_{km'} \sum_{m\in M} O_{mm'} \end{cases}$$

Then define the operator valued measure $\{W_{\ell k}\}_{L\times K}$:

$$W_{\ell k} := \sum_{m\in M} \sum_{m'\in M'} \mu_{\ell m}\, \nu_{km'}\, O_{mm'}$$

The distribution this operator valued measure generates, is a Wigner distribution for the two incompatible observables, i.e. a distribution that is not necessarily positive, but has correct marginals. (In fact $\{W_{\ell k}\}$ is a Wigner distribution if both inverse non-ideality matrices obey

$$\sum_{\ell} \mu_{\ell k} = 1 \quad ,$$

but it is not difficult to show that this is always achievable.) We are here confronted with a generalization of the relation that exists between the Husimi distribution and the original Wigner distribution [Hillery *et al.*, *op. cit.*]. This original Wigner distribution is also associated with operator ordering. For the Wigner distributions we get here this holds only in special cases (it can e.g. hold for that of def. 14, for certain choices of ρ_0 [see (89)]).

vi A result similar to (76) has been obtained for $\mathbb{C}^2$, under the assumption of covariant marginals (implying M = L = K), by Busch [Busch, *op. cit.* (1987)], extending a result by Prugovecki [Prugovecki, *op. cit.* (1977)].

vii Many of the results of this section have analogs for position-momentum. [Holevo, *op. cit.* (1982)]. Thus the treatment of this example is also of interest outside this finite dimensional context. In particular can operators be expanded into S_x^b and S_y^c in different orders, just as on the infinite dimensional space operators can be expanded into exp(i*u*P) and exp(i*v*Q). Moreover, these orederings can be shown to correspond to analogs of the Wigner distribution, the Husimi distribution, etc..

viii Such an inaccuracy relation for covariant joint measurements for position-momentum is implicit in Personick, *op. cit.*, and has also been derived by Ali & Prugovecki [S. Ali & E. Prugovecki (1977): *J. Math. Phys.* **18**, p. 219] and by others [see also S. Ali (1985): *Riv. Nuovo Cim.* **8**, p. 1 and references in § 2.1].

ix The possibility of informational completeness has prompted e.g. Busch [Busch, *op. cit.* (1985)] to argue that a matrix $\{\lambda_{km}\}$ should not be interpreted as representing some kind of inaccuracy. But we do not interpret the EVM *itself* as non-ideal. It is the marginals that are non-ideal. The joint measurement EVM can be informationally complete, and it can also be maximal. The EVM as a whole contains "correlations" that give information which cannot be obtained through the marginals alone. For this reason the properties of the marginals should be clearly distinguished from those of the EVM itself. Moreover, an argument like Busch's would hardly affect the operationalization of § 2.3, which ensures the practical meaningfulness of the name 'non-ideality' for $\longrightarrow$.

x Of course the fact that covariance is not to be considered a fundamental factor for the choice of a scale for a measuring instrument, in no way diminishes the importance of group theory in quantum mechanics. The work of e.g. Ludwig (Ludwig, *op. cit.* (1983)) shows that insistence on the unimportance of scales of measurement devices does not mean that group theory should not have a prominent position in quantum mechanics. In fact, group theory shows why hardly anybody has perceived as unsatisfactory the fact that textbooks use (almost) exclusively the von Neumann-Dirac formalism. Practical quantum problems (for the treatment of which these textbooks train the student), such as scattering, can be treated quite well using self-adjoint operators; EVMs are not at all needed. The reason for this is that in such problems "physical quantities" are involved.

All these "quantities" (position, momentum, spin, ...) have a group theoretical background [like the correspondence principle; see H. Primas (1983): *Chemistry, Quantum Mechanics and Reductionism* (2nd ed., Springer, Berlin)], and the generators of the quantum representations of the transformations associated with these "quantities" are self-adjoint operators. But in the treatment of problems like scattering, measurement is *not* involved, and therefore our conclusion regarding the relevance of covariance to the scale of a measuring instrument is not at all related to the usefulness of (self-adjoint operators or) group theory to other problems.

xi This relation would be a generalization of the covariant one of § 2.1 in the same sense in which the application of (76) to the pair of (82) generalizes (88).
Note further that, e.g., the time-of-flight proposal of Park & Margenau, *op. cit.*, is formally a special case of our approach. We can generalize def. 2 to

$$\forall_{\rho\in\Upsilon}\ \mathrm{Tr}(\rho M_k) = \sum_\ell \lambda_{k\ell}\,\mathrm{Tr}(\rho N_\ell)$$

for some subset Υ of the class of all density operators. If we take Υ to be the full class of all density matrices, our def. 2 results. The smaller we take Υ, the weaker this relation becomes. Accordingly a bound in an inaccuracy relation based on this definition will also get weaker as Υ gets smaller. If we take Υ so narrow that it contains only density operators that have equal position and momentum distributions, Park & Margenau's proposal results, including a trivial bound in (98).

xii The translation width w [§ 1.2] has a similar drawback. The above restriction of CM would imply a restriction of the set of ϵ's in (10c). Thus this restriction may lead to an increase in translation width without really affecting preparation possibilities, just as this was the case for scatter. Accordingly, because of the definition of scatter W and translation width w, shift-scatter relations fare no better than scatter relations in this respect, although they are an improvement.

CHAPTER IV

The Uncertainty Principle: Consequences

In the previous chapter we saw that there are two forms of the UP, in correspondence with the fact that there are two types of devices in QM. The inaccuracy principle limits measurement, the scatter principle limits preparation. As these principles are independent, the separation is necessary for a clear understanding of the UP [§ III.3]. But we have as yet not discussed any of the familiar examples and illustrations. The γ–microscope [ch. I], for example, is a meter (and thus not an illustration of the scatter principle), but *joint* measurement is not involved (and it therefore appears not to exemplify the inaccuracy principle). One might be tempted to conclude that there are even more forms of the UP[1]. In the following we shall investigate a number of issues concerning fundamental limits allegedly implied by QM, to see to what extent they are consequences of the UP. We will see that the aforementioned γ–microscope argument is in fact connected to the inaccuracy principle. Diaphragm, "standard quantum limit", Wigner-Araki-Yanase principle and interference versus path are also discussed. Thus the two principles presented in ch. III, are not only necessary but also sufficient for a full representation of the UP.

1 DISTURBANCE INTERPRETATIONS

Now that we have established the two UPs, we shall study several of their applications in the rest of the paper. The first of these concerns one of the most widespread views on the UP, the disturbance interpretation: a position measurement is said to "disturb" the object's momentum in an amount (at least) inversely proportional to its inaccuracy. The formulations to this effect which we find in the work of Heisenberg and others are not substantiated formally: the Heisenberg relation

$$\langle \Delta^2 Q \rangle \, \langle \Delta^2 P \rangle \geq \tfrac{1}{4} \tag{1}$$

does not involve inaccuracy or disturbance, and it has no bearing on this problem [ch. I, III]). Therefore the disturbance interpretation remains to be justified. We

[1]Cf. E. McMullin (1954): *The Principle of Uncertainty* (PhD Thesis, Catholic University of Louvain, Belgium), unpublished

might, for instance, consider whether the amount of σ_1–scatter after a σ_3–measurement is bounded from below, or even increases[2] (σ_1 denotes spin in the 1–direction; analogously for σ_3). But a σ_3–meter, e.g. a Stern-Gerlach type device [§ 5.2], that normally leaves the object approximately in a σ_3 might as well be followed by a magnetic field which turns spin 90 degrees. Then σ_1–scatter would be proportional to σ_3–inaccuracy, rather than inversely proportional. Measurement transformations therefore do not all satisfy a disturbance interpretation in this sense. It seems to be based on the idea that measurements must be intended to approach first kind measurements. But that is by no means always profitable [§ 3]. Accordingly the disturbance interpretation has been claimed to be altogether invalid by, e.g., Kraus[3]. We shall see, however, that the impossibility of certain ways of circumventing the inaccuracy principle and the scatter principle, respectively, allows us to *rigorously* derive disturbance interpretations for *both* incarnations of the UP. In both cases it is important to be precise about the meaning given to 'disturbance', just as it was important to be precise about 'inaccuracy' in the derivation of the inaccuracy principle [§ III.2].

1.1 Disturbance & the Inaccuracy Principle

Joint measurements may be realized in several ways. One way consists of performing two consecutive measurements. For example, a joint $\mathfrak{p}$,$\mathfrak{q}$–measurement may well consist of a non-destructive position meter followed by a meter designed to measure $\mathfrak{p}$. Heisenberg's γ–microscope [ch. I, fig. I.3] is a non-destructive $\mathfrak{q}$–meter. In that experiment we may thus perform another measurement on the outgoing electron in order to find out the momentum the electron had immediately prior to its reaching the microscope. But such a measurement of the electron's initial momentum is hampered by the fact that we do not precisely know how much momentum was transferred in the collision with the photon. The quality of the measurement of initial electron momentum measurement is therefore limited by the recoil momentum indeterminacy.

---

[2]Analogous to a disturbance theory by M. Srinivas (1985): *Pramana* **24**, p. 673

[3]K. Kraus (1987): *Phys. Rev. D* **35**, p. 3070

Note that, as we want to perform a joint p,q–measurement on the particle's original input state, the second meter should be designed to measure the particle's momentum as it was just before it reached the q–meter, rather than to directly measure momentum itself. After all, the position meter may well distort the particle's state such that the momentum information ends up in quite another degree of freedom. [Analogously, in the above mentioned spin example we would need to measure σ_3 to get information on σ_1 before the measurement.] If such a measurement of p–just-*before*-the-q–measurement ("undisturbed p") would be possible with unlimited accuracy, the joint p,q–measurement thus realized could violate the inaccuracy principle. A QM measurement cannot do that, however. Therefore a non-ideal q–meter n with inaccuracy $\delta_{q\to n}$ must destroy p–information in an amount at least inversely proportional to $\delta_{q\to n}$. Heisenberg's γ–microscope, as it was just discussed, illustrates this[i]. A more realistic example will be treated in § 5.1

We shall now justify this *disturbance interpretation* for the finite dimensional case we dealt with in ch. III. First we must extend the notion of state transformation accompanying a measurement. This transformation is traditionally assumed to be a measurement of the first kind [ch. I]. The simplest kind of device involving the influence of measurement on the object is undoubtedly a filter, a device that selects quantum systems. It is represented in the extended formalism by an *operation*[4]. An operation Ξ is a linear mapping of the set of trace class operators into the set of trace class operators satisfying[ii]

$$\text{(2a)} \qquad \forall_{\rho \mid \rho \geq 0 \wedge \mathrm{Tr}(\rho) = 1} \quad \Xi[\rho] \geq \mathbf{0} \ ;$$

$$\text{(2b)} \qquad \forall_{\rho \mid \rho \geq 0 \wedge \mathrm{Tr}(\rho) = 1} \quad \mathrm{Tr}(\Xi[\rho]) \leq 1 \ .$$

The norm $\mathrm{Tr}(\Xi[\rho])$ corresponds to the probability of the system passing the filter, defining an effect $\boldsymbol{M}$

$$\text{(3)} \qquad \forall_{\rho} \ \mathrm{Tr}(\rho \boldsymbol{M}) = \mathrm{Tr}(\Xi[\rho]) \quad \Leftrightarrow \quad \boldsymbol{M} = \Xi^{\dagger}[\mathbf{1}] \ .$$

---

[4]E. Davies (1976): *Quantum Theory of Open Systems* (Academic, London); G. Ludwig (1983): *Foundations of Quantum Mechanics*, vol. I (Springer, Berlin); K. Kraus (1983): *States, effects and operations* (Lecture Notes in Physics **190**; Springer, Berlin)

If it has passed the filter, the system is in a state $\Xi[\rho]/\mathrm{Tr}(\Xi[\rho])$.
Note that there are many operations corresponding to a given effect. Two examples are

$$\Xi[\rho] = M^{\frac{1}{2}}\rho M^{\frac{1}{2}} \quad , \quad \Xi[\rho] = \mathrm{Tr}(\rho M)\,\rho_f \quad (\rho_f \text{ fixed}),$$

both corresponding to the effect M. For a measuring device with discrete outcome set, each possible outcome corresponds to a filter: we can always select or ignore a system depending on whether it generated a particular outcome in the measuring device. Thus an instrument is represented by an *operation valued measure* (OVM). An OVM is, for a discrete outcome set M, a family $\mathfrak{N} = \{\nu_m\}_M$ of operations, satisfying

$$(4) \qquad \forall_{\rho \mid \rho \geq 0 \wedge \mathrm{Tr}(\rho) = 1} \quad \mathrm{Tr}(\nu_M[\rho]) = 1 \quad \Leftrightarrow \quad \nu_M^\dagger[1] = 1$$

with:

$$(5) \qquad \nu_M := \sum_{m \in M} \nu_m \quad .$$

When a measurement is performed on an object initially in state ρ, the object's state after measurement is, conditional on the measurement outcome m, given by $\nu_m[\rho]$. The norm $\mathrm{Tr}(\nu_m[\rho])$ of this state is the probability of outcome m.

Accordingly the EVM $n = \{N_m\}_M$ representing the determinative aspect of the instrument associated with $\mathfrak{N}$, is uniquely defined by:

$$(6) \qquad \forall_\rho \ \mathrm{Tr}(\rho N_m) = \mathrm{Tr}(\nu_m[\rho]) \quad \Leftrightarrow \quad N_m = \nu_m^\dagger[1] \quad .$$

We shall denote the EVM that corresponds to a given OVM by the corresponding lower case character. Just as an EVM is a generalization of a PVM, an OVM is a generalization of a measurement of the first kind.

After the object system has left the $\mathfrak{N}$–meter, we may perform other measurements on it. Suppose we apply an apparatus realizing the EVM $o = \{O_\ell\}_L$ to the object after $\mathfrak{N}$. Then the probability of finding o–outcome ℓ, given that we have found m in the $\mathfrak{N}$–measurement, is $\mathrm{Tr}(\{\nu_m[\rho]/\mathrm{Tr}(\nu_m[\rho])\}\, O_\ell)$. Obviously the joint m,$\ell$-distribution

is given by $\mathrm{Tr}(\nu_m[\rho]\, O_\ell) = \mathrm{Tr}(\rho\, \nu_m^\dagger[O_\ell])$. In other words, this experiment consisting of two consecutive measurements may also be seen as a realization of the bivariate EVM $\{\nu_m^\dagger[O_\ell]\}$.

If we do not differentiate systems leaving $\mathfrak{N}$ according to the outcome they engendered, we may represent the outgoing state by $\nu_M[\rho]$. This implies in particular that the o–marginal of the just mentioned joint measurement is given by $\nu_M^\dagger[o]$. Since ν_M is a *non-selective operation* (i.e. it satisfies $\nu_M^\dagger[1] = 1$), $\nu_M^\dagger[o]$ is an EVM. If o was performed to find out something about some PVM f as it was before the $\mathfrak{N}$ device, we should require $f \longrightarrow \nu_M^\dagger[o]$ rather than simply $f \longrightarrow o$: in the latter case the marginal $\nu_M^\dagger[o]$ could not be expected to be at all related to f. This means physically that the o meter should be designed so as to allow for the action of $\mathfrak{N}$ on the object state as much as possible. The extent to which such a compensation is possible, is given by the amount of non-ideality in the relation $f \longrightarrow \nu_M[o]$ for an optimal choice of o. This is how *disturbance* will be concretized.

Clearly the above discussion only depends on $\nu_M^\dagger[o]$ being an EVM, and thus on $\nu_M^\dagger$ being a non-selective operation. Accordingly, we formulate the disturbance notion for an *arbitrary* non-selective operation Ξ. We define the amount of f disturbance in a device that realizes an operation Ξ, as

$$(7) \qquad J_\Xi^{\delta f} := \inf_{o\,|\,f \longrightarrow \Xi^\dagger[o]} \left(J_{f \longrightarrow \Xi^\dagger[o]} \right) .$$

An application of (7) shows the amount of f disturbance in an operation corresponding to a unitary evolution to be 0. Indeed this is precisely what one would desire: normal Schrödinger evolution does not "disturb" anything. A meter, on the other hand, represented by $\mathfrak{N}$, disturbs f by an amount[iii] $J_{\nu_M}^{\delta f}$. This somewhat involved notion allows us to circumvent the aforementioned objections of Kraus[5] to a disturbance interpretation: unlike Srinivas[6] we have made no assumption about the nature of the measurement transformation. We have in particular not assumed that the OVM approximates a measurement of the first kind. This is necessary because no general

[5]Kraus, *op. cit.* (1987)

[6]Srinivas, *op. cit.*

assertions about quantum measurement processes can be based on a consideration of first kind measurements, as these can at best be seen as descriptions of idealized measurement transformations [ch. I]. A limit for the $\mathfrak{f}$ disturbance in a non-ideal $\mathfrak{e}$-measurement ($\mathfrak{e}$, $\mathfrak{f}$ arbitrary PVM's on $\mathbb{C}^n$) can then be derived from (III.76):

$$(8) \qquad J^{\delta\mathfrak{f}}_{\nu_M} + J_{\mathfrak{e}\longrightarrow n} \geq \sum_i c_i \operatorname{Tr}(G_i) \quad .$$

Inequality (8) relates two properties of a non-ideal $\mathfrak{e}$–meter: both $J^{\delta\mathfrak{f}}_{\nu_M}$ and $J_{\mathfrak{e}\longrightarrow n}$ are functionals of the OVM $\mathfrak{N}$ alone. The $\mathfrak{N}$–device does not measure $\mathfrak{f}$, nor does it in any way require an $\mathfrak{f}$–meter to be present. Ineq. (8) refers to the implication "*If* we place after $\mathfrak{N}$ a device that measures $\mathfrak{f}$–before–$\mathfrak{N}$, *then* this device cannot be arbitrarily accurate", without assuming the validity of its premise.

An interesting question is whether it is possible to process the object such as to undo the state change brought about by the measurement, i.e. whether there is an operation χ such that $\chi[\nu_M[\rho]] = \rho$ for all ρ. It is not possible (if n is non-minimal), as follows directly from the inaccuracy principle: if it were possible to completely correct the disturbance due to some measurement, it would be possible to measure accurately the *undisturbed* "value" of some incompatible observable. This violates the inaccuracy principle. Hence, in any theory encompassing an inaccuracy principle, some information must be destroyed in the measurement of non-classical quantities, and conversely non-disturbing measurements are only possible for classical quantities.

1.2 Disturbance & the Scatter Principle

Imagine a situation where you want to hit a faraway target using quantum particles. By suitably designing the source that emits the particles, and by careful operation of its controls, this aim can be approximated arbitrarily well. Now we put a diaphragm halfway between the source and the target. The result is that diffraction effects prevent the perfect directability of the beam. No matter how sophisticated our source is, or how skillfully we operate it, some particles will always miss the target. Thus

the diaphragm may be said to "disturb" momentum: no matter what source we place before the diaphragm, we can never fully control the object's momentum after it. More quantitatively, the controllability of post-diaphragm momentum is reciprocally related to the slit's width. this type of disturbance must be present in the diaphragm in order to avoid the possibility of using it to violate the scatter principle.

This reasoning is quite analogous to that in the previous section (cf. fig. 1). The diaphragm is a preparational device that changes the particle state, an analog of a non-destructive meter. Then, just as a q–meter limits the *possibilities* of ascertaining the value of p just before the q–measurement when this measurement has taken place, a diaphragm limits the possibilities of *controlling*[7] $\boldsymbol{P}$ just *after* the diaphragm with a preparator placed *before* it.

The SLP relation [§ III.1.1] can be used to substantiate the disturbance interpretation of the scatter principle mathematically. Define position projector $\boldsymbol{B}$ as in (III.6). The diaphragm is symbolized by an operation $\Theta[.]$ (§ 1.1). We assume that the diaphragm absorbs all particles outside an interval of width Dq (it is perfect as a $\boldsymbol{B}$–filter):

$$(9) \qquad \forall_{\rho \mid \rho \geq 0 \wedge \mathrm{Tr}(\rho) = 1} \ \Theta[\rho] \leq \boldsymbol{B}$$

$$\Downarrow$$

$$\forall_{\rho \mid \rho \geq 0} \forall_{\alpha} \ W_{Q,\alpha}(\Theta[\rho]/\mathrm{Tr}(\Theta[\rho])) \leq Dq \quad .$$

A diaphragm is not a non-selective operation, and correspondingly $\Theta^{\dagger}[1] \neq 1$. Therefore the normalization in (9) is needed. Now, the overall width W of (III.9) inspires the following disturbance measure:

$$(10) \qquad W^{\boldsymbol{P}}_{\Theta,\beta} := \inf_{\rho \mid \rho \geq 0} \left(W_{P,\beta}(\Theta[\rho]/\mathrm{Tr}(\Theta[\rho])) \right) \quad .$$

Then we can check by substituting the state $\Theta[\rho]$ in the SLP relation (III.8), and using the right hand part of (9), that

$$(11) \qquad W^{\boldsymbol{P}}_{\Theta,\beta} Dq \geq 4 \lambda^{-1}(\beta^2) \quad .$$

---

[7]Note that we use $\boldsymbol{P}$ to denote momentum in a preparative context, whereas p is used whenever momentum measurement is involved (cf. § III.3).

This relation can be seen as a preparative analog of inequality (8). It can be straightforwardly extended to non-perfect diaphragms[iv]. We must stress that (11) has no bearing on measurement. Both $W^{\delta P}_{\Theta,\beta}$ and Dq are properties of the diaphragm, which is not a meter at all. Moreover, they are not object properties.

Above discussion applied a preparative disturbance concept in the context of a diaphragm experiment. A similar application is possible in a multiple slit situation. Consider the periodical array of slits of fig. II.2. As we saw [§ III.1.1], the Hilgevoord-Uffink relation (III.11) is suitable for a description of this experiment. If we, like above, represent the device's action by an operation Θ, we can again use (10). Instead of (9), however, we characterize the multiple slit by

$$(12) \qquad \forall_{\rho \mid \rho \geq 0} \forall_{\alpha} \ \tilde{w}_{Q,\alpha}(\Theta[\rho]/\mathrm{Tr}(\Theta[\rho])) \leq a \ .$$

Then we can straightforwardly derive from (III.11) the inequality

$$(13) \qquad W^{\delta P}_{\Theta,\beta} a \geq C(0,\beta) \quad (\beta^2 \geq \tfrac{1}{4}).$$

Again, both $W^{\delta P}_{\Theta,\beta}$ and a are properties of the slit device, and not of the object.

1.3 Disturbance in Amplifiers

'Disturbance' is sometimes connected with the UP also in another context. It has been claimed that the UP implies that an amplifier of photon number must disturb phase[8].

> Consider a joint phase-number meter that is optimally accurate: $\Delta n \Delta \phi = \frac{1}{2}$. Then we can construct a new device by placing the amplifier in front of the meter. This new meter is more accurate as regards photon number than the original one: the inaccuracy decreases according to $\Delta n \rightarrow \Delta n/G$ (the *gain* $G > 1$). If we assume that the amplifier leaves phase unaffected, the meter's phase inaccuracy remains unchanged

[8]H. Heffner (1962): *Proc. IRE* 50, p. 1604 [reprinted in J. Wheeler & W. Zurek (eds.) (1983): *Quantum Theory and Measurement* (Princeton University Press)]

($\Delta\phi \longrightarrow \Delta\phi$). Since the new inaccuracies violate the uncertainty principle, it is concluded that there has to be some kind of noise in a quantum amplifier for fundamental reasons, increasing $\Delta\phi$.

Of course Heffner based his argument on a relation of type (1). But $\Delta\phi$ and Δn in this argument are both intended as inaccuracies, not as scatters. In the amplification process the photon number scatter *increases* [$\langle\Delta^2 N\rangle \longrightarrow G^2\langle\Delta^2 N\rangle$], in contrast to the measurement inaccuracy. Therefore the conventional uncertainty principle, represented by inequalities like (1), is not endangered at all, and there is no reason for any amplifier noise on the basis of above reasoning. The inaccuracy principle would seem to be more appropriate, and we shall see here whether we can indeed use it. We shall discuss the position-momentum case. Again, it is important to be be precise about what is meant by 'amplifier'. Consider an operation χ, defined by

$$(14) \qquad \chi{:}\rho_i \longrightarrow \rho_f = T\rho_i T^\dagger \ ,$$

with:

$$T := G^{\frac{1}{2}} \int dx \ |Gx\rangle\langle x| \ .$$

Here $G > 1$ is the amplification factor. As $\chi^\dagger[|x\rangle\langle x|\,dx] = |\frac{x}{G}\rangle\langle\frac{x}{G}|\,d\frac{x}{G}$, the operation χ satisfies

$$(15) \qquad \langle X\rangle_f = G\langle X\rangle_i \ , \ \langle\Delta^2 X\rangle_f = G^2\langle\Delta^2 X\rangle_i \ .$$

Moreover, it indeed works as a magnifying glass, increasing the accuracy of a following position meter:

$$(15') \qquad \epsilon_{\xi\longrightarrow\Xi^\dagger[m]} = \frac{1}{G}\epsilon_{\xi\longrightarrow m} \text{ for all m such that } \xi \longrightarrow m \ .$$

But, as χ is a unitary transformation, it involves no disturbance in the sense of either § 1.1 or § 1.2. A χ device "attenuates" momentum noiselessly[9].

Thus, *if* we characterize 'amplifier' by (15) or (15'), the claim that a position amplifier *must* disturb momentum, can only be justified if 'disturbance' is given a much weaker content than in § 1.1 and § 1.2. We may, however, also restrict the definition

[9] C. Caves (1981): *Phys. Rev. D* **23**, p. 1693

of 'amplifier', so as to render this notion inapplicable to (14). Consider, for instance, supplementing the characterization by an 'unbiasedness' requirement, viz.

$$(16) \qquad \forall_{\rho_i} \ \langle P \rangle_f = \langle P \rangle_i \ \Leftrightarrow \ \chi^\dagger[P] = P \ .$$

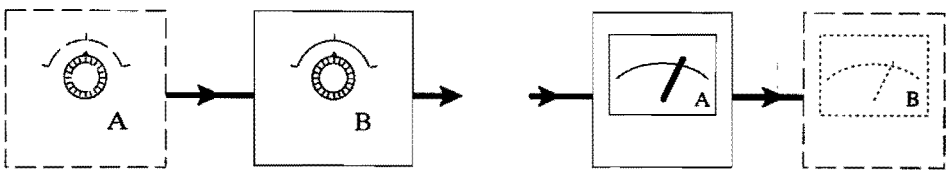
This is sufficient to disqualify (39). Such requirements do, however, not seem to be inherent in the concept of 'amplifier'. Moreover, it is unclear how (even if (14) is excluded) a violation of the inaccuracy principle is to be brought about. Thus Heffner-type reasoning appears unsuitable to establish quantum amplifier constraints. We may also look at the case where *both* position and momentum are amplified. Model calculations[10] indicate that there is in such amplifiers noise of the type we called 'disturbance' in § 1.1.

1.4 Discussion

In fig. 1 the two disturbance interpretations are symbolized. The fact that the first preparator *c.q.* the second meter need not actually be present for the disturbance notions [cf. (7) and (10)] to make sense, may be expressed by speaking of "counterfactual" scatter *c.q.* inaccuracy. One may wonder whether measurement does not also lead to a preparative disturbance of type (10). This is not necessitated by any fundamental law, however[11]. We can devise an OVM that realizes a good position measurement and leaves the object in a state with sharp momentum [See the spin example in the opening of § 1]. More generally, it can be seen that there is no difference as regards possibilities of preparing $\boldsymbol{Q}$ or $\boldsymbol{P}$ after a q-measurement. Similarly there is no difference in the possibilities as regards measurement of undisturbed [in the sense of § 1.1] q and p after a diaphragm-like device. Hence there is *no* reciprocal relation between q–measurement accuracy and $\boldsymbol{P}$–scatter after measurement

[10]H. Haus & J. Mullen (1962): *Phys. Rev.* **128**, p. 2407; Y. Yamamoto & H. Haus (1986): *Rev. Mod. Phys.* **58**, p. 1001

[11]Kraus, *op. cit.* (1987)



(a) *(b)*

fig. 1 *Symbolic representation of the two types of disturbance interpretations of the uncertainty principle:*
a) preparative disturbance limits the possibilities of controlling post-preparator quantities through pre-preparator manipulations on the object;
b) determinative disturbance limits the possibilities of future realization of pre-measurement observables.
The dashed devices indicate the "counterfactual" character of these interpretations.

(***P***–controllability), and there is no reciprocal relation between diaphragm width and possible p–measurement inaccuracy. There are no scatter-inaccuracy uncertainty relations.

The fact that there is no law from which scatter-inaccuracy complementarity as a property for *all* devices can be derived, does not imply that there are *no* devices for which it holds. If an OVM satisfies, e.g.,

(17) $\nu_M^\dagger[\mathfrak{f}] \leftarrow \mathfrak{f}$,

then ($\mathfrak{f} = \{F_k\}_K$)

$$\nu_M^\dagger[F_k] = \sum_{k' \in K} \lambda_{kk'} F_{k'} \quad ,$$

so that

(18) $$\mathrm{Tr}(\nu_M[\rho]\, F_k) = \sum_{k' \in K} \lambda_{kk'} \mathrm{Tr}(\rho F_{k'}) \quad .$$

The ***Y***–distribution ($Y = \sum_{k \in K} k F_k$) after measurement is a smeared version of the ***Y***–distribution before measurement. Therefore, *if* a non-ideal e–meter satisfies (17),

we can speak of a preparative Y-disturbance that satisfies an inequality like (8). Indeed it is quite plausible that the action of the γ-microscope is such that it satisfies (17), that there is a preparative disturbance reciprocally related to its resolution. But the requirement (17) is not necessitated by any QM law, so that the γ-microscope is in this respect not representative of QM measurement devices in general [see further § 5].

2 WIGNER–ARAKI–YANASE RESTRICTIONS ON MEASUREMENT

In the previous section we demonstrated that the *possibility* of performing a joint measurement by doing two consecutive measurements, leads to a limitation on the working of non-destructive meters. Analogously, we show in this section how the inaccuracy principle affects a scheme where conservation laws are utilized to effect a joint measurement. In this way we can reinterpret the limitations on the accuracy of measurements in the presence of conservation laws, like those derived by Wigner[12], Araki and Yanase[13] (WAY) and by others[14], as a *consequence* of the UP rather than as a restriction of quantum measurements *in addition to* the UP. Our derivation is not based on a particular model or on an analysis of the details of the measuring process.

The general scheme in this section is that of an object $\mathcal{O}$ interacting with a device $\mathcal{A}$ such that some operator L'' of the total system commutes with the Hamiltonian. In the following we shall use the law of conservation of momentum. The results can, however, be readily generalized to cases where some other (possibly non-additive) quantity $L'' = f(L,L')$ is conserved.

---

[12]E. Wigner (1952): *Zs. f. Phys.* **133**, p. 101

[13]H. Araki & M. Yanase (1960): *Phys. Rev.* **120**, p. 622; M. Yanase (1961): *Phys. Rev.* **123**, p. 666

[14]E.g. P. Busch (1985): *J. Phys. A* **18**, p. 3351; M. Ozawa (1990): "Does a Conservation Law Limit Position Measurement ?", Harvard University preprint HUTP-90/B002; A. Shimony & H. Stein (1979): *Am. Math. Mon.* **86**, p.292

2.1 WAY Interpretation of the Inaccuracy Principle

Consider the case where the apparatus A is a position meter with EVM $\mathfrak{a}$ and accuracy $\delta_{\mathfrak{q}\rightarrow\mathfrak{a}}$. Suppose the momentum distribution in the initial state of the apparatus to be known. We *may* then measure, after the $\mathfrak{a}$–measurement is completed, the momentum of the total system. Because momentum is conserved, and we know the initial momentum distribution of the apparatus, we can use the results of this last measurement to gain information about the momentum of the object before the $\mathfrak{a}$–measurement. If the distribution of the initial A–momentum is sharp, we can find out the object's pre-measurement momentum with the same sharpness. More generally, we are dealing with a non-ideal O–momentum measurement procedure. The accuracy of this indirect momentum measurement is characterized by $\langle\Delta^2 P'\rangle_i$: the initial apparatus momentum dispersion[15]. Because the inaccuracy relation (III.98) cannot be violated in this way, a relation like

$$\langle\Delta^2 P'\rangle_i \, \delta_{\mathfrak{q}\rightarrow\mathfrak{a}} \geq 1 \tag{19}$$

holds, its precise form depending on the definition of $\delta_{\mathfrak{q}\rightarrow\mathfrak{a}}$. In other words, a small $\delta_{\mathfrak{q}\rightarrow\mathfrak{a}}$ requires that $\langle\Delta^2 P'\rangle_i$ be big, so that momentum conservation is of little use in the determination of initial object momentum.

We can easily devise a model illustrating (19). Imagine that A is constituted of two sub-systems: an ancilla B and a meter C that measures $q - q_b$ ideally (B– and C–observables are denoted by corresponding indices). The operator $Q - Q_b$, corresponding to this observable, commutes with total momentum $P + P_b + P_c$ so that there are no restrictions to the precision with which $q - q_b$ can be measured. A suitable $(O+B)\leftrightarrow C$ interaction Hamiltonian can easily be given[16]: we may take it to be a function of $Q - Q_b$ and some C–operator commuting with P_c [cf. § 5]. The initial

[15]Cf. S. Personick (1971): *Bell Syst. Techn. J.* **50**, p. 213 and § III.2.6

[16]Ozawa, *op. cit.*; D. Bohm (1951): *Quantum Theory* (Prentice Hall, Englewood Cliffs NJ), ch. 22

apparatus state ρ' is known. A $q - q_b$ measurement can be regarded as a non-ideal measurement of $\mathcal{O}$–position q with non-ideality determined by the initial scatter in Q_b [ch. III]. Thus, using (96),

$$(\epsilon^{(Q)}_{q\to a})^2 = \langle \Delta^2 Q_b \rangle_i \geq \tfrac{1}{4} / \langle \Delta^2 (P_b + P_c) \rangle_i$$

and relation (19) is satisfied. If C effects an non-ideal $q - q_b$ measurement, its non-ideality as a q–meter can only increase.

2.2 WAY Interpretation of the Scatter Principle

In order to make explicit that there is a scatter analog of the reasoning leading to (19), despite the fundamental differences between scatter and inaccuracy principle, we add the following application of the scatter principle. Of course the result can also be derived straightforwardly from the Robertson relation (III.2) using momentum conservation $[(P + P')_i = (P + P')_f]$.

Consider a source $\mathcal{A}$ emitting objects $\mathcal{O}$, with position dispersion $\langle \Delta^2 Q \rangle_f$. Again total momentum is conserved, this time in the preparation process. Assume that the initial $\mathcal{O}+\mathcal{A}$ momentum distribution is known. We *may* then measure, after the preparation is completed, the $\mathcal{A}$–momentum accurately. We can then select objects with a certain momentum by using the $\mathcal{A}$–momentum result as a criterion, since knowing $\mathcal{A}$–momentum combined with knowledge of the initial $\mathcal{O}+\mathcal{A}$ momentum can be used to predict $\mathcal{O}$–momentum. If the pre-preparation $\mathcal{O}+\mathcal{A}$ momentum distribution was sharp, the final $\mathcal{O}$–momentum distribution thus obtained will also be sharp. More generally, it will have a dispersion equal to $\langle \Delta^2 (P + P') \rangle_i$. The scatter principle must hold in the $\mathcal{O}$–state, conditional on the measured $\mathcal{A}$–momentum value. Because the measurement on $\mathcal{A}$ did not directly influence the object, it cannot have changed the average variance $\langle \Delta^2 Q \rangle_f$. This leads to

$$\langle \Delta^2 (P + P') \rangle_i \, \langle \Delta^2 Q \rangle_f \geq \tfrac{1}{4} \quad . \tag{20}$$

3 TRAJECTORY & SQL

Yet another intuitively plausible consequence of the uncertainty principle is the ban quantum mechanics imposes on 'trajectory'. Indeed the apparent inability of quantum mechanics to cope with Wilson chamber tracks motivated Heisenberg to write his 1927 paper[17]. More recently, this issue has been reopened as a consequence of the ever-increasing demand for accuracy in gravitational wave detection experiments[18].

In principle position monitoring involves continuous measurement[19]. But since that is mathematically not easy to model[20], we (like most others) will restrict ourselves to two successive position measurements. Consider a mass m moving freely between two position measurements that are separated by a time interval τ (fig. 2). The interaction of the mass with the first meter is assumed to be finished by $t=0$, whereas the second one does not interact with the mass until after $t=\tau$. The model can also occur in the form of *one* device making repeated measurements on the same system. Then:

$$(21) \qquad Q \rightarrow Q(\tau) := U_{0\rightarrow\tau}^{\dagger} Q U_{0\rightarrow\tau} = Q + \frac{\tau}{m} P \ .$$

Here $U_{0\rightarrow\tau}$ generates the system's evolution in time. For this case we have the scatter relation[21]

$$(22) \qquad \langle \Delta^2 Q(\tau) \rangle \langle \Delta^2 Q \rangle \geq \left[\frac{\tau}{2m}\right]^2 \ .$$

---

[17]W. Heisenberg (1927): *Zs. f. Phys.* **43**, p. 172 [See also W. Heisenberg (1969): *Der Teil und das Ganze* (Piper, Munich).]

[18]R. Bondurant (1986): *Phys. Rev. A* **34**, p. 3927; V. Braginskii (1988): *Sov. Phys. Usp.* **31**, p. 836

[19]See e.g. G. Prosperi (1984): in *Quantum probability and applications to the quantum theory of irreversible processes* (ed. by L. Accardi, A. Frigerio & V. Gorini; Lecture Notes in Mathematics **1055**, Springer, Berlin), p. 301; V. Belavkin (1989): *Phys. Lett. A* **140**, p. 355, 359; M. Ozawa (1989): *Squeezed & non-classical light* (ed. by P. Tombesi & R. Pike; NATO ASI B **190**; Plenum, NY), p. 263

[20]But cf. W. Edelstein, J. Hough, J. Pugh & W. Martin (1978): *J. Phys. E* **11**, p. 710

[21]J.-M. Levy-Leblond (1972): *Am. J. Phys.* **40**, p. 899

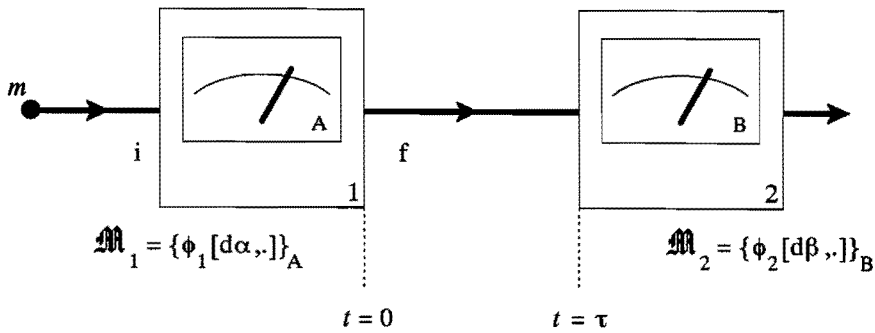


fig. 2 *Symbolic representation of measurement setup. Two consecutive non-ideal q-measurements are performed. The OVM's of the devices are indicated.*

As in the Heffner reasoning [§ 1.3], the Δ's in (22) are often not distinguished from measurement inaccuracies, so that (22) degenerates into

(22') $$\delta_{q(\tau)}\,\delta_q \geq \frac{\tau}{2m} \quad .$$

The further (natural) assumption that the two meters are identical implies $\delta_{q(\tau)} = \delta_q$. Thus

(23) $$(\delta_q)^2 \geq \frac{\tau}{2m} \quad .$$

This absolute lower limit to position tracking inaccuracy, here somewhat caricaturally presented, is called the "Standard Quantum Limit" (SQL) for position monitoring. In above form, its derivation is untenable. There is no rigorous connection between scatter and inaccuracy. Ineq. (22) cannot be used for statements about measurement accuracy [ch. I, III]. There has been a debate about whether[22] or not[23] more subtle arguments can prove an SQL.

---

[22]C. Caves (1985): *Phys. Rev. Lett.* **54**, p. 2465; Bondurant, *op. cit.*; Braginskii, *op. cit.*; Caves, *op. cit.* (1981)

[23]H. Yuen (1983): *Phys. Rev. Lett.* **51**, p. 719; (1984): *ibid.* **52**, p. 1730; S. Tsyplyaev (1989): *Sov. Phys. J.* **31**, p. 699; Ozawa, *op. cit.* (1989)

As mentioned above, the SQL discussion was instigated by the need for efficient weak force detection schemes. If we want to detect the presence of such a force, we may look at the deviations of measurement outcomes from their undisturbed values. This works best when the scatter in the outcome distributions is small. We therefore want the scatter in the second meter's outcomes to be small, in order to be able to detect the presence of a disturbing force during the time τ. Caves[24] assumes that the inaccuracy of the first measurement $\epsilon_{q\rightarrow m_1}$ [see (14)] is at least as large as the X scatter in the outgoing state:

$$(24) \qquad \forall_{\rho_i} \; (\epsilon_{q\rightarrow m_1})^2 \geq \langle \Delta^2 Q \rangle_f \; .$$

On the basis of (24) and the assumption of covariance of the meters [ch. III], Caves derives a lower bound on the scatter in the results of the second meter $\langle \Delta^2 \beta \rangle$:

$$\langle \Delta^2 \beta \rangle = \langle \Delta^2 Q(\tau) \rangle_f + (\epsilon_{q\rightarrow m_2})^2 =$$

$$= \langle \Delta^2 Q(\tau) \rangle_f + (\epsilon_{q\rightarrow m_1})^2 \geq \langle \Delta^2 Q(\tau) \rangle_f + \langle \Delta^2 Q \rangle_f \geq \frac{\tau}{m} \; .$$

Thus he has derived an upper bound for weak force detection efficiency. Unfortunately, as noted by Ozawa[25], Caves' assumption (24) need not alwys hold; there is no general relation between measurement accuracy and scatter after measurement [§ 1]. Indeed measurements have been proposed [e.g. by Yuen and by Tsyplyaev[26]] in which Caves' assumption (24) is violated.

That there is no fundamental limit on weak force detection efficiency, is indicated by the following reasoning. Consider an extremely schematic model (fig. 3), where we have a preparator that leaves the object in state ρ_f at $t=0$. The object is allowed to evolve until $t=\tau$ according to a Hamiltonian, which in the absence of the force is given by H. Now, a natural way to detect a force is through deviations in the

---

[24]Caves, *op. cit.* (1985)

[25]Ozawa, *op. cit.* (1989)

[26]Yuen, *op. cit.* (1983,1984); Tsyplyaev, *op. cit.*; Ozawa, *op. cit.* (1989)

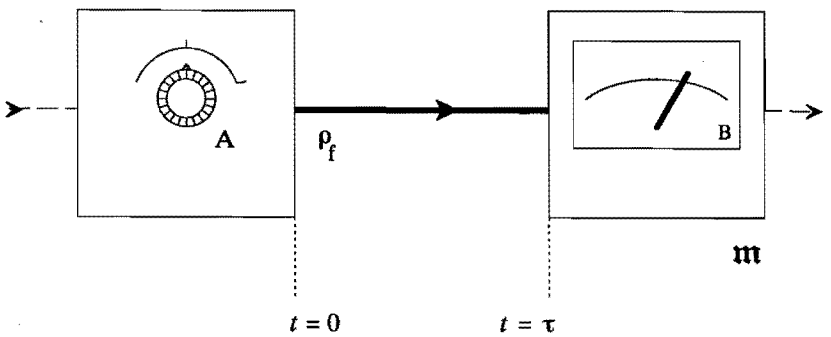


fig. 3 *Symbolic representation of measurement setup for weak force detection. A meter measures* $\mathfrak{m} \leftarrow \mathfrak{c}$ ($\mathfrak{c}$ *is the PVM of* C) *and a preparator prepares the particles in state* ρ_f *such as to have a small* $\langle \Delta^2 C(\tau) \rangle_f$.

expectation value of some operator, say C (with PVM $\mathfrak{c}$). Suppose the disturbed Hamiltonian is given by $H' = H + \eta V$. The term ηV represents the force. Then, in lowest order:

$$\langle C'(\tau) \rangle_f \simeq \langle C(\tau) \rangle_f + i\tau\eta \langle [C,V]_- \rangle_f \quad . \tag{25}$$

Eq. (25) indicates that, if C is chosen such that the second term does not vanish, the detection efficiency is determined by the measurement accuracy $\delta_{\mathfrak{c}(\tau) \rightarrow \mathfrak{m}(\tau)}$ and the pre-measurement scatter $\langle \Delta^2 C(\tau) \rangle_f$. Both $\delta_{\mathfrak{c}(\tau) \rightarrow \mathfrak{m}(\tau)}$ and $\langle \Delta^2 C(\tau) \rangle_f$ can be made arbitrarily small. Note that for this type argument it is not really relevant whether the first device is a meter, or whether the second is non-destructive. We have only taken into account the preparator-side of the first and the detector-side of the second. Corresponding to this, in this argument only the state ρ_f and the EVM $\mathfrak{m}$ were involved. If we want to use a scheme in which repeated measurements on the same object are made (analogous to fig. 2), the meter should be constructed so that the C scatter at the time τ' of the next measurement on the object is small. Hence, it should generate a post-measurement state with sharp $C(\tau')$,rather than one that gives an identical value in an *immediately* repeated measurement [i.e. a state with sharp $C(\tau)$]. Since the latter condition is characteristic of a measurement of the first kind, and in general incompatible with the condition that $C(\tau')$ is sharp, we have here a case

where we do *not* want to approach a measurement of the first kind. A measurement of the first kind is not the optimal apparatus for weak force detection.

In Yuen-Tsyplyaev type schemes the first meter, an inaccurate position meter, does not leave the object in a state such that $\langle \Delta^2 Q \rangle_f$ is small, but is cleverly chosen so as to leave the object in a state such that $\langle \Delta^2 Q(\tau) \rangle_f$ is small, analogously to the above reasoning. Then the results of the second meter may show little dispersion, enabling efficient weak force detection. The OVM of the proposals by Yuen and by Tsyplyaev represents a non-destructive device, fundamentally differing from a first kind measurement.

But can such schemes really be considered to *monitor* position ? The smallness of the variance in the results of the second measurement that arises in these schemes is to a large extent a creation of the first measurement. What we are looking at is more a track that we made ourselves than it is the "undisturbed" track: we are *controlling* the object's position rather than *measuring* it. In other words, true 'position monitoring' seems to require an interpretation of the sequential measurements of fig. 2 as one *joint* measurement[27] of the observables q and $q(\tau)$ on the input state ρ_i. Until now we assumed implicitly that the first meter measures q on ρ_i non-ideally, and that the second one measures non-ideally q on $\rho_f(\tau)$ [$\Leftrightarrow$ $q(\tau)$ on ρ_f].

We shall now formalize this reasoning, using the technique of § 1.1. The first meter has OVM $\mathfrak{M}_1 = \{\phi_1[d\alpha,.]\}_A$ on the field $\mathcal{F}_A$ on outcome set A. The second meter has OVM $\mathfrak{M}_2 = \{\phi_2[d\beta,.]\}_B$ on the field $\mathcal{F}_B$ on outcome set B. The state evolves according to

$$\rho_i \longrightarrow \rho_f = \phi_1[A,\rho_i] \longrightarrow \rho_f(\tau) := U_{0\to\tau}\, \rho_f\, U_{0\to\tau}^{\dagger} \quad ,$$

upon which the second measurement is performed. As argued above, 'monitoring' corresponds to requiring $q(\tau) \longrightarrow m_{2\circ 1}$ on the undisturbed state ρ_i, rather than

[27] This can, e.g., be done through extension of the Hilbert space: A. Holevo (1986): *Theor. & Math. Phys.* **65**, p. 1250 [cf. also ch. III].

$\mathsf{q} \longrightarrow \mathsf{m}_2$ on the disturbed state $\rho_f(\tau)$ (i.e. $\mathsf{q}(\tau) \longrightarrow U_{0\to\tau}^\dagger\, \mathsf{m}_2\, U_{0\to\tau}$ on ρ_f). We use the notation

$$\mathsf{m}_{2o1} := \{M_{2o1}(d\beta)\} = \phi_1^\dagger[\mathbb{R}, U_{0\to\tau}^\dagger\, \mathsf{m}_2\, U_{0\to\tau}] \quad . \tag{26}$$

This new requirement means physically that the second meter should be designed such as to compensate for the modification of the object's state due to the first meter [§ 1.1]. Since the observables involved are incompatible, this compensation can never be complete. This assertion constitutes the disturbance interpretation of an inaccuracy inequality these two observables. Such an inequality follows from (III.98), analogously to (8), because the $\mathsf{q},\mathsf{q}(\tau)$ case is mathematically identical to the position-momentum case [with the substitution $\hbar \longrightarrow \hbar\tau/m$, see (21)]. We have

$$\delta_{\mathsf{q}(\tau)\longrightarrow \mathsf{m}_{2o1}}\, \delta_{\mathsf{q}\longrightarrow \mathsf{m}_1} \geq \frac{\tau}{m} \quad . \tag{27}$$

We end with a special case. If

$$\phi_1^\dagger[\mathbb{R}, \mathsf{q}(\tau)] \longleftarrow \mathsf{q}(\tau) \quad , \tag{28}$$

any m_2 such that $\mathsf{m}_2 \longleftarrow \mathsf{q}$ satisfies $\mathsf{q}(\tau) \longrightarrow \mathsf{m}_{2o1}$. Therefore, if we have one position meter that makes repeated measurements on the same object (i.e. $\phi_1 = \phi_2$) and satisfies (28), we can regard its outcome *both* as a (non-ideal) position measurement *and* as a (non-ideal) measurement of the position as it would have been were the previous measurement not performed. More precisely, its POVM satisfies both $\mathsf{m}_1 \longleftarrow \mathsf{q}$ and $\mathsf{m}_{1o1} \longleftarrow \mathsf{q}(\tau)$, with $\mathsf{m}_{1o1} := \phi_1^\dagger[\mathbb{R}, U_{0\to\tau}^\dagger\, \mathsf{m}_1\, U_{0\to\tau}]$. Moreover, under condition (28) the determinative $\mathsf{q}(\tau)$ disturbance results in preparative "smearing", in extra scatter in the results of the next measurement [cf. (17)]. If we arrange it so that the two non-idealities are equal $[\delta_{\mathsf{q}\longrightarrow \mathsf{m}_1} = \delta_{\mathsf{q}(\tau)\longrightarrow \mathsf{m}_{1o1}}]$, we get from (29):

$$\delta_{\mathsf{q}\longrightarrow \mathsf{m}_1} \geq \sqrt{\tau/m} \quad . \tag{27'}$$

Yuen-Tsyplyaev type schemes do not satisfy (28), and therefore may violate (27').

The possibilities of position monitoring are indeed limited. An SQL does exist, if we interpret 'position monitoring' as referring to the undisturbed trajectory. Then we must take the disturbance due to the first measurement properly into account, and we can derive (27). Ineq. (27) indicates how successive measurements lead to a "loss of memory" in the object system: as more and more q measurements are performed on it, it becomes harder and harder to find out the properties of the original input state[28]. Similar results can be expected to hold for the measurement of any quantity that is not conserved. This result does not contradict the non-existence of a quantum bound to weak force detection efficiency. That problem is quite different from position monitoring. Our "SQL" was derived for the latter, and it does not affect the former. Position monitoring is not the optimal way to detect the presence of a weak force.

4 INTERFERENCE VERSUS PATH

One of Bohr's most famous illustrations of complementarity is the double slit [ch. II]. One way of looking at this experiment, via the Hilgevoord Uffink relations[29], was already given in § 1.2. The claim that the UP makes it impossible in this experiment to find out (with certainty) through which slit the particle has passed without (completely) wiping out the interference pattern, can also be investigated in another way, via the inaccuracy principle.

Assume that the vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ symbolize two different paths of the particle. For sake of simplicity we shall take these states orthonormal. We are here dealing with a superposition of the form $\alpha|\psi_1\rangle + \beta\exp(j\eta)|\psi_2\rangle$. In the double slit the parameter η is the consequence of a difference in path length, and it varies as the detection point varies along the screen. Other interference experiments, most notably

[28]Cf. Belavkin, *op. cit.*

[29]Cf. J. Uffink & J. Hilgevoord (1988): *Physica B* **151**, p. 309

neutron interferometry, can also be dealt with in this way[30]. We have two extreme situations:

a) We want to know precisely which path the particle took. Then we should measure the PVM $\mathfrak{e} = \{E_1, E_2\}$; $E_j := |\psi_j\rangle\langle\psi_j|$.

b) We want to get a perfect look at the interference pattern. In that case we will want to measure the PVM $\mathfrak{f} = \{F_+, F_-\}$; $F_\pm := |\xi_\pm\rangle\langle\xi_\pm|$; $|\xi_\pm\rangle := \frac{1}{2}\sqrt{2}(|\psi_1\rangle \pm |\psi_2\rangle)$.

The PVM's $\mathfrak{e}$ and $\mathfrak{f}$ are incompatible. Hence the inaccuracy principle is applicable, in its original form as well as in the disturbance form. The two demands a) and b) can be jointly met only at the expense of some non-ideality.

The relations between this treatment of the double slit, and that given in § 1.2 may not at first be obvious. In the latter case, however, we were concerned with uncertainty of prediction (i.e. scatter). We wanted on the one hand to characterize the statistical distribution of outcomes over the different interference peaks, and on the other hand characterize the uncertainty about through which slit the particle went based on knowledge of the initial state (~ position scatter). Accordingly, in § 1.2, we could only study the effect of improvement of preparative knowledge of one variable on the quality of preparation of the other. Here, however, we intend to use a measurement to really *find out* which path it took, without considering the initial state. Therefore we can here consider the case of a true joint measurement of path & interference. Whereas in § 1.2 we looked at the double slit from the preparative point of view, we consider it here in a determinative light.

[30]Uffink & Hilgevoord, *op. cit.* (1988); W. de Muynck & H. Martens (1990): *Phys. Rev. A* 42, p. 5079

5 MEASUREMENT PROCESSES

In ch. III we already did some work on the connection between non-ideality and the measuring process. We shall continue here, in preparation of some concrete schemes we will treat shortly. Our treatment will now include consideration of the state after measurement. In ch. III we compared our approach to estimation theoretic methods, and chose the Heisenberg picture. Here we shall take traditional textbook measurement theory as a starting point, and accordingly we will use the Schrödinger picture. We start with a measuring device $\mathcal{A}$, initially in state $|\xi\rangle_a \in \mathscr{H}'$. The object observable $\mathfrak{z}$ to be measured (non-ideally) is the PVM of Z with eigenstates $|z_\ell\rangle_o \in \mathscr{H}$. Traditionally[31], we assume the interaction to effect a transition

$$|z_\ell\rangle_o \otimes |\xi\rangle_a \longrightarrow |z_\ell\rangle_o \otimes |\xi_\ell\rangle_a \quad . \tag{29}$$

Such a transition occurs, e.g., with an interaction Hamiltonian that is a function of Z and some apparatus operator, either if Z is conserved or in the impulsive interaction approximation[32]. Then we assume the pointer states $|\xi_\ell\rangle_a$ to be orthogonal (perhaps as a consequence of "macroscopic distinguishability"). Reading of the corresponding PVM on $\mathscr{H}'$ leads to a measurement of the first kind, with OVM $\mathfrak{E} = \{\epsilon_m\}_M$,

$$\epsilon_m[\rho] = \sum_\ell |z_\ell\rangle_o \langle z_\ell| \, {}_o\langle z_\ell|\rho|z_\ell\rangle_o \quad . \tag{30}$$

Thus $\mathfrak{z}$ is precisely measured. In case of degeneracy, a generalization of (30) holds[33].

Of course, as emphasized before, such a measurement description is hardly realistic. The system directly interacting with the object is never "macroscopic", so that orthogonality of the $|\xi_\ell\rangle_a$'s is questionable, as is the possibility of reading out a PVM on the ancilla. Dropping both requirements, we get an OVM $\mathfrak{N} = \{\nu_m\}_M$ on $\mathscr{H}$,

$$\nu_m[\rho] = \sum_\ell \sum_{\ell'} |z_\ell\rangle_o \langle z_{\ell'}| \, {}_o\langle z_\ell|\rho|z_{\ell'}\rangle_o \, {}_a\langle \xi_{\ell'}|O'_m|\xi_\ell\rangle_a \quad , \tag{31}$$

---

[31] J. von Neumann (1932): *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin)

[32] See e.g. Bohm, *op. cit.* (1951) ch. 22

[33] G. Lueders (1951): *Ann. der Phys.* **8**, p. 322

where $o' = \{O'_m\}_M$ is the read-out EVM on $\mathscr{H}'$. The EVM $n = \{N_m\}_M$ on $\mathscr{H}$, corresponding to (31), is then given by

$$(32) \qquad N_m = \sum_\ell |z_\ell\rangle_o\langle z_\ell| \, {}_a\langle\xi_\ell|O'_m|\xi_\ell\rangle_a = \sum_\ell |z_\ell\rangle_o\langle z_\ell| \, \lambda_{m\ell} \ ,$$

with

$$\lambda_{m\ell} = {}_a\langle\xi_\ell|O'_m|\xi_\ell\rangle_a \ .$$

Therefore reading of an *arbitrary* pointer observable o leads to a non-ideal $\mathfrak{z}$–measurement with matrix $\{\lambda_{m\ell}\}$. The amount of non-ideality is related to the distinguishability of the states $|\xi_\ell\rangle_a$ by means of o. The choices leading to (30) represent simply one optimal choice. Transformation (31) leaves Z–eigenstates intact $[\nu_M^\dagger[\mathfrak{z}] = \mathfrak{z}]$. We may generalize (29) to

$$(29') \qquad |z_\ell\rangle_o \otimes |\xi\rangle_a \longrightarrow |\zeta_\ell\rangle_o \otimes |\xi_\ell\rangle_a \ ,$$

so that (31) changes into

$$(31') \qquad \nu_m[\rho] = \sum_\ell \sum_{\ell'} |\zeta_\ell\rangle_o\langle\zeta_{\ell'}| \, {}_o\langle z_\ell|\rho|z_{\ell'}\rangle_o \, {}_a\langle\xi_{\ell'}|O'_m|\xi_\ell\rangle_a \ .$$

Eqs. (32) still follow if we only require either ${}_a\langle\xi_{\ell'}|O_m|\xi_\ell\rangle_a = 0$ if $\ell \neq \ell'$ or ${}_o\langle\zeta_{\ell'}|\zeta_\ell\rangle_o = 0$ if $\ell \neq \ell'$. The first case is realized when the pointer states $|\xi_\ell\rangle_a$ are orthogonal, and o is the corresponding PVM. Then the state after measurement contains only Z information about the initial state: *every* EVM $\{M_k\}_K$ measured after $\mathfrak{N}$ can be seen as a non-ideal measurement of Z before $\mathfrak{N}$:

$$(33) \qquad \mathrm{Tr}(M_k \, \nu_m[\rho]) = \sum_\ell {}_o\langle\zeta_\ell|M_k|\zeta_\ell\rangle_o\langle z_\ell|\rho|z_\ell\rangle_o \, {}_a\langle\xi_\ell|O_m|\xi_\ell\rangle_a$$

$$\Updownarrow$$

$$\nu_m^\dagger[M_k] = \sum_\ell |z_\ell\rangle_o\langle z_\ell| \, {}_o\langle\zeta_\ell|M_k|\zeta_\ell\rangle_o \, {}_a\langle\xi_\ell|O_m|\xi_\ell\rangle_a \ .$$

All information about other observables than $\mathfrak{z}$ has been obliterated. Subsequent measurements will provide us with no additional information about the undisturbed object. This case is not very interesting.

In the second case, even though $\mathfrak{z}$ is not left untouched, $\mathfrak{z}$ is undisturbed. The $\mathfrak{z}$–information has only been transferred into another variable. We can measure $\mathfrak{z}$–before–$\mathfrak{N}$ with arbitrary accuracy. This scheme is, however, still not the most general. In (29') we assumed that, starting from a Z–eigenstate, the final $\mathcal{O}+\mathcal{A}$ state is unentangled. This will not always be true, as in case of degeneracy

$$(29'') \qquad |z_\ell\rangle_o \otimes |\xi\rangle_a \longrightarrow |\tilde{\xi}_\ell\rangle_{o+a}$$

may occur. General statements about characterizations of the measurements in terms of non-ideality become impossible. In special cases, however, the measurement can even then be non-ideal, e.g. due to symmetries in the interaction Hamiltonian [see § 5.2].

In view of (8), however, we are also interested in the disturbance of observables other than the one we measure. Therefore we shall again consider the pair of observables $\mathfrak{x}$ and $\mathfrak{y}$ we discussed in § III.2.6. These correspond to self-adjoint operators X and Y on $\mathbb{C}^n$ with complete sets of eigenvectors $(|x_k\rangle_o)_K$ and $(|y_k\rangle_o)_K$ ($K = \{0, \dots, n-1\}$) that satisfy

$$(34) \qquad {}_o\langle x_\ell | y_k \rangle_o = n^{-\frac{1}{2}} \exp(i\tfrac{2\pi}{n}\ell k) \quad .$$

For a non-ideal $\mathfrak{x}$ meter $\mathfrak{N}$ [$\mathfrak{N}$ as above, substituting X for Z], (8) reads

$$(35) \qquad J^{\delta\mathfrak{y}}_{\nu_M} + J_{\mathfrak{x}\to n} \geq \log(n) \quad .$$

For sake of definiteness we shall require that $\mathfrak{y}$–before–$\mathfrak{N}$ can be measured through $\mathfrak{y}$:

$$(36) \qquad \forall_\rho \; S_x \, \nu_M[\rho] \, S_x^\dagger = \nu_M[S_x \, \rho \, S_x^\dagger]$$

$$\Downarrow$$

$$\nu_M^\dagger[\mathfrak{y}] \leftarrow \mathfrak{y}$$

Eq. (36) is automatically satisfied when the OVM $\mathfrak{N}$ is covariant, i.e. if

$$\forall_\rho \;\; S_x \nu_m[S_x^\dagger \rho \; S_x] S_x^\dagger \;=\; \nu_{[[m+1]]}[\rho] \;\; .$$

with:

$$[[k]] \;=\; k \bmod n \;\; ; \;\; S_x \;=\; \Sigma_{k\in K} \, |x_{[[k+1]]}\rangle_o\langle x_k| \;\; .$$

This is equivalent to

$$(37) \qquad |\zeta_{[[k+1]]}\rangle_o\langle\zeta_{[[k'+1]]}| \; {}_a\langle\xi_{[[k'+1]]}|\xi_{[[k+1]]}\rangle_a \;=$$

$$= \; S_x|\zeta_k\rangle_o\langle\zeta_{k'}|S_x^\dagger \; {}_a\langle\xi_{k'}|\xi_k\rangle_a \;\; .$$

Note that in case of (36), analogous to (18), the Y distribution after $\mathfrak{N}$ measurement is a smeared version of the Y distribution before measurement, leading to preparative Y disturbance ("excess Y scatter") equal to the determinative η disturbance. Determinative η disturbance is here the accuracy a measurement of n after completionn of the interaction achieves with respect to η–before–$\mathfrak{N}$. Using (37), the non-ideality matrix for (36) can be seen to be

$$(38) \qquad \mu_{\ell k} \;=\; {}_o\langle y_\ell| \; \nu_M^\dagger[|y_k\rangle_o\langle y_k|] \; |y_\ell\rangle_o \;=$$

$$= \; \Sigma_{m\in K} \, n^{-\frac{1}{2}} \; {}_a\langle\xi_0|\xi_m\rangle_a \exp(\tfrac{2\pi}{n}\ell m) \; {}_o\langle\zeta_0|y_k\rangle_o\langle y_k|\zeta_m\rangle_o \qquad (\ell,k \in K).$$

If $|\zeta_k\rangle_o = |x_k\rangle_o$ [$\Leftrightarrow \; \nu_M^\dagger[\mathfrak{x}] = \mathfrak{x}$; viz. (29)], this reduces further to

$$(38') \qquad \mu_{\ell k} \;=\; \Sigma_{m\in K} \, n^{-\frac{1}{2}} \exp(\tfrac{2\pi}{n}\,[\ell\text{–}k]m) \; {}_a\langle\xi_0|\xi_m\rangle_a \qquad (\ell,k \in K).$$

From (38′) we see that the η disturbance is related to the Fourier transform of a column of the inner product matrix ${}_a\langle\xi_0|\xi_m\rangle_a$. Thus [cf. § III.1.2] disturbance is reciprocally related to the distinguishability of these apparatus states.

Because of the analogy with the position-momentum case, the above treatment also characterizes the γ–microscope[34]. Given that the electron is in state $|q\rangle_o$, the

[34] J. Hilgevoord & J. Uffink (1990): in *Sixty-two years of Uncertainty: Historical, Philosophical and Physical Inquiries into the Foundations of Quantum Mechanics* (ed. by A. Miller; Plenum, NY), p. 121

scattered photon arrives in a wavepacket $|\xi_q\rangle_a$ at the photographic plate. The device's resolution is then determined by the distinguishability of $|\xi_q\rangle_a$ and $|\xi_{q'}\rangle_a$ in relation to the distance $|q-q'|$. Covariance of the microscope's OVM follows from momentum conservation, if the light incident on the electron is monochromatic (and therefore homogeneous in direct space). Thus conditions analogous to the above (i.e. (36) and $\nu_M^\dagger[\mathfrak{x}] = \mathfrak{x}$) actually hold in the γ–microscope case, and we may assume that determinative momentum disturbance is here also reciprocally related to the distinguishability of the pointer states $|\xi_q\rangle_a$. Moreover, there is also be a preparative $\boldsymbol{P}$ smearing in the microscope, reciprocally related to its resolution. But, as noted above, in this latter respect the microscope is not representative of measurements in general.

5.1 Optical Kerr Effect

As a more concrete illustration of the above, we shall treat the non-destructive photon number determination using the optical Kerr effect. This effect is contained in the third order non-linear susceptibility $\chi^{(3)}$. This contributes a term[35]

$$(39) \qquad \boldsymbol{H}^{(3)} = \frac{3}{4}\iiint \mathrm{d}^3V \sum_{i,j,k,\ell} \chi^{(3)}_{ijk\ell}\, \boldsymbol{E}_i \boldsymbol{E}_j \boldsymbol{E}_k \boldsymbol{E}_\ell$$

to the energy of the electro-magnetic field. If we restrict ourselves to two modes, S and P, with frequencies ω_S and ω_P, we can rewrite (39) to

$$(40) \qquad \boldsymbol{H}^{(3)} = \boldsymbol{H}_S + \boldsymbol{H}_P + \boldsymbol{H}_I \ ,$$

with

$$(41a) \qquad \boldsymbol{H}_S = D\omega_S^2 \left[\chi^{(3)}(\omega_S;\omega_S,-\omega_S,\omega_S)\, \boldsymbol{a}_S^\dagger \boldsymbol{a}_S \boldsymbol{a}_S^\dagger \boldsymbol{a}_S + \text{5 permutations}\right] ;$$

$$(41b) \qquad \boldsymbol{H}_P = D\omega_P^2 \left[\chi^{(3)}(\omega_P;\omega_P,-\omega_P,\omega_P)\, \boldsymbol{a}_P^\dagger \boldsymbol{a}_P \boldsymbol{a}_P^\dagger \boldsymbol{a}_P + \text{5 permutations}\right] ;$$

---

[35] N. Imoto, H. Haus & Y. Yamamoto (1985): *Phys. Rev. A* **32**, p. 2287; Y. Yamamoto *et al.* (1990): *Progr. in Opt.* (ed. by E. Wolf; North Holland, Amsterdam) **28**, p. 87

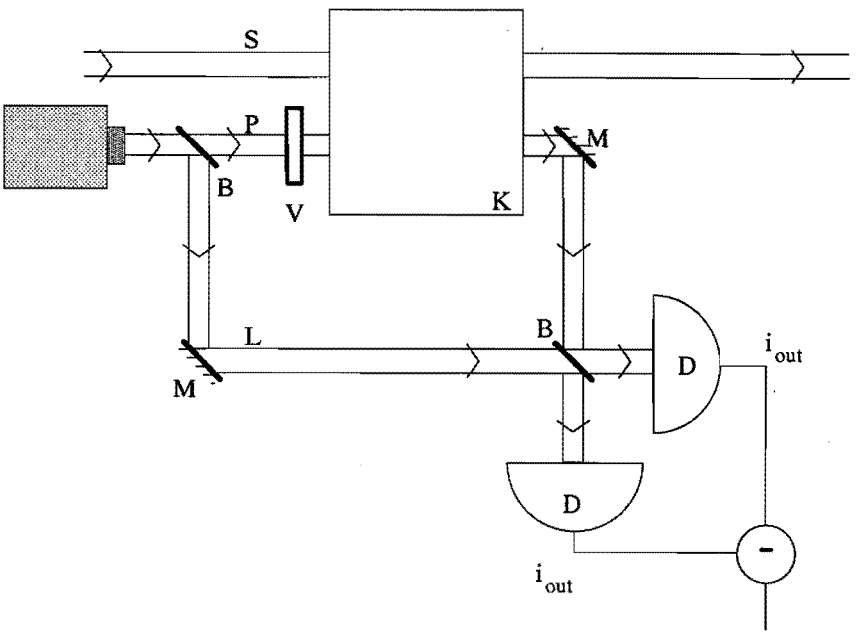


fig. 4 *Kerr QND measurement of photon number. A signal beam* S *is mixed with a probe beam* P *into a non-linear medium. The outgoing probe beam is fed into a homodyne detector. [*V *is a delay,* B*'s are beamsplitters,* M*'s are mirrors,* D*'s are detectors, and* L *indicates the homodyne local oscillator beam.]*

(41c) $$H_I = D\omega_S\omega_P[\chi^{(3)}(\omega_P;\omega_P,-\omega_S,\omega_S)\, a_P^\dagger a_P\, a_S^\dagger a_S + \text{23 perm.}] \;;$$

$$D = \frac{3}{16\, V\epsilon^2} \;.$$

[V is the quantization volume, and ϵ is the dielectric constant; a_P and a_S ($a_P^\dagger$ and $a_S^\dagger$) are the respective boson annihilation (creation) operators.] Normal ordering of the permutations gives extra terms 1 , $a_P^\dagger a_P$ and $a_S^\dagger a_S$ we shall ignore: they merely shift the frequency of the modes. Thus we can simplify (40) into

(42) $$H^{(3)} = \chi_S N_S^2 + \chi_P N_P^2 + 4\chi_I N_S N_P \;,$$

with

$$\chi_S = 6D\omega_S^2\ \chi^{(3)}(\omega_S;\omega_S,-\omega_S,\omega_S)\ ;\ N_S = a_S^\dagger a_S\ .$$

(χ_P, N_P and χ_I are defined analogously).

In (39) we can see how the Kerr effect can be seen as an intensity-dependent refractive index. Accordingly, the optical path length of the probe beam depends on the intensity of the signal beam. The Kerr effect brings about a correlation between the number of photons in the S mode and the phase of the P mode. A determination of the probe beam phase would then allow the number of photons in the signal beam to be calculated. As these photons need not be absorbed, the measurement is non-destructive. A device in which the Kerr effect is thus utilized is sketched in fig. 4. A laser beam is split in two, after which one sub-beam (P) is coupled into a non-linear medium. The signal beam (S), upon which the measurement is to be carried out, is also fed into the medium. Afterwards the P–beam is mixed with the other laser sub-beam (L), to perform a homodyne detection. Homodyne detection[36] may be considered a measurement of "position" $Q = \frac{1}{2}\sqrt{2}(a + a^\dagger)$ on the P–mode. We neglect losses in the device[v], so that we have a true 'Quantum Non-Demolition' (QND) measurement.

We shall at first neglect the first two terms ('self-phase modulation') in the interaction (42) (i.e. $\chi_S = \chi_P = 0$). The initial P–state is the coherent state[37] $|\beta\rangle_P$. If we take an initial S–state $|n\rangle_S$, it is not difficult to verify that after a time $\tau = \ell/c$ of interaction, the S+P state becomes

(43) $$|n\rangle_S \otimes |\beta\rangle_P \rightarrow |n\rangle_S \otimes |\beta_n \exp(-i\tau\omega_P)\rangle_P \exp(-i\tau\omega_S n)$$

with

$$\beta_n = \beta \exp(-i4\tau\chi_I n)\ .$$

---

[36] H. Yuen & J. Shapiro (1978): *IEEE Trans. Inf. Th.* IT-24, p. 657; (1979): *ibid.* IT-25, p. 179; (1980): *ibid.* IT-26, p. 78.

[37] See e.g. R. Loudon (1983): *The Quantum Theory of Light* (2nd ed., Oxford University Press)

Thus the Kerr-device is of type (29). The states $|\beta_n\rangle$ play the role of the pointer states. Hence performance of *any* measurement on P would give information about N_S. The homodyne detection we perform can be calculated to lead to the EVM

(44) $$P(dq) = \sum_n |n\rangle_S\langle n| \lambda_n(dq) \quad ; \quad \lambda_n(dq) = |\langle\beta_n|q\rangle|^2 dq \quad .$$

The homodyne detector is only sensitive to the relative phase of P and L beams. Therefore the phase $i\tau\omega_P$, which is present in both beams, is not involved in (44). Since, for convenience, we would like

$$\int \lambda_0(dq)\, q = 0 \quad ,$$

we require that

(45) $$\beta = i|\beta| \quad .$$

Then, as is easily seen,

(46) $$\langle q\rangle_n := \int \lambda_n(dq)\, q = \sqrt{2}\, |\beta| \sin(\tau 4\chi_I n) \quad .$$

If $\tau 4\chi_I n \ll 1$, the meter works approximately linear[vi], and $\langle q\rangle_n \simeq 4\sqrt{2}\, |\beta|\, \tau\chi_I n$. Condition (45) makes the measurement (in first order) unbiased, but it also causes the linear regime (the area of maximal sensitivity) to be applicable for low photon numbers. The condition can be fulfilled by varying the relative optical distance of P and L beams, by adjusting the delay V.

In the linear regime

(47) $$\lambda_n(dr) \simeq (2\pi\sigma_N^2)^{-\frac{1}{2}} \exp\left(-\frac{(r-n)^2}{2\sigma_N^2}\right) dr$$

with

$$r = q/[4\sqrt{2}\, |\beta|\, \tau\chi_I] \quad ;$$

$$\sigma_N = \frac{1}{8\, |\beta|\, \tau\chi_I} \quad .$$

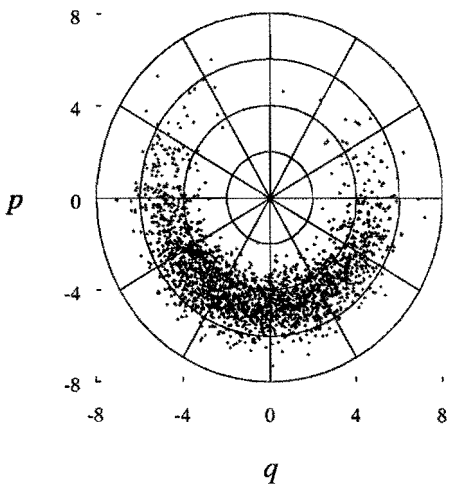


fig. 5 *Husimi distribution of a coherent state, squeezed by a Kerr QND measurement. The squeezing reduces* **N** *scatter at the cost of* ϕ *width. [Input state* $|\alpha\rangle$, $\alpha = 3\sqrt{2}$; *Kerr device with* $\tau\chi_I = 0.002$ *and* $\beta = 100i$; *squeezing conditional on measurement outcome* $q_0 = 12$; *Husimi distribution* $\langle\gamma|\rho|\gamma\rangle$ *is plotted, with* $\gamma = \frac{1}{2}\sqrt{2}(q + ip)$.*]*

The (unnormalized) S–output state, conditional on the measurement outcome q, can be calculated to be

$$(48) \qquad |\psi\rangle_S\langle\psi| \;\longrightarrow\; \mathbf{K}[dq, |\psi\rangle_S\langle\psi|] \;:=\; \boldsymbol{A}_q|\psi\rangle_S\langle\psi|\boldsymbol{A}_q^\dagger \, dq$$

with

$$\boldsymbol{A}_q \;:=\; \sum_n \langle q|\beta_n\rangle \, |n\rangle_S\langle n| \quad .$$

The transformation **K** defined by (48) is seen to satisfy the requirements of an OVM [§ 1.1]. If we look at the ***N***–distribution of the state after measurement conditional on the outcome q_0, we see that it is related to the distribution before measurement by a factor $|\langle q|\beta_n\rangle|^2$. Thus, if the original distribution is wider than the Gaussian of this factor (e.g. if the input state is coherent), the Kerr-device has the effect of narrowing it. This is called *squeezing* (fig. 5).

We are, in view of the foregoing sections, more interested in the disturbing influence on the optical phase of the S–mode. As a description of phase, we will choose an EVM introduced by Lévy–Leblond[38,vii]:

$$(49) \qquad \boldsymbol{M}(d\phi) \;:=\; |\phi\rangle\langle\phi|d\phi \quad ; \quad |\phi\rangle \;:=\; \sum_n (2\pi)^{-\frac{1}{2}} \exp(in\phi) \, |n\rangle \quad .$$

---

[38] J.–M. Levy–Leblond (1976): *Ann. of Phys.* **101**, p. 319

Note that the "states" $|\phi\rangle$ satisfy a closure relation

$$(50) \qquad \int_{-\pi}^{\pi} d\phi \, |\phi\rangle\langle\phi| \; = \; 1 \quad ,$$

but are not orthogonal:

$$(51) \qquad \langle\phi|\phi'\rangle \; = \; \frac{1}{4\pi} + \tfrac{1}{2}\delta([\phi-\phi'] \bmod 2\pi) + \frac{i}{4\pi}\cot(\frac{\phi-\phi'}{2}) \quad .$$

They are eigenstates of the shift operator e:

$$(52) \qquad e|\phi\rangle \; = \; \exp(i\phi)|\phi\rangle \; ; \; e \; = \; \Sigma_n \, |n\rangle\langle n+1| \; = \; \int_{-\pi}^{\pi} d\phi \exp(i\phi) \, |\phi\rangle\langle\phi| \quad .$$

Substituting $|\phi\rangle\langle\phi|$ into the conjugate of (48), and averaging over q gives

$$(53) \qquad \mathbf{K}^{\dagger}[\mathbb{R}, |\phi\rangle_S\langle\phi|d\phi] \; = \; d\phi \int dq \, A_q^{\dagger} \, |\phi\rangle_S\langle\phi|A_q \; =$$

$$= \; \int_{-\pi}^{\pi} d\phi' \, \mu(d\phi,\phi') \, |\phi'\rangle_S\langle\phi'|$$

with

$$(54) \qquad \mu(d\phi,\phi') \; = \; \frac{d\phi}{2\pi} \Sigma_{k\in\mathbb{Z}} \exp(ik[\phi - \phi' + \tau\omega_S]) \, \langle\beta|\exp(4i\tau\chi_I k\mathcal{N})|\beta\rangle \quad .$$

In the linear regime, states have phase distributions involving only low-frequency components. Therefore only terms with low k are important in (54), so that

$$(55) \qquad \mu(d\phi,\phi') \; \simeq \; \frac{d\phi}{2\pi} \, \theta_3[-2|\beta|^2\tau\chi_I + \tfrac{1}{2}\tau\omega_S + \frac{\phi-\phi'}{2} \; ; \; \exp(-8|\beta|^2\chi_I^2\tau^2)] \quad .$$

Here θ_3 denotes the third of Jacobi's theta functions[39] [fig. 6]. Thus both (54) and (55) are *covariant*: they only depend on $\phi - \phi'$. Therefore the amount of non-ideality can be characterized by the width of a single probability distribution [ch. III]. Note further that (an analog of) (18) is satisfied. There is both determinative and preparative disturbance, proportional to the width of the distribution (54) or (55). A

[39] See e.g. E. Whittaker & G. Watson (1927): *A Course of Modern Analysis* (4th ed.; Cambridge University Press)

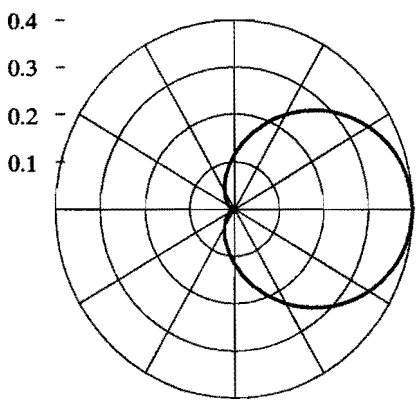


fig. 6 *Polar plot of the function* $\mu(d\phi,\phi')/d\phi$ *vs.* ϕ-ϕ'. *The linear regime is assumed, and we have taken* $|\beta|^2\chi_I^2\tau^2 = \frac{1}{8}$. *The phase bias* $2|\beta|^2\tau\chi_I + \frac{1}{2}\tau\omega_S$ *is ignored.*

convenient width measure for phase probability distributions (i.e. for probability distributions on a circle)is provided by[40]

$$V_\phi := -1 + \left| \int_{-\pi}^{\pi} \mu(d\phi,0)\exp(i\phi) \right|^{-2} \quad . \tag{56}$$

Substituting (54) in (56) gives

$$V_\phi = -1 + \exp(2|\beta|^2[1-\cos(4\tau\chi_I)]) \simeq -1 + \exp([4|\beta|\tau\chi_I]^2) \quad , \tag{57}$$

if $\tau\chi_I \ll 1$ .Phase disturbance and *N*–inaccuracy, as derived from (57) and (47) respectively, are plotted in fig. 7, along with the fundamental bound for this case:

$$V_\phi\,\sigma_N^2 \geq \tfrac{1}{4} \quad . \tag{58}$$

Because of the covariance of both (55) and (47), this bound can be straightforwardly derived from the phase-number scatter relation[41] [cf. § III.2.6]. The idealized Kerr-device we treated here is seen to be slightly sub-optimal. If $\tau\chi_I \ll 1$

$$\log(1 + V_\phi)\,\sigma_N^2 = \tfrac{1}{4} \quad .$$

---

[40]A. Holevo (1982): *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam), p. 180

[41]Levy–Leblond, *op. cit.* (1976)

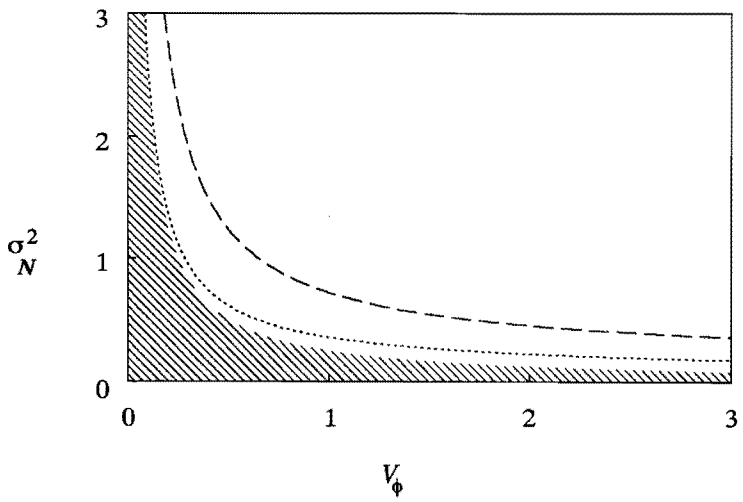


fig. 7 *Determinative N-quality versus phase disturbance plotted for the Kerr-device both without (dotted) and with self-phase modulation (dashed:* $\tau\chi_P = 0.000025$*). $\tau\chi_I$ is varied around* 0.0025*. The combinations forbidden by the inaccuracy principle are in the shaded area* [$\beta = 100i$].

If we include self-phase modulation effects, eq. (43) becomes

(43′) $$|n\rangle_S^{\otimes}|\beta\rangle_P \longrightarrow |n\rangle_S^{\otimes}|\zeta_n\rangle_P \exp(-i\tau\omega_S n)$$

with:

$$|\zeta_n\rangle = \sum_m |m\rangle \frac{\beta^m \exp(-\frac{1}{2}|\beta|^2 - i\varphi_{nm}\tau)}{\sqrt{m!}} \; ;$$

$$\varphi_{nm} = \tau\chi_S n^2 + \tau\chi_P m^2 + 4\tau\chi_I nm \; .$$

We adjust (*via* V) the relative optical distances of L and P such that $\arg(\beta) = \pi/2 + |\beta|^2\sin(2\tau\chi_P) + i\tau\chi_P$: we replace (45) by

(45′) $$\tilde{\beta} = i|\tilde{\beta}|$$

with

$$\tilde{\beta} = \beta \exp(|\beta|^2[\exp(-i2\tau\chi_P) - 1] - i\tau\chi_P)$$

$$\simeq \beta \exp(-i2|\beta|^2\tau\chi_P - 4|\beta|^2\tau^2\chi_P^2) \; .$$

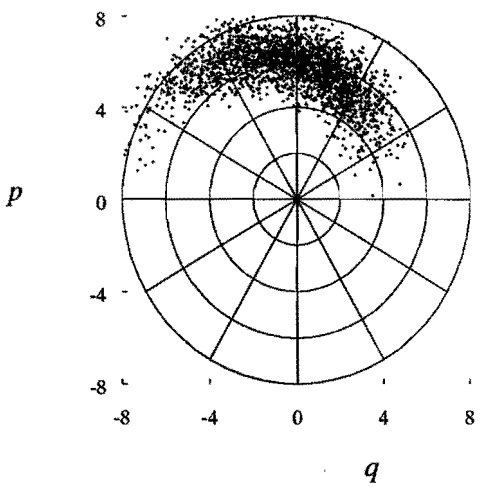


fig. 8 *Husimi distribution of self-phase modulated state. Note that, contrary to fig. 5, the N–scatter is not reduced. [Input state* $|\alpha\rangle$, $\alpha = 3\sqrt{2}\exp(-2i\tau\chi_P|\alpha|^2)$; $\tau\chi_P = 0.05$; *the distribution* $\langle\gamma|\rho|\gamma\rangle$ *has been plotted, with* $\gamma = \frac{1}{2}\sqrt{2}(q + ip)$.*]*

Then some tedious calculation shows that

$$(46') \qquad \langle q\rangle_n = \sqrt{2}\,|\tilde\beta|\,\sin(\tau 4\chi_I n) \simeq 4\sqrt{2}\,|\tilde\beta|\,\tau\chi_I\, n \;;$$

$$(59) \qquad \langle q^2\rangle_n := \int \lambda_n(dq)\, q^2 \simeq$$

$$\simeq \langle q\rangle_n^2 + (4\sqrt{2}\,|\tilde\beta|\,\tau\chi_I)^2\Big[(8|\tilde\beta|\,\tau\chi_I)^{-2} + \frac{\sinh(4|\beta|^2\tau^2\chi_P^2)}{16\tau^2\chi_I^2} +$$

$$-16\tau\chi_P(4|\beta|^2\tau^2\chi_P^2 - 1)\exp[-4|\beta|^2\tau^2\chi_P^2]\,(n - \tfrac{32}{3}n^3) +$$

$$+ (\exp[-4|\beta|^2\tau^2\chi_P^2] - 1)\,n^2\Big] \;.$$

We have assumed $\tau\chi_P \le \mathcal{O}(1/|\beta|) \ll 1$ and $8\tau\chi_I n \ll 1$. Further restriction to $\tau\chi_P|\beta| \ll 1$ simplifies these expressions to

$$(59') \qquad \langle q^2\rangle_n \simeq \langle q\rangle_n^2 + (4\sqrt{2}\,|\tilde\beta|\,\tau\chi_I)^2\Big[(8|\tilde\beta|\,\tau\chi_I)^{-2} + \frac{4|\beta|^2\tau^2\chi_P^2}{16\tau^2\chi_I^2} +$$

$$+ 16\tau\chi_P(n - \tfrac{32}{3}n^3) - 4|\beta|^2\tau^2\chi_P^2\, n^2\Big] \;.$$

We see that self-phase modulation decreases the accuracy of the meter (fig. 7). The cause of this is the effect it has on the pointer states. These states $|\zeta_n\rangle$ are no longer coherent. The self-phase modulation increases their phase-width (fig. 8). Since it is

precisely phase (corresponding to angle of rotation in the p,q–plane) that distinguishes the $|\zeta_n\rangle$ from each other, self-phase modulation makes it harder to tell them apart.

Also interesting is the effect of the self-phase modulation on the signal's phase. As is straightforwardly verified, (an analog of) (18) is no longer satisfied: a measurement of the phase of the outgoing signal mode cannot be considered a (non-ideal) measurement of the initial phase. Define generalized phase "states":

$$(60) \qquad |\phi;v\rangle \;:=\; \sum_n (2\pi)^{-\frac{1}{2}} \exp[i\phi n + \tfrac{1}{2}iv n(n+1)]\,|n\rangle \quad .$$

The state $|\phi;v\rangle$ is an eigenvector of the operator $e\exp(-iv\hat{N})$ with eigenvalue $\exp(i\phi)$. Then the outcome probability distribution associated with measuring the EVM corresponding to (60) after the Kerr-device, is

$$(53') \qquad \mathrm{Tr}(\mathbf{K}[\mathbb{R},\rho]\;|\phi;v\rangle_S\langle\phi;v|\;d\phi) \;=\; \mathrm{Tr}(\mathbf{K}^\dagger[\mathbb{R},|\phi;v\rangle_S\langle\phi;v|\;d\phi]\;\rho) \quad ;$$

$$\mathbf{K}^\dagger[\mathbb{R},|\phi;v\rangle_S\langle\phi;v|\;d\phi] \;=\; \int_{-\pi}^{\pi} d\phi'\;\mu(d\phi,\phi')\;|\phi';v'\rangle_S\langle\phi';v'|$$

with

$$(54') \qquad \mu(d\phi,\phi') \;=\; \frac{d\phi}{2\pi}\sum_{k\in\mathbb{Z}} \exp(ik[\phi-\phi'+\tau\omega_S-\tau\chi_S])\cdot \cdot\langle\beta|\exp(i4\tau\chi_I k\hat{N})|\beta\rangle \quad ;$$

$$v' \;=\; v + 2\tau\chi_S \quad .$$

Therefore a measurement of the EVM $\{|\phi;-2\tau\chi_S\rangle_S\langle\phi;-2\tau\chi_S|d\phi\}$ allows a non-ideal determination of the initial phase, with quality equal to that in the $\chi_S = 0$ case. Thus, contrary to the photon number accuracy, the phase disturbance is not affected by the self-phase modulation. This, however, is not immediately clear due to the self-phase modulation's distorting effect (53'), but it can be understood relatively easily: because the disturbance depends only on the interaction Hamiltonian and on the initial P state, and *not* on the read-out observable, the disturbance must be compatible with the accuracy achievable with an optimal choice of that read-out observable. Since the self-phase modulation corresponds to a unitary transformation, it can in principle be undone by a clever choice of read out observable. Thus the optimal photon number accuracy is not affected by the self-phase modulation either.

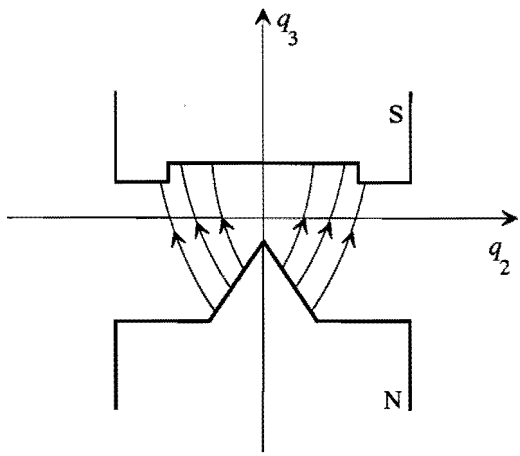


fig. 9 *Schematic cross-section of the Stern-Gerlach setup. The inhomogeneous magnetic field is indicated by the dashed lines. The electrons move in the* 1*-direction, perpendicular to the plane of the drawing*

Note further that the determinative phase disturbance is equal to the preparative phase disturbance: in order to control outgoing signal phase (i.e. achieve a small phase scatter), the initial state should be aimed to approach $|\phi;2\tau\chi_S\rangle_S\langle\phi;2\tau\chi_S|$.

5.2 Stern–Gerlach

A Stern-Gerlach device may be used to show the effects of spin[42]. In a Stern-Gerlach device (fig. 9) a beam of spin ½ particles passes in the 1–direction through an inhomogeneous magnetic field. The field is (in lowest order) given by a vector potential

(61) $$\vec{A} = (-aq_2 + bq_2q_3, 0, 0) \Rightarrow \vec{B} = (0, bq_2, a - bq_3) \ .$$

[42]Bohm, *op. cit.*; F. Schroeck (1982): *Found. Phys.* **12**, p. 479; cf. appendix B

Note that $(\vec{\nabla}\cdot\vec{B}) = 0$ [cf. app. B]. The particle's Hamiltonian is then given by

$$H = \frac{1}{2m}([P_1 - q(-aQ_2 + bQ_2Q_3)]^2 + P_2^2 + P_3^2) + \\ + \mu\,(bQ_2\sigma_2 + [a - bQ_3]\sigma_3) \tag{62}$$

with

$$\mu = \frac{e}{2m_e} \ .$$

Here σ_3 is the (Pauli) operator representing spin in the 3–direction, with eigenvalues ± 1, and PVM s_3. For neutral atoms, such as in the original experiment, $q = 0$. For electrons, on the other hand, $q = -e$ and $m = m_e$. We choose such units that $2m_e = \mu = b = e = \hbar = 1$. In the case of neutrals this scaling leads to

$$H = \frac{1}{2m}(P_1^2 + P_2^2 + P_3^2) + Q_2\sigma_2 + [a - Q_3]\sigma_3 \ , \tag{63}$$

in the electron case to

$$H = [P_1 + (-aQ_2 + Q_2Q_3)]^2 + P_2^2 + P_3^2 + Q_2\sigma_2 + [a - Q_3]\sigma_3 \ . \tag{64}$$

The initial state of the particle is assumed to be a product of the spin part ρ and the spatial part $|\xi\rangle\langle\xi|$. We shall in this section assume the spatial part $|\xi\rangle$ to be initially Gaussian, centered at $q_2 = q_3 = 0$, and characterized by the variances $\langle\Delta^2 Q_2\rangle$ and $\langle\Delta^2 Q_3\rangle$. The spin dependent force, represented by the last term of (63) or (64), will bring about a spin dependent development of the wave packet in P–space. Thus P_2 and P_3 are our read-out variables. A realization might involve a screen placed very far away.

We shall first look at the standard case (63). Because of the absence of a Lorentz-force, it is simpler. If we denote the operator that reflects the spatial part of the state in the 1,3–plane by I_2, we see that

$$[I_2 \otimes \sigma_3]\, H\, [I_2 \otimes \sigma_3]^\dagger = H \ . \tag{65}$$

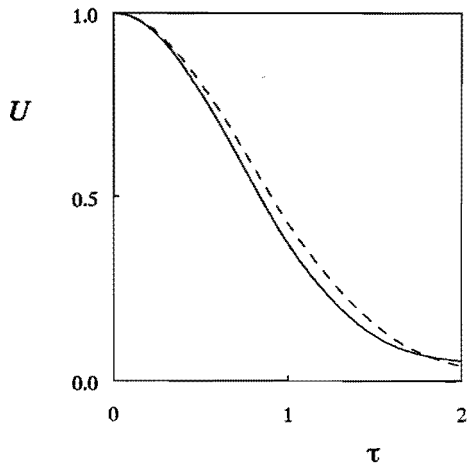


fig. 10 *Quality of the Stern-Gerlach with $a \neq 0$. The quantity U of (68) has been plotted vs. the measurement time τ. The solid line indicates neutrals, the dashed line electrons. [Neutrals: initially $\langle \Delta^2 Q_2 \rangle = \langle \Delta^2 Q_3 \rangle = 1$; $a = 20$; $P_1 = 0.03$. Electrons: initially $\langle \Delta^2 Q_2 \rangle = 1/64$; $\langle \Delta^2 Q_3 \rangle = 1$; $a = 20$; $P_1 = 0.03$.]*

Note that σ_3 inverts (or, equivalently, shifts cyclically ; viz. (III.81) on $\mathbb{C}^2$) spin in both 1– and 2–directions. Because the spatial part $|\xi\rangle$ of the wave-packet is initially symmetrical under reflections in both the 1,2– and the 1,3–plane,

(66) $$\sigma_3 \, \boldsymbol{M}(\Delta p_3) \, \sigma_3 \;=\; \boldsymbol{M}(\Delta p_3)$$

with

$$\mathrm{Tr}(\rho \, \boldsymbol{M}(\mathrm{d}p_3)) \;:=\; \mathrm{d}p_3 \int_{\mathbb{R}^2} \mathrm{d}p_2 \, \mathrm{Tr}\Big(\exp(-\mathrm{i}\tau \boldsymbol{H}) \rho^{\otimes} |\xi\rangle\langle\xi| \exp(\mathrm{i}\tau \boldsymbol{H}) \, |p_2;p_3\rangle\langle p_2;p_3|\Big) \;.$$

Here τ is the time the measurement takes. According to (66), the EVM generated by measuring the final p_3–distribution of the particles, is invariant with respect to σ_2–inversions. Thus it represents a non-ideal s_3–measurement, and we can write (using an obvious notation):

(67) $$\boldsymbol{M}(\mathrm{d}p_3) \;=\; \lambda_+(\mathrm{d}p_3) \, |\uparrow\rangle\langle\uparrow| \;+\; \lambda_-(\mathrm{d}p_3) \, |\downarrow\rangle\langle\downarrow| \;.$$

The Schrödinger equation was numerically approximated[viii]. The results, indicating the quality of the measurement, are indicated in fig. 10. As a characterization of quality we choose the indistinguishability of the two λ's in (67), viz.

(68) $$U[\lambda] \;:=\; \int_{\mathbb{R}} \sqrt{\lambda_+(\mathrm{d}p_3) \; \lambda_-(\mathrm{d}p_3)} \;.$$

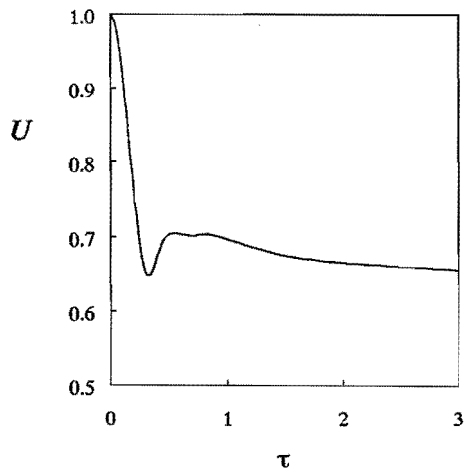


fig. 11 *Quality of the Stern-Gerlach with $a = 0$. The quality of the σ_3 marginal, indicated by the quantity U has been plotted vs. the measurement time τ. Neutral particles are involved. [Initially $\langle \Delta^2 Q_2 \rangle = \langle \Delta^2 Q_3 \rangle = 9$; $a = 0$; $P_1 = 0.03$.]*

The last term of (63) represents spin precession around an axis making an angle $\phi = \arctan(q_2/(a-q_3))$ with the 3-axis. Usually care is taken that the point $(q_2,q_3) = (0,a)$ lies far outside the beam area (i.e. $a \gg \langle \Delta^2 Q_3 \rangle$ and $a \gg \langle \Delta^2 Q_2 \rangle$), so that $1 \gg \phi$. Thus σ_3 is approximately conserved. Moreover, the spatial and spin parts of the total state after measurement are entangled, and cannot be reduced to a simple product, such as (29), or even (29'). These expressions therefore turn out to be not representative of measurement in general. The Stern-Gerlach example shows that even in the absence of such simplifying assumptions, measurement may be fitted into the non-ideality framework.

More interesting from the point of view of the joint measurement problem is the $a = 0$ case (quadrupole field). Then we may rotate our coordinate axes 45^0,

$$q_2 \longrightarrow \tfrac{1}{2}\sqrt{2}(q_2 - q_3)$$

$$q_3 \longrightarrow \tfrac{1}{2}\sqrt{2}(q_2 + q_3) \quad ,$$

giving

$$(63') \qquad H = \frac{1}{2m}(P_1^2 + P_2^2 + P_3^2) - Q_3\sigma_2 - Q_2\sigma_3 \quad .$$

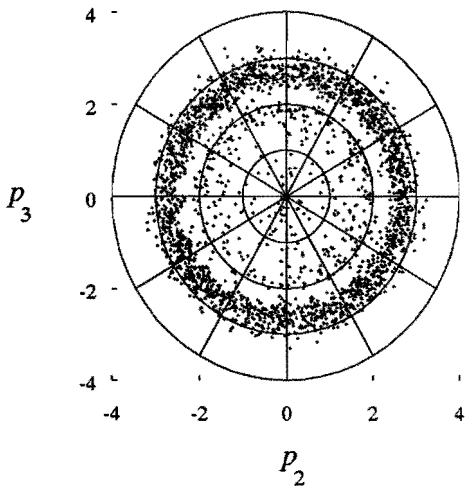


fig. 12 *Typical output of a quadrupole Stern-Gerlach. Neutral particles are involved. Conditions are as in fig. 11, but* $\tau = 2.5$.

This Hamiltonian has two symmetries, viz.

$$[I_3 \otimes \sigma_3]\, H\, [I_3 \otimes \sigma_3]^\dagger = H \;\; ; \;\; [I_2 \otimes \sigma_2]\, H\, [I_2 \otimes \sigma_2]^\dagger = H \;\; . \tag{69}$$

Consequently,

$$\begin{aligned} \sigma_3\, \boldsymbol{M}(\Delta p_2;\Delta p_3)\, \sigma_3 &= \boldsymbol{M}(\Delta p_2;-\Delta p_3) \;\; ; \\ \sigma_2\, \boldsymbol{M}(\Delta p_2;\Delta p_3)\, \sigma_2 &= \boldsymbol{M}(-\Delta p_2;\Delta p_3) \;\; ; \end{aligned} \tag{70}$$

with

$$\mathrm{Tr}(\rho\, \boldsymbol{M}(\mathrm{d}p_2;\mathrm{d}p_3)) := \mathrm{d}p_2\, \mathrm{d}p_3\, \mathrm{Tr}\Big(\exp(-\mathrm{i}\tau \boldsymbol{H})\rho \otimes |\,\xi\rangle\langle\xi\,| \exp(\mathrm{i}\tau \boldsymbol{H}) \; |p_2;p_3\rangle\langle p_2;p_3|\Big) \;\; .$$

Thus the p_2–marginal of $\boldsymbol{M}(\mathrm{d}p_2;\mathrm{d}p_3)$ is invariant under σ_2–inversion, and conversely the p_3–marginal of $\boldsymbol{M}(\mathrm{d}p_2;\mathrm{d}p_3)$ is invariant under σ_3–inversion. Our EVM represents a non-ideal joint measurement of $\mathfrak{s}_2$ and $\mathfrak{s}_3$:

$$\begin{aligned} \int_{\mathbb{R}^2} \boldsymbol{M}(\mathrm{d}p_2,\Delta p_3) &= \mu_+(\Delta p_3)\, |\!\rightarrow\rangle\langle\rightarrow\!| + \mu_-(\Delta p_3)\, |\!\leftarrow\rangle\langle\leftarrow\!| \;\; ; \\ \int_{\mathbb{R}^2} \boldsymbol{M}(\Delta p_2,\mathrm{d}p_3) &= \lambda_+(\Delta p_2)\, |\!\uparrow\rangle\langle\uparrow\!| + \lambda_-(\Delta p_2)\, |\!\downarrow\rangle\langle\downarrow\!| \;\; . \end{aligned} \tag{71}$$

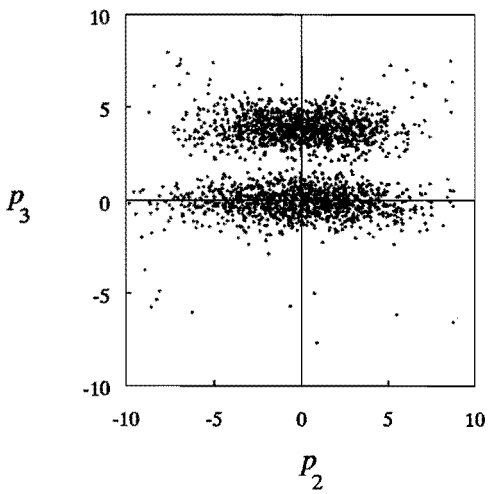


fig. 13 *Typical output of a dipole Stern-Gerlach, for electrons. Conditions are as in fig. 10, but $\tau = 2$. Note that, unlike for neutrals, the centers of the two wave-packets are not symmetrically placed around $p_3 = 0$.*

If we divide the p_2,p_3–plane into its four quadrants, the resulting quadrivariate POVM is, in view of (70), a covariant [§ III.2.6] joint measurement. The qualities of the two marginals (71) are indicated in fig. 11. A sample output distribution is seen in fig. 12.

For electrons the situation is similar to that just discussed, if we take $a = 0$. Then (64) becomes

$$(64') \qquad H = [P_1 - \tfrac{1}{2}Q_2^2 + \tfrac{1}{2}Q_3^2]^2 + P_2^2 + P_3^2 + Q_3\sigma_2 + Q_2\sigma_3 \quad .$$

We still have the two reflection symmetries (69). Thus this situation also represents a joint non-ideal measurement of spin $\frac{1}{2}$ in the two corresponding directions for electrons, just as was the case for neutral particles. We shall now proceed to consider the $a \neq 0$ case. Then, due to the Lorentz-force, we no longer even have the symmetry (65). It is disturbed by the presence of P_1. We do not have an exact non-ideal measurement. But, if the electrons move slow enough in the 1–direction, so that the P_1–term can be neglected, the σ_2–inversion invariance condition is still approximately fulfilled. This is also the case if σ_3 is sufficiently well conserved [cf. (31)]. Accordingly, even this experiment can approximately be considered as a non-ideal

σ_3–measurement (fig. 10). Moreover, as fig. 13 indicates, the beam is cleanly separated in spin-up and spin-down electrons: the upper wave packet contains practically only spin-up electrons, the lower spin-down electrons.

NOTES

i Heisenberg himself probably interpreted the disturbance in the γ-microscope in a more naive way, as scatter after measurement [see a letter to Dirac *d.d.* April 27th, 1927; reprinted in: N. Bohr (1985): *Collected Works*, vol. 6 (ed. by J. Kalckar; North Holland, Amsterdam)]. An interpretation of this imaginary experiment rather more along our lines is given by Reichenbach [H. Reichenbach (1980): *Selected Writings 1909-1953*, vol. 2 (ed. by M. Reichenbach & R. Cohen; Reidel, Dordrecht), p. 215-216].

ii Actually one should demand *complete* positivity instead of (2a) = positivity [Kraus, *op. cit.* (1983); see app. A]. Note that the relation $\longrightarrow$ satisfies the very natural requirement

$$m \longrightarrow n \implies \Xi^\dagger[m] \longrightarrow \Xi^\dagger[n]$$

for any operation Ξ. This is an extra argument for $\longrightarrow$ not mentioned in ch. III.

iii This disturbance concept gives rise to interesting generalizations of *measurement of the first kind* and *destructive measurement* (using the equivalence definition of ch. III [$\mathfrak{M}$ as in (4)]:

An OVM $\mathfrak{M}$ is of the first kind

$$:\Leftrightarrow$$

$$\nu_M^\dagger[n] \leftrightarrow n$$

An OVM $\mathfrak{M}$ is destructive

$$:\Leftrightarrow$$

$$\forall_{\text{EVM } o}\ \nu_M^\dagger[o] \leftrightarrow i$$

Here i denotes the trivial EVM $\{\mathbf{1}\}$.

iv Relations of the type (11) can also be obtained for a non-perfect filter, i.e. if Θ satisfies:

$$\forall_{\rho \mid \rho \geq 0}\ W_{Q,\alpha}(\Theta[\rho]/\mathrm{Tr}\{\Theta[\rho]\}) \leq Dq$$

only for all $\alpha \leq \alpha_0$, for some constant $\alpha_0 < 1$.

v Efficiencies may be as good as .97 (cf. N. Imoto, S. Watkins & Y. Sasaki (1987): *Opt. Comm.* 61, p. 159). For the parameters $\tau\chi_I$ and $\tau\chi_P$ values of the order 10^{-14} are possible (*ibid.*). We have here chosen larger non-linearities, in order to exhibit the quantum effects more clearly. A more realistic treatment can be found in; H. Martens & W. de Muynck (1990): "The uncertainty principle in a QND measurement of photon number", paper to be presented at the International Workshop on *Quantum Aspects of Optical Communications*, Paris (France), proceedings to be published by Springer, Berlin.
Note that inefficiencies in the detectors (D in fig. 4) lead to further non-ideality: the "position" measurement becomes a non-ideal "position" measurement (see H. Martens (1987): "Gelijktijdige Meting van Incompatibele Observabelen" [Eindhoven University of Technology, Dept. of Theoretical Physics, Internal Report 1987-18]).

vi More accurately, the device realizes a non-ideal measurement of the PVM corresponding to $\sin(4\tau\chi_I N)$. For low S-intensities, this is approximately equivalent to a non-ideal measurement of N.

vii A better known description of phase is due to Carruthers & Nieto (P. Carruthers & M. Nieto (1968): *Rev. Mod. Phys.* 40, p. 411), who use self-adjoint cosine and sine operators C and S. They are related to e [defined in (52)] by $e = C + iS$. Although these operators, being Hermitian, are perhaps more familiar objects than EVMs, they have rather undesirable properties. They are, e.g., incompatible, and satisfy no straightforward commutation relation with N, unlike e:

$$\exp(iaN)\, e^b = e^b \exp(iaN) \exp(-iab) \quad (a \in \mathbb{R},\, b \in \mathbb{N})$$

[cf. the pair (III.82)]. Thus we have chosen Levy-Leblond's solution.

viii The calculations were done using a straightforward integration procedure of second order in position, and first order in time (implicit). The data are:

fig. 10, neutrals (solid):	Mesh width $(q_2 \times q_3)$ 20 by 50, mesh resolution 64 by 64 points; $\Delta\tau = 0.002$
fig. 10, electrons (dashed):	Mesh width $(q_2 \times q_3)$ 5 by 50, mesh resolution 32 by 128 points; $\Delta\tau = 0.0003$
fig. 11, neutrals:	Mesh width $(q_2 \times q_3)$ 35 by 35, mesh resolution 64 by 64 points; $\Delta\tau = 0.0025$

CHAPTER V

Evaluation

As sketched in the introduction and in ch. I, the status of the Uncertainty Principle (UP) is somewhat ambiguous. On the one hand there are a number of mathematically well-defined relations. These limit statistical dispersion. On the other hand, 'the Uncertainty Principle' refers to a cluster of concepts of much wider significance. It is perhaps natural to turn first to Niels Bohr's writings when stuck with problems regarding the meaning of the UP. For Bohr the QM formalism is *only* of instrumentalistic significance. Its elements, wave functions, density matrices, etc., are only computational tools without further physical significance. They do not correspond to physical object properties. Bohr thinks we can only comprehend in terms of CM. Thus, whereas *quantitative* calculation of a physical situation involves QM, *qualitative* understanding proceeds along the lines of CM. Understanding can therefore not be completely definite; it must leave room for probabilistics. To put it somewhat extremely, calculation and interpretation are always in accord without there being a direct connection, like perfectly synchronized watches. This is in contrast to more conventional approaches to interpretation, where the elements of the formalism are interpreted. The interpretation of concrete situations is then guided directly by their formal description.

Bohr's insistence on the impossibility of extending the CM conceptual framework is essential to complementarity, and at the same time its major weakness. Attempts to show discrepancies between complementarity and QM, or inconsistencies in complementarity itself seem futile[1]. But surely the thesis that CM concepts are somehow the only concepts in which we can visualize, is highly dubious. The classical concepts themselves have evolved over a long time (Aristotelians knew no 'momentum', and even a century before Bohr such "natural" concepts as 'electrical field' were not natural at all), so that there appears to be no reason for assuming that our conceptual apparatus cannot be refined any further[2].

Indeed Bohr's attitude can have a number of unfortunate consequences. As we saw (ch. II, app. B) Bohr considered the spin of free electrons to be unmeasurable.

---

[1]For instance Popper [K. Popper (1982): *Quantum Theory and the Schism in Physics* (ed. by W. Bartley; Rowman & Littlefield, Totowa, NJ)] proposes an experiment to decide between "the Copenhagen interpretation" and QM.

[2]See e.g. P. Feyerabend (1958): *Proc. Arist. Soc.* (suppl. Vol.) **32**, p. 83.

Originally he came to this conclusion because he thought that spin was a fundamentally quantum mechanical quantity without a classical analog, it being in magnitude proportional to ħ. Subsequent discussions brought him to a more subtle point of view, widely reproduced in the literature[3]. But even that viewpoint is incorrect: spin measurement is for electrons, due to charge effects, undoubtedly more difficult than for neutral atoms, but by no means excluded for fundamental reasons (ch. IV). Bohr's dislike of real QM calculations perhaps caused him not to actually perform the analysis of the spin-measurement quantitatively. Such an analysis would have put the matter beyond philosophical dispute.

Bohr undoubtedly gave the UP a much wider content than that of the Heisenberg-type relations. But, as could be expected, he never substantiated this wider relevance, apart from philosophical arguments and discussions of a few highly idealized situations. Moreover, Bohr's discussions are phrased in such vague and general terms that they do not specify exactly which types of non-Heisenberg uncertainty relation actually hold, or can be expected to hold. Bohr's discussions do not offer any starting points for a rigorous derivation of new forms of the UP. More generally, the extreme difficulty of applying complementarity in non-toy situations[4] severely restricts the value of complementarity from a physical point of view, as opposed to the philosophical point of view.

Thus Bohr's philosophy offers little inducement for a further investigation into the gap between formal and intuitive meaning of the UP. Moreover, as measurements have, until recently, been very far away from the (supposed) quantum limits[5], they could usually be treated (semi–)classically, ignoring the finer points of quantum measurement theory. A closer investigation of the nature of quantum bounds was not imperative.

---

[3]E.g. in N. Mott & H. Massey (1965): *The Theory of Atomic Collisions* (Clarendon Oxford). See also app. B.

[4]Cf. even Pauli's and Bohr's difficulties in defending Bohr's Stern-Gerlach argument against ever more complicated devices [O. Darrigol (1985): *Hist. Studies Phys. Sc.* **15**, no. 1, p. 39] and his papers with Rosenfeld on field theory [N. Bohr & L. Rosenfeld (1933): *Mat.–Fys. Medd. Dan. Vidensk. Selsk.* **12**, no. 8; (1950): *Phys. Rev.* **78**, p. 794].

[5]Cf. Jammer (1974): *The Philosophy of Quantum Mechanics* (Wiley, NY), p. 82

But in recent years measurement quality has improved to such an extent that quantum measurement theory has become relevant. A case in point is the discussion surrounding the so-called "Standard Quantum Limit" for position tracking (SQL; see ch. IV). Here the ever-increasing demand for position measurement accuracy for gravitational wave detectors, instigated a closer investigation of the properties of such accurate meters. The intuitively plausible interpretation of the UP as a limit to the accuracy in joint measurement of incompatible observables, together with seemingly natural requirements on the back-action of the instrument on the object, led to the conclusion that there exists an absolute lower limit to the accuracy with which the position of a quantum object can be monitored: the SQL. But intuitive reasoning, because of its inherent vagueness, is of limited value. The confusion in the ensuing discussion on the SQL clearly evidences this. Accordingly, closer formal investigations seem appropriate to clear up such matters.

Bohr's philosophy gives, as we saw, little hold in formal investigations. It is, however, not the general attitude towards QM. Textbook QM, despite (frequent) use of Bohr's name, actually has little conceptual overlap with QM according to Bohr. People like von Neumann and Dirac, rather than Bohr, should be regarded as originators of the attitude towards QM these books have. On the one hand, the emphasis is strongly towards instrumentalism: the larger part of the textbooks is devoted to the development of algorithms for specific calculations, without bothering too much about what they mean. On the other hand, interpretation (where it is involved) tends towards the other extreme, i.e. quantum realism. Things like wave functions and/or density matrices are discussed as if they were object properties. The interpretational side is, however, hardly coherent: in other places "quantities" (momentum, angular momentum, position, etc.) are uncritically used in a classical way. As regards the UP, many textbooks present the narrow mathematics alongside the wide physics, without noticing the discrepancy. The narrowness of the mathematics is the result of the formalism's possibilities. Because of interpretational prejudices, this was set up (axiomatized) so as to contain only a very restricted class of measurements (ch. I). This formalism blocked the way towards investigations of the UP going beyond the Heisenberg-type relations.

Nevertheless the von Neumann-Dirac approach, when rid of unnecessary interpretational ballast (i.e. when sufficiently generalized, see ch. III & IV), offers a suitable context for a discussion of the full significance of the UP. There are two ways of connecting macro-devices with QM objects. We may first attempt to produce, or prepare, quantum systems with "properties" approaching some prescribed characteristics. We may try to *control* certain aspects of the object. Secondly, we may attempt to find out something about certain object "properties". We may try to *determine* certain aspects of a quantum object. The first connection may, somewhat oversimplifyingly, be said to refer to the future, the second to the past. The two connections are represented in the generalized framework by two different entities (density matrices and effect-valued measures, respectively; see ch. III). Of course most devices will combine both types of connection to some extent, e.g. when a meter does not only measure some "property", but also, depending on the measurement outcome, influences the object state accordingly (corresponding to a representation by an operation-valued measure in the formalism; see ch. IV).

In the generalized von Neumann-Dirac scheme the inaccuracy interpretation of the UP, more properly called the inaccuracy principle, can be mathematically derived (ch. III). Such a derivation determines fixes the 'inaccuracy' notion it involves. Together with the better-known scatter relations, *two* incarnations of the UP are then justified. These two incarnations are of general validity. Hence they are certainly applicable to special sub-classes of measurement devices. For instance, a quantum source combined with subsequent devices further manipulating the quantum object may be considered as a (bigger) quantum source. Similarly, a non-destructive measurement combined with subsequent further determinations of object "properties", may be considered a larger quantum meter. To such larger devices the UP is applicable. Thus a lower bound to the amount of disturbance caused by a measurement may be derived from the impossibility of using the measurement in combination with subsequent measurements to violate the inaccuracy principle. Analogous reasoning holds true for the influence of quantum object manipulations. These applications of the UP in its two forms show two types of disruption of the connection between past and future development of micro-systems by the intervention of macro-systems.

In a similar vein, the well-known Wigner-Araki-Yanase theorems can be incorporated into the UP. In this way we have, besides achieving some unification, reduced the conceptual vagueness surrounding most forms of the UP thus far. They have become suitable for concrete applications, as can be illustrated by the SQL discussion (see above). An "SQL" of sorts can be derived from the inaccuracy principle. Unfortunately the plausible assumption in the usual SQL derivation regarding instrument back-action[i], turns out to be invalid as a general property of quantum measurement. As a result, the SQL that does hold is extremely weak. In particular no absolute upper limit to gravitational wave detector sensitivity *via* position measurement can be derived. Indeed sufficiently cleverly chosen measurement operations[6] can be shown to increase this sensitivity arbitrarily.

The role of the UP in QM is highly analogous to that of the Lorentz contraction in relativity theory (RT). Just as 'Lorentz contraction' refers to an element of RT that is surprising only from the point of view of CM (the preceding theory), so does the UP highlight limitations to the micro-macro connections that appear new from the point of view of CM. The two forms of the uncertainty principle, corresponding to the two types of micro-macro connections, are only restrictions when seen from the CM point of view. Quantum preparations and measurements actually show much more variation than their classical counterparts, corresponding to more varied behavior of quantum objects (e.g. interference with material particles). In a similar vein the contraction as a physical effect appears to vanish when seen from RT ("is Lorentz contraction 'real' ?"). And like the contraction does not need a force causing it, the UP is of a logical nature, not the consequence of uncontrollable disturbances[7]. The analogy can be pursued even further. When phrasing RT in terms of Minkowskian space-time, the objects in the theory (4–vectors, etc.) simply do not allow the description of an object that does not contract. Thus the choice of conceptual forms suitable to the theory under consideration, reduces aspects that appeared surprising from a classical point of

[6]H. Yuen (1984): *Phys. Rev. Lett.* 52, p. 1730. See also other references in ch. IV.

[7]Of course this situation would change drastically if some neo-classical theory would turn out to be behind QM, but this remains speculation at this time.

view, to non-effects. Similarly, the QM formalism simply cannot describe macro-devices that do not satisfy the UP. The derivation of the two forms of the UP *directly* from the formalism reflects this. It would appear then that conceptual forms suitable for the interpretation of QM would make a device that does not satisfy the UP simply unimaginable. Classical concepts would retain their validity only in limiting cases, and then only approximately. But since the interpretation of QM in general is not the subject of this thesis, we will end the discussion of this highly controversial issue here.

The results in ch. III and IV were achieved without recourse to models. Results obtained by model analysis are not necessarily generally valid. Like measurement models, restrictions on measurement (unbiasedness, covariance, etc.; see ch. III)[ii] may aid in the obtaining of results. But these restrictions should be seen for what they are: they were introduced for reasons of convenience. Thus any results derived with their help, like those obtained by model analysis, can only be of temporary significance. As soon as results can be achieved without (some or all of) the restrictions, they should be dropped. Unfortunately, restrictions that were originally introduced for convenience, to facilitate subsequent reasoning, tend to lead a life of their own. The results in this thesis show the importance of viewing things from the most general point of view that is practicable.

NOTES

i The classical spirit is responsible for the overly restrictive measurement notion of text-book QM, as we saw in ch. I. In particular, the classical measurement ideal is behind the 'measurement of the first kind' transformation. The requirement referred to here (see ch. IV) is a generalization, connecting measurement accuracy to scatter-after-measurement, probably inspired by the idea that if a measurement of the first kind is not realizable, one should at least try to approximate it.

ii In particular use of the generalizations of first kind measurements, such as those referred to in note i, strongly affects the generality of the results obtained with them, precisely because such transformations are not representative.
Note further that the restriction to finite dimensional spaces is not a great drawback from a pragmatic point of view because, as already said in ch. III, infinite dimensional results can be approximated arbitrarily well by results on finite dimensional spaces.

APPENDICES

Appendix A
Some Mathematics

1 Convexity

A subset $\mathscr{R}$ and a finite subset $\mathscr{S} = \{s_j\}_M$ (M = {1, ..., m}) of an n-dimensional real linear space $\mathscr{E}$ are given. Then[1]

DEFINITION Let $\{p_j\}_M$ be a set of non-negative real numbers satisfying $\sum_{j \in M} p_j = 1$. Then $r = \sum_{j \in M} p_j s_j$ is a *convex combination* of the s_j's with coefficients (weights) $\{p_j\}_M$.

DEFINITION The *convex span* $C(\mathscr{S})$ of $\mathscr{S}$ is the collection of all convex combinations of the elements of $\mathscr{S}$.

DEFINITION The *convex hull* $H(\mathscr{R})$ of $\mathscr{R}$ is the collection of all convex combinations of finite subsets of $\mathscr{R}$.

DEFINITION $\mathscr{R}$ is *convex* if $H(\mathscr{R}) = \mathscr{R}$.

DEFINITION An element r of a set $\mathscr{R}$ is *extreme* if it belongs to $\mathscr{R}$, but does not lie between any two points of $\mathscr{R}$. The set of extreme elements of a convex set $\mathscr{R}$ is denoted by $\partial_E(\mathscr{R})$.

Then[2]:

THEOREM If $\mathscr{R}$ is a convex, compact (i.e. bounded and closed) set, then $H(\partial_E(\mathscr{R})) = \mathscr{R}$.

---

[1]P. Kelly & M. Weiss (1979): *Geometry and Convexity* (Wiley, NY); A. Holevo (1982): *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam).

[2]Kelly & Weiss, *op. cit.*

THEOREM Each $h \in H(\mathcal{R})$ is a convex combination of at most n + 1 points of $\mathcal{R}$.

The first of these theorems is sometimes referred to as Minkowski's theorem, the second is Carathéodory's theorem. Extension of these theorems on infinite dimensional spaces is not quite trivial[3].

2 Extended Quantum Formalism

Suppose we have two preparators that can feed objects into a measurement device. If we use the first preparator with probability λ, and the second with probability $(1-\lambda)$, it is clear that we will see outcomes corresponding to the first preparator with probability λ, and outcomes corresponding to the second with probability $(1-\lambda)$. In other words, the outcome probability distribution of the mixed preparator is given by a convex combination of those of the two preparators. Given that preparators are represented by density operators ρ on a Hilbert space $\mathcal{H}$, and that mixing corresponds to convex combinations of ρ's, measurement outcome probability distributions must be generated by affine functionals of ρ:

DEFINITION The functional f over a convex set $\mathcal{R}$ is *affine* if

$$f[\lambda r_1 + (1-\lambda)r_2] = \lambda f[r_1] + (1-\lambda)f[r_2] \tag{1}$$

for all $r_1, r_2 \in \mathcal{R}$ and $\lambda \in [0,1]$.

All affine functionals over the density operators can be extended to linear functionals over the trace class operators[4] $\mathcal{T}(\mathcal{H})$. Since the set of bounded operators $\mathcal{B}(\mathcal{H})$ is the dual of the set of trace class operators, there is a 1–1 correspondence between

[3]See e.g. N. Dunford & J. Schwartz (1957): *Linear Operators*, part I (Interscience, NY); V. Klee (1957): *Arch. Mat.* 8, p. 234, and references in Holevo, *op. cit.*.

[4]Holevo, *op. cit.*; cf. Kelly & Weiss, *op. cit.*, p. 89.

affine functionals and bounded operators. Thus the affinity criterion together with natural requirements on probability, leads immediately to the effect notion (III.27). Similarly the operation notion (IV.2) can be derived. More technically, the relevant definitions are[5]

DEFINITION Consider an n–dimensional Hilbert space $\mathscr{H}_n$ and a linear mapping ϕ of $\mathscr{T}(\mathscr{H})$ into $\mathscr{T}(\mathscr{H})$. The Hilbert space $\mathscr{H}\otimes\mathscr{H}_n$ consists of n×n matrices (τ_{ij}) of $\mathscr{T}(\mathscr{H})$ elements. Then ϕ can be extended into a map $\phi_n:(\tau_{ij}) \rightarrow (\phi[\tau_{ij}])$ of $\mathscr{H}\otimes\mathscr{H}_n$ into $\mathscr{H}\otimes\mathscr{H}_n$. Now, ϕ is called *n–positive* if ϕ_n is positive [(IV.2.a)]; it is called *completely positive* if it is n–positive for all n.

DEFINITION An *operation* ϕ is a completely positive mapping of $\mathscr{T}(\mathscr{H})$ into $\mathscr{T}(\mathscr{H})$, satisfying (IV.2.b). If $\mathrm{Tr}(\phi[\tau]) = \mathrm{Tr}(\tau)$ for all τ in $\mathscr{T}(\mathscr{H})$, it is *non-selective*.

Operations can be represented in terms of operators in the following way[6] (cf. (IV.53)):

THEOREM For any operation ϕ there exist operators $\{A_k\}_K$ (K finite or countably infinite) on $\mathscr{H}$ satisfying

(2) $$\sum_{k\in K_0} A_k^\dagger A_k \leq 1 \quad \text{for all finite subsets } K_0 \text{ of } K,$$

such that for all τ in $\mathscr{T}(\mathscr{H})$ and B in $\mathscr{B}(\mathscr{H})$

(3) $$\phi[\tau] = \sum_{k\in K} A_k \tau A_k^\dagger \quad ; \quad \phi^\dagger[B] = \sum_{k\in K} A_k^\dagger B A_k \ .$$

The effect F corresponding to ϕ is then given by

(4) $$F = \sum_{k\in K} A_k^\dagger A_k \ .$$

---

[5]E. Davies (1976): *Quantum Theory of Open Systems* (Academic, London); K. Kraus (1983): *States, Effects and Operations* (Lecture Notes in Physics **190**, Springer, Berlin); Holevo, *op. cit.*

[6]Kraus, *op. cit.*

(If K is infinite, the two latter summations converge ultraweakly, whereas the first sum in (3) converges in the trace norm topology). Conversely, any set $\{A_k\}_K$ satisfying (2) defines an operation via (3).

Naĭmark's theorem III.2 can also be phrased in another form[7], viz.

THEOREM Any EVM $\{M(da)\}_A$ on $\mathcal{H}$ can be written as

$$M(da) = E' \, E(da) E' \quad , \tag{5}$$

where $\{E(da)\}_A$ is a PVM on a Hilbert space $\mathcal{H}' \supseteq \mathcal{H}$, and E' is the projector of $\mathcal{H}'$ onto $\mathcal{H}$.

Operations can also be extended onto larger Hilbert spaces, like in the following theorem[8]:

THEOREM For any two operations ϕ and $\bar{\phi}$, such that $\phi + \bar{\phi}$ is non-selective, there exist a Hilbert space $\mathcal{H}'$, density operator ρ' on $\mathcal{H}'$, effect F' on $\mathcal{H}'$ and unitary operator U on $\mathcal{H} \otimes \mathcal{H}'$ such that

$$\phi[\tau] = \mathrm{Tr}_{\mathcal{H}'}[(1 \otimes F')U(\tau \otimes \rho')U^\dagger] \quad , \tag{6}$$

$$\bar{\phi}[\tau] = \mathrm{Tr}_{\mathcal{H}'}[(1 \otimes (1' - F'))U(\tau \otimes \rho')U^\dagger] \quad , \tag{7}$$

for all τ in $\mathcal{T}(\mathcal{H})$.

Again, the converses of these two theorems also hold: relations (5) and (6)&(7) define EVMs and operations, respectively.

---

[7]Holevo, *op. cit.*

[8]Kraus, *op. cit.*; M. Ozawa (1984): *J. Math. Phys.* **25**, p. 79

3 Stochastic Matrices

We consider square matrices $\{\lambda_{jk}\}_{K\times K}$ over some countable index set K. They are required to satisfy

(8) $$\forall_{j,k\in K}\ \lambda_{jk} \geq 0 \quad ;$$

(9) $$\sum_{j\in K} \lambda_{jk} = 1 \quad .$$

If the set K labels physical state, λ_{jk} may be interpreted as the probability that the system will at time t_{n+1} be in state j, if it was at time t_n in state k. The matrix is interpreted as a set of transition probabilities. These are independent of the time and of the history of the system: we speak of a *Markov chain* with stationary transition probabilities[9]. Depending on the properties of the matrix, we can discern different classes of states. First introduce the n–step transition probabilities $\lambda^{(n)}_{\ell k}$:

(10) $$\lambda^{(1)}_{jk} := \lambda_{jk} \ ; \ \lambda^{(n+1)}_{jk} = \sum_{\ell\in K} \lambda_{j\ell}\, \lambda^{(n)}_{\ell k} \quad .$$

DEFINITION Consider the class of all n such that $\lambda^{(n)}_{kk} > 0$. The greatest common divisor of these n's is called the *period*, and denoted by d_k. If there are no such n's, the period is undefined.

DEFINITION A nonempty subset $J \subseteq K$ is *closed* if

(11) $$\forall_{j\in J}\ \sum_{j\in J} \lambda_{jk} = 1 \quad .$$

J is called *minimal closed* if it has no proper subsets that are closed.

[9] K. Chung (1967): *Markov Chains with Stationary Transition Probabilities* (2nd ed., Springer, Berlin)

The total set K can then be divided into classes, according to the elements' period. All elements in a minimal closed set have the same period. The Perron-Frobenius theorem[10] implies

THEOREM Let J be finite and minimal closed. Then $\{\lambda_{jk}\}_{J\times J}$ has a simple eigenvalue 1. The corresponding eigenvector $(x_k)_J$ may be chosen such that it consists of non-negative elements. Thus there is one unique probability distribution $(p_k)_J$ that is mapped onto itself by $\{\lambda_{jk}\}_{J\times J}$

More generally[11],

THEOREM Let J be a minimal closed set of period d. Then the solution of the equation

(12) $$x_j = \sum_{k\in J} \lambda_{jk} x_k \quad (j \in J) \ ,$$

such that $\sum_{k\in J} x_k < \infty$, is unique up to a multiplicative constant.

[10]J. Ortega (1987): *Matrix Theory* (Plenum, NY)

[11]Chung, *op. cit.*, p. 35

Appendix B
Ad chapter II

The essential points of complementarity were (ch. II):

(i) Objectifying description in terms of the quantities from CM: the classical quantities, in as far as they are well-defined [cf. (ii)], are object-properties[1].

(ii) The measuring instrument must be described completely classically; the uncertainty principle is not relevant for its working. The unanalyzability (~ indivisibility) of the object-meter system is symbolized by the uncertainty principle. This principle shows that well-definedness of some classical quantities in the interaction inevitably leads to unanalyzability in others. This unanalyzability may, for instance, appear in the guise of an "uncontrollable momentum exchange" [2]. The precise nature of the measuring instrument determines how well-defined the quantities in the interaction and, *as a consequence*, those describing the object [cf (i)] are.

(iii) Wave particle duality plays no role. Electrons are particles and light consists of waves[3].

(iv) The QM formalism, used for quantitative calculation, is unvisualizable (~ ununderstandable). It is only of symbolic (~ instrumentalistic[4]) value.

---

[1]Cf. Hooker [C. Hooker (1972): in *Paradigms and Paradoxes* (ed. by R. Colodny, Pittsburgh University Press), p. 67], p. 75 and p. 134 (B1).

[2]Cf. Hooker, *op. cit.*, p. 135

[3]Cf. D. Murdoch (1987): *Niels Bohr's Philosophy of Physics* (Cambridge University Press), p. 70

[4]P. Feyerabend (1981): *Realism, Rationalism and Scientific Method*, vol. 1 (Cambridge University Press), p. 258

The exposition in the following is intended *only* to justify above four points. As a consequence the treatment of certain imaginary experiments, most notably the Einstein-Podolsky-Rosen experiment, is concise: an evaluation of all of (or even some of) the issues surrounding these experiments has not been attempted, and the discussion is therefore caricatural in certain respects.
Besides Bohr's own papers, use has been made primarily of an essay by Hooker[5].

1 JUSTIFICATION

When reading Bohr, it is important that one keeps the development of his views in mind. That such a development took place is evident (and hardly surprising): in the Como lecture[6], the first two points are present, although the second one only occurs in an embryonal form: certain formulations in the Como lecture may suggest an interactionalistic view on the uncertainty principle. The third point, however, is absent in 1927: Bohr still sticks to the concept of wave-particle duality from old QM, and goes even so far as to suggest wave-particle complementarity[7] [see ad (iii) further below]. Because of the development in Bohr's views[8] his famous 1949 Einstein essay[9] is also somewhat difficult to interpret. This essay is intended to give an account of the discussions with Einstein, and thus follows the historical course of

---

[5]Hooker, *op. cit.*

[6]N. Bohr (1927) [**Como Lecture**]: *Atti del Congresso Internazionale dei Fisici 1927, Como-Pavia-Roma* (Nicola Zanichelli, Bologna), p. 565. The Como lecture was, with some minor changes [see N. Bohr (1985): *Collected Works*, vol. 6 (ed. by J. Kalckar, North Holland, Amsterdam)] also published in Nature: N. Bohr (1928) [**Nature paper**]: *Nature* (Suppl.) **121**, p. 580.

[7]Bohr (1928) p. 581: "[In the] problem of the nature of the constituents of matter [...] we are not dealing with contradictory, but with complementary pictures of the phenomena."

[8]Bohr himself acknowledged this development [see e.g. N. Bohr (1929c) [**ATDN Introduction**]: p. 1 of N. Bohr (1934): *Atomic Theory and the Description of Nature* (Cambridge University Press)], so that he can hardly be accused of dogmatism [Feyerabend, *op. cit.*, p. 293].

[9]N. Bohr (1949) [**Einstein essay**]: p. 201 of the Schilpp volume [P. Schilpp (ed.) (1949): *Albert Einstein, Philosopher-Scientist* (Open Court, Evanston IL) [reprinted on p. 9 of J. Wheeler & W. Zurek (eds.) (1983): *Quantum Theory and Measurement* (Princeton University Press)], quotations from the 1983 reprint.

development. As a result, it is not always clear whether the opinions given in the essay are his views at the time of the episode he is describing, or his 1949 views.

It is commonplace[10] to claim that the Einstein-Podolsky-Rosen (EPR) paper (see § 4) brought about a major change, even a break, in Bohr's ideas. I do, however, not see any real evidence for this[11]. It seems especially that the interactionalist doctrine, allegedly advocated by Bohr prior to EPR, is more due to Heisenberg[12]. Bohr, though admittedly wording the Como lecture in an interactionalistic way[13], takes already well before[14] 1935 the limitation on the applicability of classical concepts as primary and the uncontrollable interaction as a reflection of this (rather than *vice versa*)[15]: "[The uncertainty principle] defines the latitude in the application of classical concepts [...]" and[16] "We cannot hope to overcome this limitation of applicability because it lies absolutely embodied in the process of definition of space and time". The interactionalist misunderstanding of Bohr is not new, already in 1928 the editorial comment on Bohr's Nature paper[17] blames the "vagueness of position and path of an electron" on the necessity of employing a disturbing interaction in a quantum measurement. Thus the effect of EPR was probably a break in Bohr as people understood him, rather than a (radical) change in the views of Bohr himself[18].

In the development of Bohr's views the effect of his discussions with Einstein (few as they were) has, I think, been overestimated at the cost of especially the developments in particle physics in the early thirties. A substantial part of these latter developments indeed took place in Bohr's own institute, and they must have occupied Bohr. In

---

[10]See e.g. M. Jammer (1974): *The Philosophy of Quantum Mechanics* (Wiley, NY) or K. Popper (1972): *The Logic of Scientific Discovery* (6th. rev. impr., Hutchinson, London)

[11]Cf. Hooker, *op. cit.*, p. 149ff. A change as regards the wave-particle issue took place ***before*** EPR [see ad (iii)].

[12]Hooker, *op. cit.*, p. 157

[13]But cf. Bohr, *op. cit.* (1928) p. 582

[14]Cf. N. Bohr (1930) **[Faraday Lecture]**: *J. Chem Soc* **26** (1932), p. 349

[15]Bohr, *op. cit.* (1930) p. 351

[16]N. Bohr (1931b) **[Wills Lecture]**: p. 361 of Bohr, *op. cit.* (1985)

[17]*Nature* (Suppl.) **121**, p. 579 (1928); See also Jammer, *op. cit.*, p. 160 ff.

[18]Hooker, *op. cit.*, p. 149

particular the development of complementarity from "just" an interpretation of QM to a framework of more general relevance probably occurred in this period: at several times Bohr thought that the end of QM was in sight, and he must have thought about the fate of complementarity in this light. The wider methodological applicability of complementarity was one of its major characteristics, as we saw in ch. II.

We shall first give some concrete textual evidence for our reading of complementarity. We shall then discuss some characteristic Bohrian thought experiments. [Quotations in the following denoted by "WZ" are from the Einstein essay, in the Wheeler & Zurek, *op. cit.* (1983) reprint.]

ad (i) For Bohr QM is a generalization of CM[19] (the correspondence principle !; cf. ch. I). Accordingly, in his work the insistence on classical concepts as the means of *understandable* description (as opposed to *symbolic* description, see below) is almost ubiquitous, and epitomized by the remark that[20] "the language of Newton and Maxwell will remain the language of physicists for all time", already quoted in ch. II. Classical concepts are necessary because *only* in classical terms can we conceptually comprehend something. For Bohr it is a mistake to think that new forms of thought should be developed in order to obtain an understanding of QM[21].

With the classical concepts the objectifying mode of description is naturally brought along: Bohr uses it without qualification[22] in phrases like "[...] a particle which has momentum p [...]" (WZ p. 21), "[...] location

---

[19]His favorite phrase is "rational generalization", e.g. on p. 2 of N. Bohr (1958): *Philosophy in the Mid-Century* (ed. by R. Klibansky, La Nuovo Italia, Florence) p. 508 [reprinted on p. 1 of N. Bohr (1963): *Essays 1958-1962 on Atomic Physics and Human Knowledge* (Wiley, NY)]. Quotations here and in the following from the 1963 reprint version.

[20]N. Bohr (1931a) [**Maxwell Lecture**]: *Nature* (Suppl.) **128**, p. 691

[21]E.g. Bohr, *op. cit.* (1929c) p. 16: "[...] it would be a misconception to believe that the difficulties of the atomic theory may be evaded by replacing the concepts of classical physics by new conceptual forms". Here Bohr also speaks of QM as a "rigorous reinterpretation" (p. 18) of CM.

[22]Therefore I cannot concur with Murdoch [Murdoch, *op. cit.*, p. 102] that Bohr nowhere insists that microphysical systems also have to be described classically. In fact Bohr never described them otherwise. Especially Bohr's views on the Stern-Gerlach (§ 3) show the importance he attaches to to classical concepts even in the microdomain.

of the particle in the plane [...]" (WZ p. 22), "[...] the time an electron takes [...] to come from the diaphragm to some other place [...]" (WZ p. 29), "[...] the energy of the particle [...]" (WZ p. 29).

But there is one important limitation: *the objectivism reaches only as far as the classical quantities are well-defined.* The particular measuring arrangement under consideration determines how well-defined the quantities are [see ad (ii)]. Therefore complementarity only ends the *full* and uncontextual applicability of classical description, not its applicability *per se* or its objectifying character. The lack of a *full* applicability, however, leads to the impossibility of a classically complete description, and thus entails the recourse to statistics that is also reflected by the Heisenberg relations[23].

ad (ii) Again Bohr has been clear about this issue: already in 1929 he talks about[24] "the invocation of classical ideas necessitated by the very nature of measurement", and in 1930 of[25] "the classical mechanical concepts on which all measurements must be interpreted". The parts constituting the meters are so large and massive that the uncertainty principle is *practically irrelevant* for their working[26], despite the QM nature of the particles that ultimately constitute the device. But as regards their description (epistemology) he emphasizes the full applicability of CM to the measuring device. Bohr goes so far as to demand that the uncertainty principle

---

[23]N. Bohr (1939) [**Warsaw lecture**]: *New Theories in Physics* (International Institute of Intellectual Co-operation, Paris), p. 11. See p. 19. Cf. also Bohr, *op. cit.* (1958)

[24]N. Bohr (1929b) [**Scandinavian Natural Scientists Congress Lecture**]: p. 102 of Bohr, *op. cit.* (1934), see esp. p. 114.

[25]N. Bohr, *op. cit.* (1930) p. 377. Cf. Bohr, *op. cit.* (1931a).

[26]In the EPR answer (N. Bohr (1935) [**EPR answer**]: *Phys. Rev.* **48**, p. 696), when talking about test-bodies with which he intends to measure a diaphragm's momentum, he argues that their momentum has to be controlled. "It is true that such a control will essentially depend on an examination of the space-time course of some process to which the ideas of classical mechanics can be applied; if, however, all spatial dimensions and time intervals are taken sufficiently large, this involves clearly no limitation as regards the accuracy of the control of their space-time coordination" (p. 698; a similar remark occurs in the following sentence). Cf. Bohr, *op. cit.* (1929c) p. 11 and Jammer, *op. cit.*, p. 207

is *fundamentally inapplicable* to it. Thus, for example, he says in the Warsaw lecture[27]: "[...] a measurement can mean nothing else than the unambiguous comparison of some property of the object under investigation [cf. (i)] with a corresponding property of another system, serving as a measuring instrument, and for which this property is directly determinable according to its definition in everyday language or in the terminology of classical physics. [...] the necessity of basing the description of the properties and manipulation of the measuring instruments on purely classical ideas implies *the neglect of all quantum effects in that description* [my italics], and in particular the renunciation of a control of the reaction of the object on the instruments more accurate than is compatible with the [uncertainty relations]". In the Einstein essay we find: "The special aim of the figure is to underline that a clock is a piece of machinery, *the working of which can be completely accounted for by ordinary mechanics* [my italics] and will be affected neither by a reading of its hands, nor by the interaction between its accessories and an atomic particle." (WZ p. 28), "[...] pertaining to *individual* [~ indivisible] phenomena appearing under conditions described by classical physical concepts" (WZ p. 46)[28]. The fully classical description of the measuring instrument is, for Bohr, an epistemological need. It is a necessary condition for the communication of the measurement results to others[29].

The fully classical description of the meter, together with the indivisibility of the quantum of action, *causes* the impossibility of a joint application of all classical concepts to the object. The elementary character (~ indivisibility) of the quantum of action is a reflection of the complementarity of the pictures involved in the description of the interaction [cf. ad (i)]. Any attempt to *further* analyze the interaction can only render the instrument useless (WZ p. 29)[30] (cf. the discussion of single and double slit in ch. II). Thus the need for classical description of

---

[27]Bohr, *op. cit.* (1939) p. 19; cf. the quote from the EPR-answer in the previous footnote.

[28]See also Bohr, *op. cit.* (1939) p. 28 and Bohr, *op. cit.* (1958)

[29]Bohr, *op. cit.* (1958) p. 3

[30]Bohr, *op. cit.* (1935) p. 697 and p. 698

the measuring instruments is, for Bohr, not in contradiction with the QM nature of the components constituting the device, or with use of the QM formalism for quantitative results as regards the outcome probabilities in different situations (WZ p. 31)[31] [cf. ad (iv)].

The fact that the quality of definition of the object properties depends on the definability of the interaction with the meter indicates an operationalistic element in Bohr's philosophy. This is also reflected by his opinion that the purpose of[32] "[...] the description of nature [...] is not to disclose the real essence of the phenomena, but only to track down, so far as it is possible, relations between the manifest aspects of our experience". But Bohr (contrary, perhaps, to Heisenberg; see ch. I) is not an operationalist: the concepts themselves are *not* defined by measuring arrangements (ch. II), only the extent to which they are well-defined in a particular situation is determined by the classical environment. The content of the concepts themselves is fixed instead by their use in ordinary language and CM. Furthermore, Bohr takes the reality of the microdomain seriously. But our knowledge about the microdomain is limited: we cannot comprehend it beyond the usability of classical concepts, and thus cannot comprehend the microdomain by itself (independent of measuring arrangements)[33].

It is sometimes said that Bohr felt that irreversibility was an essential ingredient in a quantum measurement. While this is true, this aspect of measurement was, for Bohr, nothing new or specific to quantum mechanics, and therefore not related to complementarity[34].

ad (iii) As we saw earlier, Bohr in 1927 still perceived wave-particle duality (or complementarity) as important. In fact his views in QM took wave particle duality as a *starting point* (as opposed to discontinuity, from

---

[31]Cf. also a letter from Bohr to Pauli from 1947 [Bohr, *op. cit.* (1985) p. 451], and p. 315 of N. Bohr (1948): *Dialectica* 2, p. 312.

[32]Bohr, *op. cit.* (1929c) p. 18

[33]Hooker, *op. cit.*, p. 134 (B2 & B4) and p. 155.

[34]Bohr, *op. cit.* (1958) p. 3

which Heisenberg worked[35]). The later (i.e. p,q) complementarity is inherent in the earlier as Bohr saw it then, because he associated p with 'wave' and q with 'particle'. [Such an association is highly misleading, however: classical waves on the one hand can be localized instead of monochromatic, but usually have *neither* a sharp k ($= p/\hbar$) *nor* a sharp q, whereas on the other hand classical particles have *both*.] In early 1929 Bohr still sticks to the views expressed in the Nature paper, judging from his references to it in his Planck paper[36]. Later in 1929, the change approximately coinciding with the choice for 'complementarity' instead of 'reciprocity', Bohr has become more careful[37]: "It is true that light quanta and matter waves are invaluable expedients in the formulation of the statistical laws governing such phenomena as the photo-electric effect and the interference of electron rays. However, these phenomena belong, indeed, to a domain in which it is essential to take into account the quantum of action and where an unambiguous description is impossible. The *symbolical* character, in this sense, of the *artifices* [twice my italics] mentioned also becomes apparent in that an exhaustive description of the electromagnetic wave field leaves no room for light quanta and in that, in using the conception of matter waves, there is never any question of a complete description similar to that of the classical theories". In the Faraday lecture[38] and the Warsaw lecture[39] analogous reasoning occurs. In the Einstein essay there is no clear statement on this subject, although he does not present the double slit (WZ p. 24–27) as an example of wave-particle duality, even though it would have been a natural place to do so. [see also ad (iv)].

---

[35]Cf. Bohr, *op. cit.* (1985) p. 27ff. and Jammer, *op. cit.*. See also W. Heisenberg (1967): in S. Rozental (ed.) (1967): *Niels Bohr: his life and works as seen by his friends and colleagues* (North Holland, Amsterdam), p. 94; see esp. p. 106.

[36]N. Bohr (1929a) **[Planck Paper]**: *Naturwiss.* **17**, p. 483, see esp. p. 484

[37]Bohr, *op. cit.* (1929c) p. 17; cf also Bohr, *op. cit.* (1929b) p. 111 and p. 113

[38]Bohr, *op. cit.* (1930) p. 370 and p. 374; cf also Bohr, *op. cit.* (1931a)

[39]Bohr, *op. cit.* (1939) p. 16 and p. 237

There are, however, places where Bohr lapses: for instance in the Warsaw lecture[40] and on p. 25 (WZ) of the Einstein essay (cf. § 2). I do, however, not think that too much weight should be attributed to these slips of the pen in view of Bohr's clarity when he is discussing the matter directly[41].

ad (iv) Bohr takes a fundamentally instrumentalistic attitude towards the QM formalism. Despite this attitude Bohr is not an instrumentalist. He is a realist, and in fact attributes properties to the objects [viz ad (i)]. Bohr, in view of his opinion that all understanding must take place in classical terms, has no choice but to attribute no more than instrumentalist significance to the QM formalism. Although the formalism is necessary for quantitative calculations, it does not lead to understanding. In the EPR-answer he says[42]: "[...] there can be no question of an unambiguous use of the symbols of QM other than that embodied in the well-known rules which allow to predict the results [...]". In other places he says[43]: "The symbolic garb of the methods in question [i.e. QM] closely corresponds to the fundamentally unvisualizable character of the problems concerned" and[44] "The true significance of the wave formalism [of Schrödinger] as a most practical means of expressing the statistical laws [...]". Bohr's instrumentalism is in close connection with his view on the symbolical nature of material waves and photons [see ad (iii)], since the latter are used by Bohr as tools in qualitative analogs of exact QM calculations. He discusses, for instance, the interference pattern as it is built up out of many spots, the distribution of which "[...] follows a simple law derivable from the wave analysis" (WZ p. 24). In fact[45], "[...]

---

[40]Bohr, *op. cit.* (1939) bottom of p. 15

[41]Murdoch, *op. cit.*, p. 79

[42]Bohr, *op. cit.* (1935) p.701

[43]Bohr, *op. cit.* (1929c) p.12

[44]Bohr, *op. cit.* (1939) p. 16; cf also Bohr, *op. cit.* (1928) p. 586, Bohr, *op. cit.* (1930) p. 370, Bohr, *op. cit.* (1931b) p. 370 and Bohr, *op. cit.* (1958) p. 5

[45]Bohr, *op. cit.* (1939) p. 237; cf also Bohr, *op. cit.* (1949) p. 37 and Bohr, *op. cit.* (1929c)p. 17

> the duality between the undulatory and corpuscular conceptions [...] is only one aspect of a symbolical formalism [...]".
>
> Von Weizsäcker has distinguished between "parallel" (e.g. p versus q) and "circular" complementarity (causal description versus description in terms of the Schrödinger function)[46]. The latter form, like "wave particle complementarity", involves concepts that are not combined in any classical model, and therefore (ch. II) it is not surprising that Bohr categorically rejected circular complementarity.

In ch. II we noted that Bohr applied the name 'measurement' also to procedures we prefer to call 'preparation'. Such nomenclature is natural from a classical point of view (ch. I). Still, Bohr certainly was aware of the difference, and explicitly stated that *both* "types of measurement" are needed for a completely specified experimental arrangement[47]: "[...] all unambiguous use of the QM formalism involves the fixation of the external conditions, defining the initial state of the atomic system concerned and the character of the possible predictions as regards subsequent observable properties of that system. Any measurement in quantum theory can in fact only refer either to a fixation of the initial state or to the test of such predictions, and it is first [*sic*; B. means 'only' ?] the combination of measurements of both kinds which constitutes a well-defined *phenomenon* [my italics]".

Here (in 1939) we find the first use by Bohr of the word 'phenomenon' in a technical sense, referring to "[...] observations obtained under specified circumstances, including an account of the whole experimental arrangement" (WZ p. 46). In earlier essays, Bohr uses the word in a more informal way. The quotation also shows us how Bohr has to talk about preparation. Because a preparation by itself is not a phenomenon, it has to be discussed involving counterfactual measurements ("possibilities of prediction"). In other words, if we have a preparative measurement, which an analysis has shown to be describable in terms of the classical concepts to some extent, this immediately gives the extent to which the results of future determinative measurements can be predicted, if performed. Similarly the determinative context can

---

[46] Jammer, *op. cit.*, p. 103

[47] Bohr, *op. cit.* (1939) p. 20

be used for retrodictive deductions. If determinative and preparative context (properly transformed according to their time evolution) are incompatible, this means that predictions about measurement results based on knowledge of the preparation process become impossible (i.e. the results show a large amount of scatter). More accurately, the predictions loose their validity at least as much as they suggest precision beyond the uncertainty principle[48]. So, for example, we have the case where a particle with known momentum [preparative p knowledge] impinges on a photographic plate[49]: "Far from meeting any contradiction with the uncertainty relations, we have clearly here to do with a measurement arrangement [i.e. the p–preparator + photographic plate] which is not suited to define a phenomenon involving a test of predictions as regards the location of the object". Of course there remain two incompatible views on the situation in between the two measurements, but that is purely academic because it cannot lead to a conflict as regards the prediction of any phenomenon[50]: "We are dealing with an abstraction from which no unambiguous information concerning the previous or future behavior of the object can be obtained". Extrapolation from either of the two measurements cannot be pursued past the following *c.q.* preceding measurement because[51] "any observation takes place at the cost of the connection between the past and the future course of phenomena". If, for example, we prepare electrons with sharp momentum, pass them through a diaphragm and then measure their momentum, the slit's presence causes predictions as regards the measurement outcome based on knowledge of the preparation to be inexact. The disruption of the evolution is caused by the incomplete analyzability of the interaction of the object with the measuring arrangement, and is another way of talking about the disturbance in measurements. Accordingly, despite Bohr's emphasis

---

[48]Although Bohr in his examples usually discusses only extremes, in more abstract formulations he does take the full uncertainty principle, including intermediate cases, into account [e.g. Bohr, *op. cit.* (1939) p. 18].

[49]Bohr, *op. cit.* (1939) p. 23

[50]Bohr, *op. cit.* (1928) p. 583. Bohr here discusses two successive position measurements, which also provide incompatible contexts because $[X(t),X(t')]_- \neq 0$ (in general) if $t \neq t'$. The fact that he regards extrapolating away from the direct context not as *a priori* senseless (the problems occur not because of the extrapolation itself, but because of incompatibility with the following context), shows that already in 1928 Bohr was not really an interactionalist.

[51]Bohr, *op. cit.* (1929c) p. 11; see also p. 18

on classical description, terms like "uncontrollable disturbance" do not refer to the change of value of certain *autonomously possessed* magnitudes[52].

Summarizingly, we see that a complete phenomenon requires for Bohr two measurements, requires both preparation and determination. Thus a particle emerging from a preparator of a classical "property" cannot be unambiguously said to have that property: preparation does not by itself constitute a phenomenon. The case of incompatible determinative and preparative contexts illustrates this ambiguity. Experimental outcomes in such a situation are only probabilistically predictable. If the contexts are compatible, analysis in terms of some classical quantity is possible, and predictions are (like in CM) determinate. This constitutes an important qualification of the remarks as regards objectivity of properties made in ad (i).

In the following we shall further illustrate Bohr's views in a number of imaginary situations. The diaphragm and the related double slit experiments have been discussed in ch. II, so we will not discuss them in detail here. In Bohr's discussion of them (WZ p. 20–30), points (i) and (ii) are clearly present. We must, however, keep in mind that Bohr's wording in the Einstein essay is in several places intended to reflect the opinions he or Einstein held at the time of the episode described.

2 PHOTON IN THE BOX (WZ p. 28–29)

At the 1930 Solvay conference Einstein discussed a *Gedanken* experiment intended (according to Bohr)[53] "[...] to show that the foregoing considerations [i.e. complementarity] are valid as long as the viewpoint of general relativity is left out of the discussion". Einstein conceived of a box, in which a photon and a clock are present. The clock controls the shutter, so that the time of emission of the photon can be

---

[52]Hooker, *op. cit.*, p. 155

[53]Bohr, *op. cit.* (1931b) p. 368. It is, however, quite plausible that Einstein intended it for an altogether different purpose, namely as a predecessor of EPR [D. Howard (1990): in *Sixty-two Years of Uncertainty* (ed. by A. Miller, NATO ASI Series B 226, Plenum, NY), p. 61].

determined. He then proposed also to measure its energy by weighing ($E = mc^2$) the box before and after the emission. The precision of the energy measurement need not be at least reciprocally related to the opening interval of the shutter. It seems that the experiment is completely analogous to the diaphragm, where a momentum measurement before and after scattering may be carried out (ch. II). Thus Bohr could have sufficed with an analogous answer: if we perform the first energy measurement (weighing) with accuracy δ_E, events within the box after the weighing cannot be coordinated in time with those before, better than $1/\delta_E$. Similar arguments apply to the second weighing. In other words: the opening and closing time of the shutter may be well-defined *within the box*, but they are not so *to an observer outside it*. This seems to be the gist of the remark: "If we are interested in such conclusions [i.e. concerning *E*] we must, of course, use an arrangement where the shutter devices can no longer serve as accurate clocks [cf. ad (ii)], but where the knowledge of the moment when the hole of the diaphragm is open involves a latitude connected with the accuracy of the energy measurement by the [uncertainty relation]" (WZ p. 29). Possibly in view of his understanding of Einstein's goal (see above), or perhaps to heighten the dramatic effect, Bohr used a more explicit argument involving Einstein's own theory of general relativity. Somewhat simplified[i], it amounts to the following:

> In the box a timer is contained in addition to the photon. This timer is designed to open a hole in the side of the box for a pre–set period of time. The box is allowed to fall freely in a gravitational field of strength g. On t_1 and t_2 (both before the box opens) the box's momentum is measured (e.g. through collision with an object with known mass). Then the initial mass m_i follows from the results of these two measurements:

(1) $$m_i = \frac{p_{t_2} - p_{t_1}}{g(t_2 - t_1)} \quad .$$

> The final mass m_f is determined analogously. Therefore the photon energy measurement inaccuracy δ_E satisfies:

(2) $$\delta_E = c^2 \delta_{m_i - m_f} \geq c^2 \delta_{m_i} = \frac{c^2}{g(t_2 - t_1)} \delta_{p_{t_2} - p_{t_1}} \gtrsim \frac{c^2}{g(t_2 - t_1)} \delta_{p_{t_1}} \quad .$$

During the *m*–measurement procedure (between t_2 and t_1) there is a latitude δ_q in the definition of the box's height. This leads, via the gravitational redshift

(3) $$\frac{\Delta T}{T} = \frac{g\,h}{c^2}$$

to an uncertainty δ_t in the delay of the timer due to the gravitational field. Thus $[T \rightarrow (t_2 - t_1)\,;\; h \rightarrow \delta_q\,;\; \Delta T \rightarrow \delta_t]$:

(4) $$\delta_t\,\delta_E = \frac{g\,\delta_q}{c^2}(t_2 - t_1)\,\delta_E \gtrsim \frac{g\,\delta_q}{c^2}(t_2 - t_1)\,\frac{c^2}{g(t_2 - t_1)}\,\delta_p = \delta_{p_{t_1}}\,\delta_q \geq 1 \quad .$$

This uncertainty in timer delay will lead to a corresponding uncertainty about the photon's emission time. Analogously to the above, the second weighing will make it impossible to improve the definition of time by opening the box *afterwards* and comparing the reading of the timer to that of a standard clock. Both weighings imply a rupture in the box's development as regards the time picture. On the other hand opening the box to compare the timer to a standard clock *before* the second weighing will render the conservation laws inappropriate and thus disturb the determination of the mass difference.

Both elements (i) and (ii) are present. The discussion, however, contradicts (iii). The photon concept may not be used in the above way because it cannot aid in conceptual understanding and because no uncertainty relations of the type $\delta_t\,\delta_E \geq 1$ or $\delta_p\,\delta_q \geq 1$ exist for photons. Einstein, however, proposed the experiment *with* photon, so I think this inconsistency should not be taken too seriously. Moreover, the argument remains just as valid in other respects when 'electron' is substituted for 'photon'.

3 SPIN MEASUREMENT

Another interesting experiment concerns the possibility of determining the magnetic moment (spin) of an electron with a Stern-Gerlach device. Bohr's doubts as regards the feasibility of such an experiment seem at first sight in contradiction to the astronomical precision with which the magnetic moment of electrons is known (viz. the anomalous g_s–factor[54]). The situation of the experiment is depicted in fig. 1. Electrons enter the setup in a beam of width Δq_2 moving in 1–direction with momentum p_1. The magnetic field is inhomogeneous and has in the area of the beam ($q_2 \simeq 0$, $q_3 \simeq 0$) the value $\vec{B} \simeq (0, bq_2, a - bq_3)$, with in the beam area $a \gg bq_3$. Note that $(\vec{\nabla} \cdot \vec{B}) = 0$ is satisfied. This is crucial in Bohr reasoning, as we shall see.

The working of the Stern-Gerlach device is based on the existence of a spin-dependent force in the 3–direction, of magnitude

$$F_{q_3,s} = g_s \mu_B \frac{\partial B_{q_3}}{\partial q_3} \simeq \frac{eb}{m} \quad (\hbar = 1). \tag{5}$$

We also have the Lorentz-force, however. Its 3–component is not constant over the 2–width of the beam. It varies by an amount

$$\Delta F_{q_3,L} = \frac{ep_1}{m} \frac{\partial B_{q_2}}{\partial q_2} \Delta q_2 = \frac{eb}{m} p_1 \Delta q_2 \ . \tag{6}$$

For a correct operation of the device we must have at least

$$F_{q_3,s} \geq \Delta F_{q_3,L} \Rightarrow p_1 \Delta q_2 \lesssim 1 \Rightarrow \Delta q_2 \lesssim \lambda/2\pi \tag{7}$$

This last condition is, according to Bohr, in contradiction to the particle conception in the sense that use of the notion of 'beam' on which the whole analysis is based, is excluded by it. As we narrow down the beams in order to attempt to comply with (7), diffraction effects in the 1,2–plane will increase so that the attempt cannot succeed. In

[54] F. Combley (1979): *Rep. Progr. Phys.* **42**, p. 1887

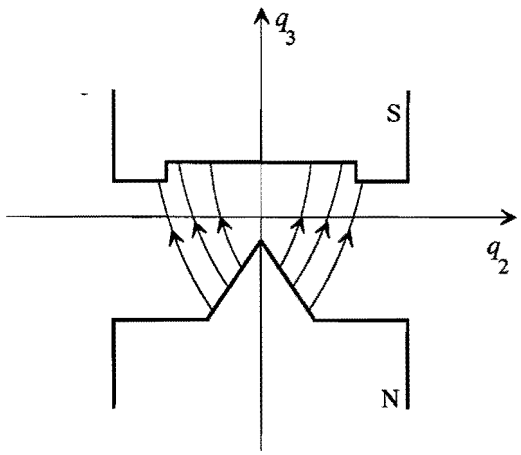


fig. 1 *Schematic cross-section of the Stern-Gerlach setup. The inhomogeneous magnetic field is indicated by the dashed lines. The electrons move in the* 1*–direction, perpendicular to the plane of the drawing*

Bohr's words[55,ii], "[...] it is impossible to reconcile the idea of a well-defined direction of the beam, with the condition that the variation of the Lorentz-force within the beam shall not exceed the deflecting force due to the electron magnetization". [Of course in the original experiment *neutral* (silver) atoms were used, so that the Lorentz-force does not come into consideration there.]

Bohr has additional arguments. In the first place angular momentum L and angle Θ are canonically conjugate, so that $\Delta L\, \Delta\Theta \geq 1$. Since moreover $\Delta\Theta \leq 2\pi$, $\Delta L \geq 1/2\pi$ implying a blurring of all effects due to spin[56]. [This argument is incorrect: due to the boundedness of Θ no such simple uncertainty principle for the L,Θ–pair[57] exists.

[55]N. Bohr, *op. cit.* (1929d): "The Magnetic Electron" p. 333 of Bohr, *op. cit.* (1985); quote from p. 335 [This manuscript's date is uncertain; in particular it is not clear whether it was written before or after Bohr, *op. cit.* (1929c).]

[56]Bohr, *op. cit.* (1985) p. 348 (from the discussions at the 1930 Solvay conference)

[57]See e.g. J. Uffink (1990): *Measures of Uncertainty and the Uncertainty Principle* (PhD Thesis, Utrecht University), unpublished. Bohr's argument is all the more strange because the problems with L-Θ had been pointed out to him by Schroedinger as early as May 1928. (Bohr, *op. cit.* (1985) p. 46).

Bohr's reasoning implies that angular momentum is not be accurately measurable at all !] In the second place, Bohr argues, the electron's magnetic moment is proportional to $\hbar$, so that[58] "any effect of the magnetic moment ascribed to the electron will disappear in the region of legitimate application of classical mechanics which involves the neglect of the quantum of action".

At first Bohr interpreted this reasoning as implying that the spin of a free electron could not be measured at all[59]. [When bound to an atom, such as in the original Stern-Gerlach experiment, the electron-spin is part of the atom's angular momentum, which does have a classical analog.] We hear a remainder of this in the remark[60] "[...] the possibility of a direct observation of the magnetic moment of the electron would be inconsistent with the fundamental principles of quantum mechanics". But as Mott showed that through spin-dependent scattering a method of preparing [~ measuring in Bohr's terminology] could be devised, Bohr withdrew this radical point of view. He writes to Mott[61]: "In fact the argument I used in the discussion of the Stern-Gerlach effect was not strict due to the mixture of classical mechanics and wave theory in the very region where no sharp distinction is possible. Thus the argument tells only that as regards the electron magnetization we cannot as in the ordinary discussion of the Stern-Gerlach effect base our considerations on classical pictures of a moving magnet, but not that the closer quantum mechanical treatment will never give a positive effect".

Summarizing, we can say that Bohr's point of view is that a Stern-Gerlach entails an object description in terms of space-time pictures. Spin does not belong to this cluster of concepts, and the Stern-Gerlach is therefore unsuitable for spin determination[62]. Comparison with anomalous g experiments is not as trivial as it may seem. Such experiments do not use a Stern-Gerlach. This would seem to be in accord with Bohr's reasoning. But in fact Bohr's experiment and anomalous g experiments have

---

[58]Bohr, *op. cit.* (1929d) p. 333; cf. Bohr, *op. cit.* (1929c) p. 13 and Bohr, *op. cit.* (1930) p. 391

[59]Cf. Bohr, *op. cit.* (1985) and O. Darrigol (1984): *Historical Studies in the Physical Sciences* **15**, no. 1, p. 38

[60]Bohr, *op. cit.* (1929d)

[61]Letter quoted on p. 56 of Darrigol, *op. cit.*

[62]Bohr, *op. cit.* (1929c) p. 13

completely different goals: the former attempts to find out the value of z spin (i.e. up or down) for *individual* electrons, whereas the latter are intended to determine the value of g_s, which is a constant of the *species* 'electron'. Bohr's reasoning should be interpreted as giving an absolute lower limit to the accuracy with which an electron-spin measurement using a Stern-Gerlach can be carried out[63]. As polarized beams of electrons can be prepared (as Bohr himself acknowledges), there is no reason why the accuracy of a Stern-Gerlach device could not be determined with arbitrary precision. The g factor could then be derived out of a comparison of the experimental value of this accuracy with a theoretical result. Consequently Bohr's reasoning is not only not at variance with g factor determination in general, but not even with g factor determination using a Stern-Gerlach.

Especially Bohr's early opinion that electron spin could not be measured because it does not have a classical analog, is evidence of (i). His discussion with Mott shows that he nevertheless had no objection to QM calculations to get *quantitative* answers [(iv)]. In ch. V a more quantitative discussion of a Stern-Gerlach is performed. We see there that, contrary to Bohr's opinions, it functions reasonably well for electrons.

4 EPR

The most famous and thought provoking thought experiment discussed by Bohr is undoubtedly the one that was proposed by Einstein, Podolsky and Rosen[64] (EPR) in 1935. In the following we shall discuss it with special regard for Bohr's opinions (§ 4.3), in as far as these are important to establishing (i)–(iv). We shall not consider the locality discussion it has engendered. First, however, we discuss the experiment itself in some detail, since the EPR paper probably was a rather awkward reflection of Einstein's views (the EPR paper was not written by Einstein, but by Podolsky[65]).

---

[63]Darrigol, *op. cit.*

[64]A. Einstein, B. Podolsky & N. Rosen (1935): *Phys. Rev.* **47**, p. 777. See for the history of this experiment e.g. Howard, *op. cit.*

[65]See e.g. M. Jammer (1985): in *Symposium on the Foundations of Modern Physics*, Joensuu, Finland (ed. by P. Lahti & P. Mittelstaedt, World Scientific, Singapore), p. 129.

4.1 EPR I

In formulations of QM such as we will discuss in ch. III, the state↔observable distinction as a reflection of the preparation↔measurement dichotomy (ch. I) is particularly clear. We saw in ch. I, however, that this distinction is not quite natural from the classical point of view current in the early stages of QM. The objectivism of classical measurement reduces it the distinction to a conceptual triviality. This also plays a role in the EPR debate. In fact different EPR-arguments can be distinguished according to how far we pursue the state↔observable distinction. First we shall discuss the original EPR version, in which the viewpoint is quite classical.

EPR want to find out if QM is *complete*. A theory T is defined to be complete when it reflects all of reality:

DEFINITION T is complete := Every element of reality corresponds to an element of T .

EPR clearly find completeness an important property for a theory. They go (implicitly; but cf. the last paragraph of the paper) so far as think that a theory that is not complete, can only be of temporary significance. Incompleteness is, in their view, evidence of the existence of a deeper, complete, theory. Thus an incomplete theory can never be a candidate for the "ultimate" theory ('finality'). Completeness is a necessary (rather than only a sufficient) condition for finality:

(8) T is final $\Rightarrow$ T is complete .

Here we already see that the finality itself plays no role in the EPR reasoning. It is the validity of the implication (8) rather than the validity of its premise which lends completeness its importance. Moreover, we have to distinguish between *finality* and *experimental adequacy*: EPR find QM not acceptable as the final theory *even if* its experimental domain of validity would be infinite. In the latter case EPR would feel that there are parts of reality, not directly accessible, that a complete theory should describe, but which QM does not. Experimental adequacy was, for EPR, not at stake. Similarly the domain of validity was not the subject of discussion for Bohr[iii] either,

although the distinction between finality and experimental adequacy made no sense for him[66].

To be able to show incompleteness in a clean way, one would really need an look at reality from a higher meta-level. If the theory itself is used to judge what is real and what is not, the reasoning leading to the decision about completeness would run the risk of being circular. One would need extra-theoretical arguments for the existence of certain elements of reality[67]. Such arguments are not available. Therefore EPR introduce a reality criterion that is only intuitively plausible:

CRITERION We can predict the value of quantity $\mathcal{A}$ with certainty (i.e. with probability 1) without disturbing the system

$$\Downarrow$$

There is an element of physical reality corresponding to $\mathcal{A}$.

EPR then consider the case where we have two particles A and B, spatially separated. The joint system is in a state where $P_a + P_b$ and $Q_a - Q_b$ are both sharp and both known (preparative knowledge; particle operators are denoted by corresponding indices). If we measure P_a, we can deduce the value of P_b. Similarly a measurement of Q_a leads to a prediction about the value of Q_b. Since in both cases there can be no question of a physical influence on the second system (spatial separation), we can employ the criterion: both P_b and Q_b correspond to elements of reality. Since they are incompatible, QM cannot describe this. Therefore QM must be incomplete.

The reality criterion is about the reality of *observables*. It can be seen as an adaptation of the classical measurement concept, designed to avoid direct interaction between object and meter. In that way the criterion evades reasoning (Heisenberg) based on the physically disturbing influence of measurements in QM[68]. The classical measurement concept contains, however, a counterfactual element: if a certain value for a quantity turns up in a measurement, we assume *that the quantity would have also had that value if we had not made the measurement.* This counterfactual element is a

---

[66]Bohr, *op. cit.* (1949) p. 37

[67]E. McKinnon (1982): *Scientific Explanation and Atomic Physics* (University of Chicago Press), p. 339

[68]W. de Muynck (1986): *Found. Phys.* **16**, p. 973

consequence of the *non-contextual* objectivism of CM: measurement outcomes are regarded as reflections of properties *independently* possessed by the object system. Counterfactuality is not explicitly mentioned in above criterion, but EPR obviously do talk about such classical reality. Accordingly, it may be argued that if we can predict the value a quantity has *without disturbance*, we may safely assume, that the object had the same value as a property before the measurement that allowed us to make the prediction, took place. In fact, we may (according to EPR) attribute the value as a property to the object *even if that measurement does not take place at all*. EPR do use counterfactuality, albeit implicitly, and only thus can they jointly attribute P_b and Q_b values to system B.

Consequently EPR seem to take for granted that measurements bring to light pre-existing quantities: the EPR reasoning uses classical, non-contextual, objectivism as regards measurement (ch. I), with all its disadvantages. One may argue that the 'no disturbance' proviso is useless: even if we assume that the value of the quantity is changed in the measurement, we get value attributions not reproducible in QM from the classical measurement objectivism *by itself*. If we, on the other hand, assume that a measurement *creates* a property[69], the EPR measurement scheme creates the P_b and Q_b values to system B in different situations. They do not exist jointly. Whereas a change in a pre-existing quantity may obviously be regarded as a disturbance, it is not clear whether the creation of an element of reality where there was none before, can be regarded as such. Thus the EPR reasoning leaves a lot to be desired. In the following sub-section we present an argument that is purer, and probably closer to what Einstein actually had in mind.

4.2 EPR II

We consider, like in EPR I, a two particle system in an entangled state:

(9) $$|\psi\rangle_{a+b} \sim |\phi\rangle_a \otimes |\chi\rangle_b + |\phi'\rangle_a \otimes |\chi'\rangle_b \ .$$

[69]Such a view on measurement was advocated by e.g. Jordan [Jammer, *op. cit.* (1974) p. 161], although Bohr did not approve of it [Bohr, *op. cit.* (1939, 1949)].

We also assume, like above, that the two systems are spatially separated and do not interact. We then make some measurement on system A. The state giving the QM description of system B afterwards will depend on the measurement outcome and on the kind of measurement performed[iv]. We now assume:

ASSUMPTION If a system is not interacting with other systems, the system independently possesses a property ($\Rightarrow$ element of reality) called the 'state' of the system.

Even without making completely specific what we mean by 'state', it is clear that wave vector and density operator are the only QM candidates for the job. Since a A–measurement cannot physically disturb B, we must assume that the B–state resulting from the A–measurement really is that system's state. But the same reasoning can be applied to all the different B–states resulting from different *types* of A–measurement. (Note that the dependence on the *outcome* of the A–measurement corresponds to a classically understandable correlation, and is unproblematic.) All may be considered as representing the second system's real state. This is absurd. Thus neither QM state description is satisfactory, and we must conclude that there is no element of QM corresponding to the element of reality 'state'. QM is incomplete.

This EPR formulation does not make any idealizing assumptions about measurement. We have in particular not assumed that the measurement is anything like an ideal classical measurement, or like a first kind measurement. Thus many of the objections against EPR I are invalid here. This allows us to focus more clearly on the real issue, as embodied in the assumption, without the quasi-logical clouding of the EPR-paper. In view of the formulations Einstein himself chooses[70] to describe the EPR situation, there can be little doubt that this reasoning is much closer to his views than that in the EPR paper. Pauli perceived this at once. He writes to Heisenberg[71]: "Aber woran

---

[70]See e.g. his reply to the essays in the Schilpp volume [A. Einstein (1949): in Schilpp, *op. cit.*, p. 665], a letter to Popper [K. Popper, *op. cit.*(1972), app. *xii], a letter to Schroedinger [quoted on p. 173 of Murdoch, *op. cit.*] and a 1948 paper [A. Einstein (1948): *Dialectica* 2, p. 320]. Cf. also Howard (1989), Murdoch, *op. cit.*, p. 166 and Hooker, *op. cit.*, p. 88

[71]"But what [Einstein] is annoyed about in this conection is the way in which in QM two systems constitute one joint system". W. Pauli (1979): ***Wissenschaftlicher Briefwechsel*** (ed. by A. Herman, K. von Meyenn & V. Weisskopf, Springer, Berlin), vol. II, letter no. [412]

[Einstein] sich in Verbingung damit stößt, ist die Weise, wie in der Quantenmechanik zwei Systeme zu einem Gesamtsystem zusammengesetzt werden". It is unfortunate that Bohr's reaction, discussed in the next section, was focused on the EPR reasoning, rather than on Einstein's own.

4.3 Bohr's answer

Bohr answered the EPR challenge almost immediately[72]. From his answer it is clear that he regards the EPR imaginary experiment as just another simple example, similar to the ones he has treated earlier (diaphragm, double slit). Thus he starts the paper by giving a general overview of complementarity, leaving the mathematics of EPR to a footnote. Then he goes over to the EPR case, "[...] which does not actually involve any greater intricacies than the simple examples discussed above" (EA p. 699; quotations denoted by EA are from Bohr, *op. cit.* (1935)) [v]. In his discussion Bohr then shows that EPR's reality criterion is ambiguous, and that their conclusion is therefore untenable. Before we discuss Bohr argument in more detail, we can on the basis of the earlier section already see that that must be Bohr's way out: as we noted, EPR talk about *independently possessed* real attributes (in EPR I these correspond to quantities, in EPR II to states). For Bohr, however, we must talk about reality in classical terms [(i)], and which terms are applicable is *determined by the experimental surroundings*. The indivisibility of the quantum of action implies the impossibility of all concepts' being applicable (beyond the uncertainty principle). Thus we cannot talk about reality independent of the means by which we investigate it. EPR's reality criterion cannot be but ambiguous.

Consider a slightly more general example than that introduced by EPR (the reason for the generalization will become clear later), where W_a and X_b are "correlated", and so are Y_a and Z_b. The operators W_a and Y_a are taken such that they do not commute, and similarly X_b and Z_b do not commute. In the EPR case, of course, $W_a = Q_a$, $X_b = Q_b$, $Y_a = P_a$ and $Z_b = P_b$. How does this "correlation" come about ? In QM

[72]Bohr, *op. cit.* (1935)

it is inherent in the state description, and from this we conclude that it must come from the state preparation. Thus Bohr, because he cannot use the QM state to understand the experiment [(iv)], must give an account of the preparation. He has to give the "measurement" that brought the "correlation" into being[73]. Indeed Bohr spends the latter half of p. 699 (EA) on a discussion of a device consisting of a double slit diaphragm and a momentum meter that can (approximately) prepare two particles in the state used by EPR. Now we can, e.g., perform an W_a–measurement on system A. Then obviously the W_a–picture is incompatible with the picture in which the correlations are defined, in particular with the Y_a,Z_b–side of it. As discussed earlier, a measurement must make predictions based on earlier measurements (in the Bohrian sense) worthless (at least) in as far as they would otherwise conflict with the uncertainty principle. Thus the W_a–measurement will impair predictions based on the Y_a,Z_b–"correlation", whereas the W_a,X_b–"correlation" remains usable. Analogously we can discuss the case of an Y_a–measurement. Thus different measurements give rise to different "types of prediction" (EA p. 700) as regards system B. The phrase "types of prediction" refers to the types of B–arrangement in which the knowledge of A+B preparation combined with A–measurement can be put to use. An W_a–measurement will enable us to make predictions about X_b , if we perform that measurement. If we perform a measurement other than of X_b (the choice is ours), then the predictions become useless because of the incompatibility of the observable measured and the one we can predict. In fact a phrase similar to "types of prediction" also occurs in the Warsaw lecture[74], where Bohr is describing state preparation (see above). Thus the EPR-setup (A+B–preparation combined with A–measurement) is for Bohr just another state preparation procedure. In this sense Bohr's reasoning closely reflects the EPR-description in the QM formalism.

The extra complication, from EPR's point of view, is of course that we can prepare system B in different states (viz. EPR II) without tampering with it. EPR argue that if the QM state description is complete, this can only be understood via a (non-local) physical interaction. This reasoning is, however, only compelling when one thinks of the 'state' as an independent object-property. For Bohr there is nothing in system B

[73]de Muynck, *op. cit.* (1986)

[74]Bohr, *op. cit.* (1939) p. 20

that is mechanically influenced by the type of measurement we choose to perform. Instead the type of A–measurement we choose to perform, determines (*via* the A+B–preparation) the type of B–preparation arrangement (EA p. 700; the preparator corresponds to the QM B–state description). As these arrangements must be analyzed in different, incompatible, pictures, they are mutually exclusive. This illustrates the more general rule that *there can be no talk of a 'state' independent of the arrangement* (i.e. the preparator) *that brought it about*[75]. Our earlier remarks regarding the limited ontical significance of a preparation by itself (a preparation alone is not a phenomenon), also apply here. Accordingly, we here cannot talk about properties of system B when no measurement is performed on it[76]. Only if we perform the B–measurement the setup allows us to predict, we can consistently ascribe a property to system B.

We have chosen a somewhat more general description of Bohr's answer, because the original one is, quite understandably, bound to the particular form EPR gave their objection. Thus Bohr's answer in above form can be more closely related to EPR II. It has the further advantage that certain misunderstandings as regards Bohr's answer can be cleared up. Usually the fact that Bohr (unlike EPR) explicitly describes an arrangement that prepares the EPR correlation, is overlooked[77]. But then we lack the connection needed to base B–predictions on A–measurement. In the original EPR example momentum is correlated with momentum, and position with position. That suggests the possibility of fixing the problem by assuming that a Q_a–measurement makes momentum inapplicable to the whole setup, i.e. *to both system A and system B.* Then the possible type of prediction for system B is always the same as that which is measured on system A, regardless of the type of initial "correlation" (A+B preparation). On the basis of this reasoning a viable answer to EPR can be given, but *only when their formulation is strictly adhered to.* If EPR had chosen to take an initial state such that Q_a is correlated with P_b and P_a with Q_b, this form of Bohr's answer would have no longer made sense. That it is incorrect can also be seen by a consideration of the further arguments to support it given by Hooker[78]. He discusses a more concrete

---

[75]Bohr, *op. cit.* (1939) p. 21

[76]de Muynck, *op. cit.* (1986)

[77]E.g. by Hooker, *op. cit.*, McKinnon, *op. cit.* (1982) and by de Muynck, *op. cit.* (1986)

[78]Hooker, *op. cit.*, p. 222ff

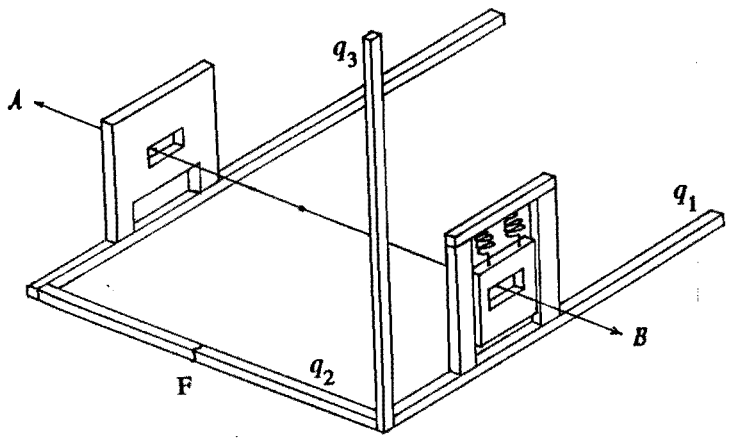


fig. 2 *Hooker's gedanken experiment. The frame* F, *to which a movable diaphragm (~ p) on the B-side and a fixed one (~ q) on the A-side are attached, explicitly defines the coordinate system (figure taken from Hooker,* op. cit.*).*

device (fig. 2). A Q_a–device and a P_b–device are both firmly bolted to a frame F. He then argues that in the Q_a–measurement an uncontrollable amount of momentum is passed into F, obscuring the relation (in p picture) between F and the momentum meter on the B–side. Now the P_b–measurement can be saved, according to Hooker, by sawing F in half, but then we lose the connection between the two measurements. The flaw in the argument is that for the quality of the P_b–measurement not the indeterminateness of the frame's momentum is important, but the indeterminateness of its velocity. By increasing F's mass we can reduce this unwanted effect indefinitely. Moreover, Hooker's reasoning is in fact in conflict with Bohr's views. For Hooker the indeterminate momentum exchange leads to an indeterminate frame momentum. But that argument is only correct when momentum conservation is applied to system A+F (including diaphragm). But then application of space-time pictures to the A-side is out of the question: the diaphragm no longer measures position. The amount of momentum flowing into the diaphragm's support is in fact untraceable *because the support is unmovable*. Bohr always effectively takes the frame's mass to be infinite, so that momentum is not conserved. Thus Hooker's version of Bohr's answer is in

conflict with Bohr's views, totally ignores (unlike Bohr) the "correlation" in the initial state, and can be countered by a simple change in the EPR argument.

In Bohr's answer especially (i) is clear. In his intuitive explanation Bohr has to choose a very awkward wording ("an influence on the very conditions which define the possible types of predictions regarding the future behavior of the system", EA p. 700), because of the necessity of describing the situation in classical terms and because a preparation by itself does not constitute a phenomenon (see above discussion about Bohr & preparation). Bohr cannot have recourse to the QM state description [(iv)], as it does not lead to conceptual comprehension (EA p. 701). In conclusion, Bohr finds that QM is not incomplete as it takes all available information in a consistent way into account[79] (EA p. 700). 'Completeness' in the Einsteinian sense is, for Bohr, simply not a well-defined notion.

NOTES

i Bohr's original version of the experiment is more complicated, and involves a balance. His description of the experiment is rather too condensed to be easily accessible. So, e.g., it seems to imply that one would have to read the balance inaccurately in order to obtain an accurate weighing. The version given here is simpler and just as suitable.

ii Bohr's reasoning is rather strange (cf. Casimir's remark on p. 111 of Rozental, *op. cit.*). The fact that the magnetic field is divergence-free is crucial. Thus his argument would falter if magnetic monopoles would turn out to exist. Similarly, Bohr's use of general relativity to defend a part of QM in § 2 is questionable from a methodological point of view (cf. Jammer, *op. cit.* (1974), p. 136ff).

iii Bohr always kept a (more than) open mind to the possibility that QM would end (cf. ch. II and, for instance, the closing paragraphs of the Faraday lecture [Bohr, *op. cit.* (1930)]), and he suggested fundamental revisions on a number of occasions (none of them turned out to be necessary; cf. Bohr, *op. cit.* (1985) and E. McKinnon (1985): *Niels Bohr, a Centenary Volume* (ed. by A. French & P. Kennedy; Harvard University Press), p. 101). Bohr, among the founders of QM, was probably the one with least confidence in it.

[79]McKinnon, *op. cit.* (1982) p. 345

iv Note that the state description of CM (probability distribution over phase space) also exhibits such dependencies. There, however, they are resolved by the realization that this state description is incomplete. The "real" state (point in phase space) is never correlated.
It is sometimes claimed (e.g. by H. Margenau (1936): *Phys. Rev.* **46**, p. 240) that the postulate (I.2.e) of the measurement of the first kind is the cause of the EPR trouble. But this is not the case. Postulate (I.2.e) is not used at all. It is a statement about what happens to system 1 after it has been subjected to measurement. The fate of system 1, however, is of no concern to EPR. The measurements involved may be destructive (viz. the Aspect experiment), a situation certainly not described by (I.2.e). In fact, the "empirical" argument given by von Neumann in favor of (I.2.e) [ch. I], is based on a similar misunderstanding.
In the EPR-II argument we may as well use the more general formalism to be discussed in ch. III, without affecting essentials.

v Bohr seems to imply that EPR's argument is, from his point of view, a needless complication. If the EPR challenge would really have been so important as to cause a major change in Bohr's views, his style of presentation in the EPR answer (and this remark in particular) would reduce to pure demagogy.

Appendix C
Ad chapter III: proofs

For the proofs in this appendix we will assume that all elements of an EVM are non-zero. It can straightforwardly be checked that this is possible without loss of generality. We shall further assume all outcome sets to be subsets of some fixed countably infinite outcome set **F**.

1 AD § 2.2

LEMMA 1 Any two equivalent and pairwise linearly independent EVMs are equal up to their labeling.

PROOF OF LEMMA 1

Suppose two EVM's $n = \{N_\ell\}_L$ and $o = \{O_m\}_M$ are given such that $n \longleftrightarrow o$ and such that both EVM's are pairwise linearly independent. There exist non-ideality matrices $\{\lambda_{\ell m}\}$ and $\{\mu_{m\ell}\}$ such that:

$$N_\ell = \sum_{m\in M} \lambda_{\ell m} O_m \quad (\ell \in L); \quad O_m = \sum_{\ell\in L} \mu_{m\ell} N_\ell \quad (m \in M).$$

Define:

(1) $$\gamma_{km} := \tfrac{1}{2}(\delta_{km} + \sum_{\ell\in L} \mu_{k\ell}\lambda_{\ell m}) \quad (k,m \in M).$$

Then:

(2) $$O_k = \sum_{m\in M} \gamma_{km} O_m \quad (k,m \in M).$$

Using the notation $p_m = \mathrm{Tr}(\rho\, O_m)$ (ρ arbitrary), this amounts to the eigenvalue equation $p_k = \sum_{m\in M} \gamma_{km} p_m$ (eigenvalue 1).

A square matrix like $\{\gamma_{km}\}$ can be used to represent a Markov chain with stationary transition probabilities[1] (app. A). Seen from that point of view, eq. (2) means that this chain has a summable stationary distribution (one for every ρ).

If a non-empty set $J \subseteq M$ has the property that $\forall_{m\in J} \sum_{k\in J} \gamma_{km} = 1$, we call it *closed*. If the set J has no proper closed subsets, it is *minimal closed* or, equivalently, *essential* (see app. A). It can be shown that, given that (2) is satisfied and that $\forall_{m\in M} O_m \neq 0$, a non-empty proper subset of M is closed if its complement in M is either closed or empty. We introduce the index set I,

(3) $$I := \{m \in M |\ \gamma_{mm} < 1\} \quad ,$$

which we shall assume to be non-empty. Then $\bar{I}$, the complement of I in M, is either empty or closed in M , which means I is closed, too. Look at the smallest closed set $J_i \subseteq I$ containing $\{i\}$ for some $i \in I$. Since I is closed and not empty there is such a subset. On J_i the eigenvalue equation reduces to:

(4) $$p_k = \sum_{m\in J_i} \gamma_{km} p_m \quad (k \in J_i).$$

The matrix $\{\gamma_{km}\}_{J_i \times J_i}$ has, because of (1), a *period* equal to 1. [The period of a stochastic matrix Γ is the smallest number $m > 0$ such that the diagonal elements of Γ^m are non-zero[2].] Since the set J_i is minimal closed by construction, the stationary summable distribution is unique up to a scalar factor[3]. This means that (4) has a solution:

$$p_m = \lambda a_m \quad (m \in M;\ \lambda \geq 0);$$

for some fixed non-negative sequence (a_m). This would imply for (2) that $O_m \sim O_k$ if $k,m \in J_i$. Since o was pairwise linearly independent by

---

[1]K. Chung (1967): *Markov Chains with Stationary Transition Probabilities* (2nd ed.; Springer, Berlin)

[2]Chung, *op. cit.*; App. A

[3]For finite J_i this is a consequence of the Perron-Frobenius theorem [J. Ortega (1987): *Matrix Theory* (Plenum, NY); see app. A].

assumption, this means that J_i cannot contain other elements than i . But then the eigenvalue 1 can only be achieved if $\gamma_{ii} = 1$, contradicting the definition (3) of I .

Since p was arbitrary, I does not contain any minimal closed subset and must be empty. Hence $\gamma_{km} = \sum_{\ell \in L} \mu_{k\ell}\lambda_{\ell m} = \delta_{km}$. Similarly one can prove that $\sum_{m\in M} \lambda_{km}\mu_{m\ell} = \delta_{k\ell}$. It is not difficult to see that this implies that $\{\lambda_{\ell m}\}$ and $\{\mu_{m\ell}\}$ must be permutation matrices.

□

PROOF OF THEOREM 4

An arbitrary EVM $\mathfrak{m} = \{\boldsymbol{M}_k\}_K$ is given. Divide K into subsets:

(5) $$\mathscr{K} := \left\{ I \subseteq K \;\middle|\; \forall_{k\in I}\left[\forall_{m\in K}\ \boldsymbol{M}_k \sim \boldsymbol{M}_m \iff m \in I\right]\right\} .$$

Obviously $\mathscr{K}$ can be mapped into $\mathbf{F}$. We have:

$$\forall_{k\in K}\ \exists!_{J\in\mathscr{K}}\ k \in J .$$

This allows us to define the EVM $\mathfrak{m}' = \{\boldsymbol{M}_I'\}_{\mathscr{K}}$:

(6) $$\boldsymbol{M}_I' := \sum_{k\in I} \boldsymbol{M}_k \quad (I \in \mathscr{K}).$$

This EVM is pairwise linearly independent by construction and obviously $\mathfrak{m} \longleftrightarrow \mathfrak{m}'$. For another EVM $\mathfrak{n}$ in $\mathfrak{m}$'s equivalence class we can similarly construct a pairwise linearly independent EVM $\mathfrak{n}'$. Then $\mathfrak{n}'$ and $\mathfrak{m}'$ are equal upto their labeling (lemma 1). This proves the theorem.

□

PROOF OF THEOREM 5

⇒

Since our Hilbert space is finite dimensional, there is an affine decomposition into (not necessarily orthogonal) one dimensional projectors for

every positive operator, and consequently for every $M_k \in \{M_k\}_K$, too:

(7) $$M_k = \sum_{m=0,1,\ldots} c_{km} |\psi_{km}\rangle\langle\psi_{km}| \quad (k \in K;\ c_{km} \geq 0).$$

The set $L := \{(k,m) \in \mathbb{F} \times \mathbb{N} \mid c_{km} \neq 0\}$ can be mapped into $\mathbb{F}$. Consequently this decomposition gives us an EVM $n = \{N_{km}\}_L$ satisfying $n \longrightarrow m$:

(8) $$N_{km} := c_{km} |\psi_{km}\rangle\langle\psi_{km}| \quad [(k,m) \in L].$$

Hence, for a given EVM m there is always an EVM $n = \{N_\ell\}_L$ such that $n \longrightarrow m$ and $\forall_{\ell \in L} N_\ell \in \partial\mathcal{B}^+$. But since we assumed m to be maximal, we must have

(9) $$N_\ell = \sum_{k \in K} \lambda_{\ell k} M_k \quad (\ell \in L;\ \{\lambda_{\ell k}\} \text{ a non-ideality matrix}).$$

The fact that $\forall_{\ell \in L} N_\ell \in \partial\mathcal{B}^+$ means that all non-zero elements on the right-hand side of (9) must be $\sim N_\ell$. Since there can be no $k \in K$ for which there is no $\ell \in L$ such that $\lambda_{\ell k} > 0$, it follows that $\forall_{k \in K} M_k \in \partial\mathcal{B}^+$.

$\Leftarrow$

Suppose that m satisfies the premise (i.e. $\forall_{k \in K} M_k \in \partial\mathcal{B}^+$). We can, just as in the proof of th. 4, construct a pairwise linearly independent EVM $m' = \{M'_m\}_M$ such that $m' \longleftrightarrow m$ and $\forall_{m \in M} M'_m \in \partial\mathcal{B}^+$.
Suppose further that there is a pairwise linearly independent EVM $n = \{N_\ell\}_L$ satisfying $n \longrightarrow m'$. Then there exists a non-ideality matrix $\{\lambda_{m\ell}\}$ such that:

$$M'_m = \sum_{\ell \in L} \lambda_{m\ell} N_\ell \quad (m \in M).$$

Since $\forall_{m \in M} M'_m \in \partial\mathcal{B}^+$ and m' is pairwise linearly independent, there is for every $\ell \in L$ precisely one $m \in M$ such that $N_\ell \sim M'_m$. Moreover, since $\sum_{m \in M} \lambda_{m\ell} = 1$, for every $\ell \in L$ such that $N_\ell \neq 0$ there is a $m \in M$ such that $N_\ell = M'_m$. Hence $m \longrightarrow n$.

□

LEMMA 2 For every EVM $\mathfrak{m}$ there is a self-extremal and pairwise linearly independent EVM $\mathfrak{n} = \{N_\ell\}_L$ such that $\mathfrak{n} \xrightarrow{i} \mathfrak{m}$ and $\forall_{\ell \in L} \exists_{k \in K} N_\ell \sim M_k$.

PROOF OF LEMMA 2

The set $B(\mathfrak{m})$ is closed since $K(\mathfrak{m})$ is closed and it is bounded since:

$$\sup_{X, Y \in B(\mathfrak{m})} \|X - Y\| \leq 2 \sup_{X \in B(\mathfrak{m})} \|X\| \leq$$

$$\leq 2 \sup_{X \in B(\mathfrak{m})} \mathrm{Tr}(X) = 2 < \infty .$$

The set $B(\mathfrak{m})$ is also convex. It is a subset of the space $\mathscr{B}(\mathscr{H})$ of linear operators on $\mathscr{H}$. Since $\mathscr{H}$ is finite dimensional, so is $\mathscr{B}(\mathscr{H})$. Therefore, using Minkowski's and Carathéodory's theorems (see app. A), we see that every element of $B(\mathfrak{m})$ can be written as a convex sum of finitely many elements of $\partial_E B(\mathfrak{m})$. It is not difficult to see that[4]:

(10) $$\partial_E B(\mathfrak{m}) = B(\mathfrak{m}) \cap \partial K(\mathfrak{m}) .$$

For these reasons there is a countable set $\{X_\ell\}_L \subseteq \partial_E B(\mathfrak{m})$ such that we can write:

$$M_k = \sum_{\ell \in L} \alpha_{k\ell} X_\ell \quad (k \in K;\ \alpha_{k\ell} \geq 0).$$

We can assume that for every $\ell \in L$ there is a $k \in K$ such that $\alpha_{k\ell} > 0$. If this were untrue for some $\ell_0 \in L$, we would not have included X_{ℓ_0} in $\{X_\ell\}_L$. Moreover, $\sum_{k \in K} \alpha_{k\ell}$ has to be finite because $\sum_{k \in K} M_k = 1$ and $X_\ell \neq 0$ for all $\ell \in L$. Now define:

$$N_\ell := (\sum_{k \in K} \alpha_{k\ell})\, X_\ell \quad (\ell \in L);$$

$$\lambda_{k\ell} := \alpha_{k\ell} (\sum_{k \in K} \alpha_{k\ell})^{-1} \quad (k \in K;\ \ell \in L).$$

---

[4]See e.g. G. Jameson (1970): *Ordered Linear Spaces* (Lecture Notes in Mathematics **141**, Springer, NY), § 1.9.

It follows that $n = \{N_\ell\}_L$ is a self-extremal and pairwise linearly independent EVM, and that $\{\lambda_{k\ell}\}$ is a non-ideality matrix. Hence $n \rightarrow m$.

Since $\forall_{\ell \in L} N_\ell \in K(m)$, there must be a non-negative matrix $\{\beta_{\ell k}\}$ such that:

$$N_\ell = \sum_{k \in K} \beta_{\ell k} M_k \quad (\ell \in L).$$

Since also $\forall_{\ell \in L} N_\ell \in \partial K(m)$, all non-zero terms on the right-hand side must be $\sim N_\ell$. Consequently for every $m \in L$ there is an $\ell \in L$ such that $N_m \sim M_\ell$.

□

Note that, although the relation of lemma 2 is stronger than $\xrightarrow{i}$, we cannot in general achieve equivalence.

PROOF OF THEOREM 6

$\Rightarrow$

We can, for a given EVM m, construct a pairwise linearly independent EVM $n = \{N_\ell\}_L$ such that $n \xrightarrow{i} m$ and $\forall_{\ell \in L} N_\ell \in \partial K_{max}(m)$, using an algorithm similar to the one we used in the proof of lemma 2.
Assume m is i–maximal. We must then also have:

(11) $$N_\ell = \sum_{k \in K} \lambda_{\ell k} M_k \quad (\ell \in L;\ \{\lambda_{\ell k}\} \text{ a non–ideality matrix}).$$

Since $\forall_{\ell \in L} N_\ell \in \partial K_{max}(m)$ all non–zero terms on the right–hand side of (11) must be $\sim N_\ell$. There can be no $k \in K$ for which there is no $\ell \in L$ such that $\lambda_{\ell k} > 0$, so it follows that $\forall_{k \in K} M_k \in \partial K_{max}(m)$.

$\Leftarrow$

Suppose that m satisfies the premise (i.e. $\forall_{k \in K} M_k \in \partial K_{max}(m)$). We can, just as in the proof of th. 4, construct a pairwise linearly independent

EVM $\mathfrak{m}' = \{M'_m\}_M$ satisfying:

$\forall_{m\in M} M'_m \in \partial K_{max}(\mathfrak{m}) = \partial K_{max}(\mathfrak{m}')$ and $\mathfrak{m}' \longleftrightarrow \mathfrak{m}$.

Suppose $\mathfrak{n} = \{N_\ell\}_L$ satisfies $\mathfrak{n} \xrightarrow{i} \mathfrak{m}$, and is pairwise linearly independent. Hence there exists a non-ideality matrix $\{\lambda_{m\ell}\}$ such that:

$M'_m = \sum_{\ell\in L} \lambda_{m\ell} N_\ell \quad (m \in M)$.

Since $\forall_{m\in M} M'_m \in \partial K_{max}(\mathfrak{m}')$ and $\mathfrak{m}'$ is pairwise linearly independent we have $\forall_{\ell\in L} \exists!_{m\in M} N_\ell \sim M'_m$. Combining this with $\sum_{m\in M} \lambda_{m\ell} = 1$ gives $\forall_{\ell\in L} \exists_{m\in M} N_\ell = M'_m$. Hence $\mathfrak{m} \rightarrow \mathfrak{n}$.

□

PROOF OF THEOREM 7

This proof of is simple: every EVM $\mathfrak{n}$ such that $\mathfrak{n} \leftarrow i$ is equivalent to i, and for every EVM $\mathfrak{n}$ it is true that $i \leftarrow \mathfrak{n}$.

□

PROOF OF THEOREM 9

If $\dim(\mathcal{H}) = 2$, we only have to show that if $\mathfrak{m}$ is i–maximal and not minimal, it is maximal.
We can always write:

$M_k = c_{k1}E_k + c_{k2}\bar{E}_k = (c_{k1} - c_{k2})E_k + c_{k2}1 \quad (k \in K)$.

Here E_k is a projector onto a 1–dimensional subspace, and $\bar{E}_k$ is its orthogonal complement. We assume (without loss of generality) that $c_{k1} \geq c_{k2}$ for all $k \in K$. There must be at least one $k_0 \in K$ for which $c_{k_0 1} > c_{k_0 2}$, since $\mathfrak{m}$ is not minimal (th. 7). This means that

$c_{m1} = c_{m2} \ (k_0 \neq m \in K) \Rightarrow M_m \in L(i) \subsetneq K(\mathfrak{m})\backslash\partial K(\mathfrak{m})$.

Since m is i–maximal, there can be no such m (th. 6). Consequently we must have $c_{k1} > c_{k2}$ for all $k \in K$. Then it follows that

$$\partial K(\mathfrak{m}) = \{X \mid \exists_{k\in K}\, 0 \leq X \sim E_k\} \quad .$$

Since $\forall_{k\in K}\, \{E_k, \bar{E}_k\} \subseteq \partial\mathcal{B}^+$ and an i–maximal EVM is self-extremal, this means (th. 5) that m is maximal.

If $\dim(\mathcal{H}) > 2$, the PVM $\{E,\bar{E}\}$ $(E \neq 0,\ E \neq 1)$ is i–maximal, but neither maximal nor minimal.

□

PROOF OF THEOREM 10

For the first part of the theorem we only need to prove that an i–minimal EVM is minimal (cf. th. 8). Suppose m is i–minimal, but not minimal. Choose an EVM n with the same outcome as m, such that $\forall_{k\in K}\, N_k = c_k 1 \neq 0$. Define:

$$O_k := \lambda M_k + (1-\lambda)N_k \quad (k \in K;\ 0 < \lambda < 1).$$

Obviously $o = \{O_k\}_K$ is an EVM satisfying $m \xrightarrow{i} o$. Because of our assumption we must also have $o \longrightarrow m$, and thus:

$$M_k = \sum_{\ell\in K} \gamma_{k\ell} M_\ell \quad (k \in K;\ \gamma_{k\ell} \text{ a non-ideality matrix}).$$

In view of the definition of o, the matrix $\gamma_{k\ell}$ can be chosen:

$$\gamma_{k\ell} = \lambda\mu_{k\ell} + (1-\lambda)\left[\sum_{m\in K} \mu_{km} c_m\right] \quad ;$$

$\{\mu_{km}\}$ a non–ideality matrix for $o \longrightarrow m$.

It can be seen that for the matrix $\{\gamma_{k\ell}\}$ the whole index set K is minimal closed. Hence we can use reasoning similar to that of the proof of lemma 1 to complete this proof.

The second part of the theorem is immediately clear if one considers th. 7, def. III.3 and the above.
□

PROOF OF THEOREM 11

For a given EVM $\mathfrak{m}$ one can find a maximal EVM $\mathfrak{n} \rightarrow \mathfrak{m}$ simply by using (8), thus proving the first part of the theorem. Because the decomposition (8) is usually not unique, there exist in general many non-equivalent maximal EVM's related to a given EVM.

We now proceed with the second assertion of the theorem. For a given EVM $\mathfrak{m}$ one can find an i–maximal EVM $\mathfrak{n} \xrightarrow{i} \mathfrak{m}$ by repeating the construction of the proof of lemma 2, using $K_{\max}(\mathfrak{m})$ instead of $K(\mathfrak{m})$. Because the corresponding base $B_{\max}(\mathfrak{m})$ is not always a simplex, this i–maximal EVM is in general also not unique.

The third assertion of the theorem is trivial.

The fourth is a straightforward corollary of th. 10.
□

Note that, if $\dim(\mathscr{H}) = 2$, the conditions of th. 12 and th. 13 coincide because of th. 9. This proves the converse part of th. 12 and th. 13 for $\dim(\mathscr{H}) = 2$.
In the following two proofs we shall use the fact that i–maximal EVM's are self–extremal (def. III.8).

PROOF OF THEOREM 12

First the counter–example. If $\dim(\mathscr{H}) > 2$ there is a PVM $\{E_1, E_2, E_3\}$ ($\forall_i$ $E_i \neq 0$). Define $\mathfrak{m} = \{M_1, M_2\} := \{E_1 + \frac{1}{2}E_2 , \frac{1}{2}E_2 + E_3\}$. This binary EVM is i–maximal, pairwise linearly independent and $B(\mathfrak{m})$ is a simplex. It is not extreme.

The major part of the theorem's proof consists of four stages:

a) First assume that m is not pairwise linearly independent. Then (viz. proof of th. 4) there is a pairwise linearly independent EVM $o = \{O_\ell\}_L$ such that

$$M_k = \sum_{\ell \in L} \lambda_{k\ell} O_\ell \quad (k \in K)$$

with non-ideality matrix $\{\lambda_{k\ell}\}$ satisfying $\forall_k \exists!_\ell \lambda_{k\ell} > 0$. The matrix $\{\lambda_{k\ell}\}$ cannot consist entirely of 1's and 0's because m is not pairwise linearly independent. Such a matrix $\{\lambda_{k\ell}\}$ can be written as a convex sum of two different non-ideality matrices. Consequently m can be written as a convex sum of two other EVM's, and is not extreme.

b) Now assume that m is pairwise linearly independent, but not self-extremal (def. III.8). Then (lemma 2) there is a pairwise linearly independent self-extremal EVM $o = \{O_\ell\}_L$ such that $o \rightarrow m$ and:

$$M_k = \sum_{\ell \in L} \lambda_{k\ell} O_\ell \quad (k \in K);$$

$$\forall_{\ell \in L} \exists_{k \in K} \forall_{m \neq \ell} \lambda_{km} = 0 \ .$$

The matrix $\{\lambda_{k\ell}\}$ cannot consist entirely of 1's and 0's because m is not self-extremal. We can use the above reasoning (stage a) to complete stage b.

c) Suppose m is self-extremal and pairwise linearly independent, but $B(m)$ is not a simplex. Since $L(m)$ must be finite dimensional, there must be a finite set $J \subseteq K$ such that $B(\{M_k\}_J)$ is not a simplex ($\Longleftrightarrow$ the number of elements of J is greater than $\dim[L(\{M_k\}_J)]$; see app. A).

Then there is an operator $X \in K(\{M_k\}_J)$ that can be written as $X = \sum_{k \in J} \nu_k M_k$ for *two* distinct non-negative finite sequences $(\nu_k)_J$.

Hence there is a non–trivial bounded sequence $(a_k)_K$ such that $\sum_{k\in K} a_k M_k = 0$. Define:

$$M_k^{\pm} := (1 \pm \lambda a_k) M_k \quad (k \in K;\ 0 < \lambda < \left(\sup_{k\in K}(|a_k|)\right)^{-1}).$$

Under these conditions $\{M_k^{\pm}\}_{k\in K}$ are different EVM's, and:

$$M_k = \tfrac{1}{2}(M_k^{+} + M_k^{-}) \quad (k \in K).$$

Hence m is not extreme.

d) Now suppose $B(\mathfrak{m})$ is a simplex, m is self-extremal and pairwise linearly independent, but not i–maximal. Then (proof of th. 11 (part 2), lemma 2) there is a pairwise linearly independent i–maximal EVM

$$\mathfrak{o} = \{O_\ell\}_L \xrightarrow{i} \mathfrak{m} \ .$$

If the matrix $\{\lambda_{k\ell}\}$ in $M_k = \sum_{\ell\in L} \lambda_{k\ell} O_\ell$ does not consist entirely of 1's and 0's, we can use the reasoning of stage a to see that m is not extreme. If this matrix $\{\lambda_{k\ell}\}$ does consist entirely of 1's and 0's, the set $B(\mathfrak{o})$ cannot be a simplex because $L(\mathfrak{o}) = L(\mathfrak{m})$ and $B(\mathfrak{o})$ contains more extreme elements than $B(\mathfrak{m})$. Hence o is not an extreme EVM (stage c) and can be written as the convex sum of two different EVMs. These lead, through this same $\{\lambda_{k\ell}\}$, to two different EVMs of which m is a convex sum. Consequently under these circumstances m is not extreme either.

□

PROOF OF THEOREM 13

First the counter–example. A PVM $\{E,\bar{E}\}$ ($E \neq 0$, $E \neq 1$) is extreme, but neither minimal nor maximal if $\dim(\mathscr{H}) > 2$.

If m is minimal, self–extremal and pairwise linearly independent it must be equal to i up to labeling. Such an EVM is clearly extreme.

A more meaningful statement results if we assume that m is maximal and satisfies the other criteria, too. We then have (use th. 5):

$$M_k = \lambda M_k^{(1)} + (1-\lambda) M_k^{(2)} \Rightarrow (\lambda = 0 \vee M_k^{(1)} \sim M_k) \; ;$$

for all EVMs $\{M_k^{(1)}\}_{k\in K}$ and $\{M_k^{(2)}\}_{k\in K}$.

Therefore there is a non–negative sequence $(\nu_k)_K$ such that $M_k^{(1)} = \nu_k M_k$. But since m is self-extremal, pairwise linearly independent and $B(m)$ is a simplex there can be only one sequence $(\nu_k)_K$ such that $\sum_{k\in K} \nu_k M_k = 1$, namely $\nu_k = 1$. Hence it follows that $M_k^{(1)} = M_k$.

□

2 AD § 2.3

PROOF OF THEOREM 17

$$\sum_{k\in K} |\mathrm{Tr}\left[(\rho_1-\rho_2)M_k\right]| = \sum_{k\in K} |\mathrm{Tr}\left[(\rho_1-\rho_2)\sum_{\ell\in L}\lambda_{k\ell}N_\ell\right]| \leq$$

$$\leq \sum_{k\in K}\sum_{\ell\in L} |\mathrm{Tr}\left[(\rho_1-\rho_2)\lambda_{k\ell}N_\ell\right]| =$$

$$= \sum_{k\in K}\sum_{\ell\in L} \lambda_{k\ell}|\mathrm{Tr}\left[(\rho_1-\rho_2)N_\ell\right]| =$$

$$= \sum_{\ell\in L} |\mathrm{Tr}\left[(\rho_1-\rho_2)N_\ell\right]| \; .$$

□

Theorem 18 is a simple consequence of th. 17 and def. III.4. Theorems 19 and 20 are corollaries of the following lemma:

LEMMA 3 Suppose that $\mathfrak{n}$ and $\mathfrak{m}$ are two EVMs as above, such that $\mathfrak{n} \longrightarrow \mathfrak{m}$ with non-ideality matrix $\{\lambda_{k\ell}\}$. A sequence $(g_k)_K$ is given such that $\sum_{k\in K} g_k M_k$ and $\sum_{k\in K} g_k^2 M_k$ converge. Define:

$$f_\ell := \sum_{k\in K} g_k \lambda_{k\ell} \quad . \tag{12}$$

Then $\sum_{\ell\in L} f_\ell N_\ell = \sum_{k\in K} g_k M_k$ and:

$$\sum_{\ell\in L} f_\ell^2 N_\ell \leq \sum_{k\in K} g_k^2 M_k \quad .$$

PROOF OF LEMMA 3

It is easily seen that the first part is true. We shall proceed with the proof of the second part:

$$\sum_{k\in K} g_k^2 M_k - \sum_{\ell\in L} f_\ell^2 N_\ell =$$

$$= \sum_{\ell\in L} N_\ell \left\{ \sum_{k\in K} g_k^2 \lambda_{k\ell} - \left(\sum_{k\in K} g_k \lambda_{k\ell}\right)^2 \right\} =$$

$$= \sum_{\ell\in L} N_\ell \sum_{k\in K} \sum_{k'\in K} g_k \left\{ \delta_{kk'} \lambda_{k\ell} - \lambda_{k\ell}\lambda_{k'\ell} \right\} g_{k'} =$$

$$= \sum_{\ell\in L} N_\ell \sum_{k,k'\in K} (g_k \sqrt{\lambda_{k\ell}}) \underbrace{\left\{ \delta_{kk'} - \sqrt{\lambda_{k\ell}}\sqrt{\lambda_{k'\ell}} \right\}}_{:= A_{kk'}^{(\ell)}} (g_{k'} \sqrt{\lambda_{k'\ell}}) \quad .$$

If we regard $\{A_{kk'}^{(\ell)}\}$ as a matrix-valued function of ℓ, it is seen to be equal to the identity matrix minus the projector onto the vector $(x_k^{(\ell)})_K$; $x_k^{(\ell)} = \sqrt{\lambda_{k\ell}}$. This vector has norm 1, since:

$$\sum_{k\in K} (x_k^{(\ell)})^2 = \sum_{k\in K} \lambda_{k\ell} = 1 \quad .$$

Therefore the summation over k and m always results in a non-negative number, so that:

$$\sum_{k \in K} g_k^2 \, M_k - \sum_{\ell \in L} f_\ell^2 \, N_\ell \geq 0 \quad .$$

□

Theorems 19 and 20 are corollaries of lemma 3

PROOF OF THEOREM 21

The first part of the theorem is obvious. We therefore proceed with the proof of the second part. We can construct, just as in the proof of th. 4, two pairwise linearly independent EVM's m' and n' such that $m' \longleftrightarrow m$ and $n' \longleftrightarrow n$. If a sequence $(f_\ell)_L$ for n is such that there is none better (alternative i) false), this sequence must have the property that $f_\ell = f_m$ for all $\ell, m \in L$ for which $N_\ell \sim N_m$. Through $m \longleftrightarrow m' \longleftrightarrow n' \longleftrightarrow n$ we can then for every such sequence explicitly construct a sequence $(g_k)_K$ for m that satisfies the demands.

□

Note that lemma 3 implies that alternative i) of th. 21 is true iff there is a sequence $(g_k)_K$ such that:

$$\sum_{k \in K} g_k M_k = F \wedge \sum_{k \in K} g_k^2 \, M_k \begin{Bmatrix} \leq \\ \neq \end{Bmatrix} \sum_{\ell \in L} f_\ell^2 \, N_\ell \quad .$$

3 AD § 2.4

Theorem 22 is a consequence of (III.67) and the following lemma[5]:

LEMMA 4 Suppose two non-ideality matrices $\{\lambda_{k\ell}\}$ and $\{\mu_{\ell m}\}$ ($k \in K$, $\ell \in L$, $m \in M$) are given. Then:

$$- \quad I\left(\{\textstyle\sum_{\ell\in L} \lambda_{k\ell}\mu_{\ell m}\}\,;\,(p_m)\right) \;\le\; I\left(\{\lambda_{k\ell}\}\,;\,(\textstyle\sum_{m\in M} \mu_{\ell m}p_m)\right) \quad ;$$

$$- \quad I\left(\{\textstyle\sum_{\ell\in L} \lambda_{k\ell}\mu_{\ell m}\}\,;\,(p_m)\right) \;\le\; I\left(\{\mu_{\ell m}\}\,;\,(p_m)\right) \quad ;$$

for all probability distributions $(p_m)_{m\in M}$.

PROOF OF THEOREM 23

First the counter-example. Consider a binary PVM $\mathfrak{e} = \{E_1, E_2\}$ and the equivalent EVM $\mathfrak{e}' = \{E_1, \frac{1}{2}E_2, \frac{1}{2}E_2\}$. We take as weighting distributions (p_ℓ) the sequences $(\frac{1}{2},\frac{1}{2})$ and $(\frac{1}{2},\frac{1}{4},\frac{1}{4})$, respectively. Note that these distributions correspond to the same density operator ρ. Only under that restriction are J's for $\mathfrak{e}$ and $\mathfrak{e}'$ comparable. We can write

$$J \;=\; -\textstyle\sum_\ell p_\ell \log(p_\ell) \;-\; I$$

I satisfies lemma 4, so that in particular $I_{\mathfrak{e}\to\mathfrak{n}} = I_{\mathfrak{e}'\to\mathfrak{n}}$. The entropies of the two distributions (p_ℓ) are, however, not equal. Therefore $J_{\mathfrak{e}\to\mathfrak{n}} \neq J_{\mathfrak{e}'\to\mathfrak{n}}$ for any non-ideal version $\mathfrak{n}$ of $\mathfrak{e}$.

Right-order preservation follows analogously to lemma 4.

□

[5]See e.g. R. McEliece (1977): *The Theory of Information and Coding* (Addison-Wesley, London), p. 27

4 AD § 2.5

Introduce two maximal PVMs $\mathfrak{x} = \{E_k\}_K$ and $\mathfrak{y} = \{F_k\}_K$,

(13) $$E_k = |x_k\rangle\langle x_k| \quad \text{and} \quad F_k = |y_k\rangle\langle y_k| \quad ,$$

where $(|x_k\rangle)_K$ and $(|y_k\rangle)_K$ (K = 0, ... , n–1) are orthonormal bases for $\mathbb{C}^n$ with

(14) $$f_{k\ell} := \langle x_k|y_\ell\rangle \quad .$$

$$X := \sum_{k',\ell'} C_{k'\ell'} E_{k'} F_{\ell'} + \sum_{k',\ell'} C^*_{k'\ell'} F_{\ell'} E_{k'} \geq 0$$

There are complex $U_{k\ell}$, $V_{k\ell}$, $\alpha_{k\ell}$ and real β_k, γ_ℓ such that:

(15)
$$\begin{cases} C_{k\ell} = \alpha_{k\ell} + \beta_k + \gamma_\ell \\ U_{k\ell} V^*_{k\ell} = \alpha_{k\ell} f_{k\ell} \\ 2\beta_k \geq \sum_\ell |U_{k\ell}|^2 \\ 2\gamma_\ell \geq \sum_k |V_{k\ell}|^2 \end{cases}$$

[$\{E_k\}_K$ and $\{F_k\}_K$ as in (13)].

PROOF OF LEMMA 5

Define the following vectors:

$$|c_{k\ell}\rangle := U_{k\ell}|x_k\rangle + V_{k\ell}|y_\ell\rangle \quad ;$$

and the projector (up to a scalar factor):

$$N_{k\ell} := |c_{k\ell}\rangle\langle c_{k\ell}| \geq 0 \quad .$$

Explicit calculation using (15) shows:

$$N_{k\ell} = |U_{k\ell}|^2 E_k + |V_{k\ell}|^2 F_\ell + \alpha_{k\ell} E_k F_\ell + \alpha^*_{k\ell} F_\ell E_k \ .$$

Moreover:

$$\sum_{k,\ell} N_{k\ell} \leq 2 \sum_k \beta_k E_k + 2 \sum_\ell \gamma_\ell F_\ell + \\ + \sum_{k,\ell} \alpha_{k\ell} E_k F_\ell + \sum_{k,\ell} \alpha^*_{k\ell} F_\ell E_k = X \ .$$

□

PROOF OF THEOREM 24

We shall first prove the theorem for maximal PVMs ξ and η as in (13). Consider the EVM $\mathfrak{m} = \{M_{k\ell}\}_{L\times M}$:

$$M_{k\ell} := \sum_{k',\ell'} C^{(k\ell)}_{k'\ell'} E_{k'} F_{\ell'} + \sum_{k',\ell'} C^{(k\ell)*}_{k'\ell'} F_{\ell'} E_{k'} \tag{16}$$

with:

$C^{(k\ell)}_{k'\ell'}$ complex $(k',\ell' \in K)$

$$\sum_k \sum_\ell C^{(k\ell)}_{k'\ell'} = \tfrac{1}{2} \ ; \ \sum_\ell C^{(k\ell)}_{k'\ell'} = A^k_{k'} \ ; \ \sum_k C^{(k\ell)}_{k'\ell'} = B^\ell_{\ell'}$$

That $M_{k\ell}$ is self-adjoint and that $\sum_k \sum_\ell M_{k\ell} = 1$ is evident. If each element of the operator valued measure defined by (16) satisfies the condition of lemma 5, $\mathfrak{m}$ is an EVM. Then the fact that $\mathfrak{m}$ is associated with a joint non-ideal measurement of ξ and η can be seen from:

$$\sum_{k\in L} M_{k\ell} = \sum_{m\in K} \lambda^{(2)}_{\ell m} F_m \ ; \ \sum_{\ell\in M} M_{k\ell} = \sum_{m\in K} \lambda^{(1)}_{km} E_m \ ;$$

with:

$$\lambda^{(2)}_{\ell m} = 2 \operatorname{Re}\left[B^\ell_m\right] \ ;$$

$$\lambda^{(1)}_{\ell m} = 2\,\mathrm{Re}\left[A^{\ell}_{m}\right] \quad .$$

Given that $\mathfrak{m}$ is an EVM, the matrices $\lambda^{(1)}_{k\ell}$ and $\lambda^{(2)}_{k\ell}$ are non–ideality matrices because $\mathfrak{x}$ and $\mathfrak{y}$ are PVM's.

This leaves us with the task of providing a set of coefficients $C^{(k\ell)}_{k'\ell'}$ that satisfies lemma 5, and the assertions of the theorem. First note that the EVM $\mathfrak{m}$ is a non–trivial joint measurement of $\mathfrak{x}$ and $\mathfrak{y}$ if we cannot separate $C^{(k\ell)}_{k'\ell'}$ into a sum of the form $\beta^{(k\ell)}_{k'} + \gamma^{(k\ell)}_{\ell'}$. This requirement, as well as invertibility, is satisfied by the following choice:

(17) $$C^{(k\ell)}_{k'\ell'} = \epsilon\Big[2\delta_{[[k-k']]0}\delta_{[[\ell-\ell']]0} + \\ + \sqrt{F}\,\delta_{[[\ell-\ell']]0} + \sqrt{F}\,\delta_{[[k-k']]0}\Big]$$

with:

$$[[a]] := a \bmod n \;\; ; \;\; L = M = K \;\; ;$$

$$F := \max_{k,\ell}\left\{|f_{k\ell}|^2\right\} \;\; ; \;\; \epsilon := (4 + 4n\sqrt{F})^{-1} \quad .$$

This example is also in accord with lemma 5, as we can see if we take:

$$U^{(k\ell)}_{k'\ell'} = V^{(k\ell)}_{k'\ell'} = \sqrt{2\epsilon}\,\delta_{[[k-k']]0}\,\delta_{[[\ell-\ell']]0}\sqrt{|f_{k'\ell'}|}\,\exp(i\tfrac{1}{2}\theta_{k'\ell'})$$

$$\alpha^{(k\ell)}_{k'\ell'} = 2\epsilon\,\delta_{[[k-k']]0}\,\delta_{[[\ell-\ell']]0}$$

$$\beta^{(k\ell)}_{k'} = \gamma^{(k\ell)}_{k'} = \epsilon\sqrt{F}\,\delta_{[[k-k']]0}$$

with:

$$f_{k\ell} = |f_{k\ell}|\exp(i\theta_{k\ell}) \;\; ;$$

and note that:

$$\sqrt{F} \geq |f_{k\ell}| \quad .$$

For the non–ideality matrices we have in this case:

(18) $$\lambda^{(1)}_{k\ell} = \lambda^{(2)}_{k\ell} = 2\epsilon\left[(2 + n\sqrt{F})\delta_{k\ell} + \sqrt{F}\right]$$

Now to the general case. Any two PVMs $e = \{\tilde{E}_I\}_{\mathcal{K}_e}$ and $f = \{\tilde{F}_J\}_{\mathcal{K}_f}$ can be decomposed into maximal PVMs as in (13):

(19) $$\tilde{E}_I = \sum_{k\in I} E_k \quad (I \in \mathcal{K}_e); \quad \tilde{F}_J = \sum_{k\in J} F_k \quad (J \in \mathcal{K}_f).$$

We have chosen (without loss of generality) the outcome sets $\mathcal{K}_e$ and $\mathcal{K}_f$ to be partitions of $K = \{0, 1, \ldots, n-1\}$. Then we can apply the analog of (19) to $m = \{M_{k\ell}\}_{K\times K}$ (defined by (16) and (17)):

$$\tilde{M}_{IJ} := \sum_{k\in I}\sum_{\ell\in J} M_{k\ell} \quad (I \in \mathcal{K}_e, J \in \mathcal{K}_f).$$

The EVM $\{\tilde{M}_{IJ}\}_{\mathcal{K}_e\times\mathcal{K}_f}$ thus defined is easily seen to satisfy the requirement of the theorem because of the particular form (18) of the non-ideality matrices above.

□

Note that formula (16) can easily be extended to the joint measurement of arbitrary non–coexistent EVM's. The analogous extension of (15) might not be so easy. In any case, however, we see that positivity conditions like (15) will prevent such a joint measurement from having arbitrarily small non–ideality.

PROOF OF THEOREM 25

For an EVM $o = \{O_j\}_J$ such that $o \leftarrow \mathfrak{x}$, the expression (III.78) can be rewritten as

(20) $$J_{e\longrightarrow n} = \sum_{j\in J} \frac{1}{n} \mathrm{Tr}(O_j)\, H_X\left[O_j/\mathrm{Tr}(O_j)\right] \quad ,$$

where of course

(21) $$H_X[\rho] := -\sum_{m\in M} \mathrm{Tr}(\rho\, E_m) \log\left[\mathrm{Tr}(\rho\, E_m)\right] \quad .$$

If o is in fact the marginal $m^{(1)}$ of an EVM $m = \{M_{ij'}\}_{J\times J'}$ that represents a joint non-ideal measurement of e and f (def. III.11) we get, in view of the concavity of the entropy functional H_X,

(22) $$J_{e\to m^{(1)}} \geq \sum_{j\in J} \frac{1}{n}\mathrm{Tr}(M^{(1)}_m) \sum_{j'\in J'} \frac{\mathrm{Tr}(M_{jj'})}{\mathrm{Tr}(M^{(1)}_j)} H_X\left[M_{jj'}/\mathrm{Tr}(M_{jj'})\right] \quad .$$

Proceeding analogously for the other marginal EVM ($m^{(2)}$) we get, using a functional H_Y defined analogously to H_X in (21),

(23) $$J_{e\to m^{(1)}} + J_{f\to m^{(2)}}$$

$$\geq$$

$$\sum_{j\in J}\sum_{j'\in J'} \left[H_X\left[M_{jj'}/\mathrm{Tr}(M_{jj'})\right] + H_Y\left[M_{jj'}/\mathrm{Tr}(M_{jj'})\right]\right] \frac{1}{n}\mathrm{Tr}(M_{jj'}) \, .$$

We resolve cases where the PVM's e and f have eigenspaces in common by introducing a third PVM $\{G_i\}_I$ satisfying

(III.75) $$\forall_i \, (\forall_k \; [G_i, E_k]_- = 0 \;\wedge\; \forall_\ell \; [G_i, F_\ell]_- = 0) \quad .$$

We loose no generality here, because this PVM may be minimal (i.e. $\{G_i\}_I = \{1\}$, which occurs when e and f have no eigenspaces in common). Its properties insure that

(24) $$H_X\left[M_{jj'}/\mathrm{Tr}(M_{jj'})\right] = H_X\left[\textstyle\sum_{i\in I} G_i M_{jj'} G_i/\mathrm{Tr}(M_{jj'})\right] \geq$$

$$\geq \textstyle\sum_{i\in I} \dfrac{\mathrm{Tr}(G_i M_{jj'})}{\mathrm{Tr}(M_{jj'})} H_X\left[G_i M_{jj'} G_i/\mathrm{Tr}(G_i M_{jj'})\right]$$

and analogously for H_Y. At this point we can use

(25) $$H_X[C/\mathrm{Tr}(C)] + H_Y[C/\mathrm{Tr}(C)] \geq \mathrm{Tr}(C\,R)/\mathrm{Tr}(C)$$

with $C \geq 0$ and

(26) $$R = \sum_i G_i c_i$$

(III.77) $$c_i := -2\log[\tfrac{1}{2} + \tfrac{1}{2}\max_{k,\ell}\|E_k F_\ell G_i\|^{\frac{1}{2}}] \quad ,$$

an inequality that can be derived[6] from

(III.5) $$\arccos\left[\langle A\rangle^{\frac{1}{2}}\right] + \arccos\left[\langle B\rangle^{\frac{1}{2}}\right] \geq \arccos\left[\|AB\|^{\frac{1}{2}}\right] \quad ,$$

which holds for any two projectors A and B. Eq. (25) limits the entropies on the right-hand side of (23), and yields, since $\sum_{j\in J}\sum_{m\in M} M_{mj} = 1$, the inequality

(III.76) $$J_{\mathfrak{e}\to\mathfrak{m}(1)} + J_{\mathfrak{f}\to\mathfrak{m}(2)} \geq \sum_{i\in I} \mathrm{Tr}(G_i)\, c_i \quad .$$

For the case of maximal (non-degenerate) PVMs as in (13) the Maassen-Uffink relation (III.3) can be used to improve (III.76) to

(27) $$J_{\mathfrak{x}\to\mathfrak{m}(1)} + J_{\mathfrak{y}\to\mathfrak{m}(2)} \geq -\sum_{i\in I} \mathrm{Tr}(G_i)\log[\max_{k,\ell}\|E_k F_\ell G_i\|] \quad .$$

The bound of (27) is non-trivial if the maximal PVMs have elements in common. The case where $\mathfrak{x}$ and $\mathfrak{y}$ are two maximal PVMs with

(28) $$\{f_{k\ell}\} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ 0 & \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{bmatrix} \quad ,$$

already treated in ch. III, furnishes an example. By taking $\{G_i\}_I = \{G_1, G_2\} = \{E_1, E_2+E_3\}$, we get the bound log(2) for (27). If, however, we have an inner product matrix close to this one, but not reducible to block form, like (28), our bound will be quite bad. This suggests that a better operator R in (25) can be found. Similarly, the bound of (III.76) is not optimal either.

---

[6] H. Maassen & J. Uffink (1988): *Phys. Rev. Lett.* **60**, p. 1103

Results of numerical work along these lines for three– and four–dimensional spaces inspire the conjecture that a bound as low as:

$$2\log(n) + \tfrac{1}{n}\sum_{k\in K}\log\left[\max_{k'\in K}|f_{kk'}|^2\right]$$

is achievable for (27).

□

5 AD § 2.6

PROOF OF THEOREM 26

$$S_x X S_x^\dagger = [[X-1]] = X - 1 + nE_0 \ ;$$

$$S_x X S_x^\dagger = S_x(\sum_{k\in K} f_k O_k)S_x^\dagger = \sum_{k\in K} f_{[[k-1]]} O_k \ .$$

Combining these two equations:

$$E_0 = \sum_{k\in K}\tfrac{1}{n}(1-f_k+f_{[[k-1]]})\,O_k \ .$$

Analogously we have:

$$S_x^2 X (S_x^\dagger)^2 = [[X-2]] = X - 2 + n(E_0 + E_1) \ ;$$

and accordingly:

$$E_1 = \sum_{k\in K}\tfrac{1}{n}(2 - f_k + f_{[[k-2]]})O_k - E_0$$

$$= \sum_{k\in K}\tfrac{1}{n}(1 - f_{[[k-1]]} + f_{[[k-2]]})O_k \ .$$

In this way we prove that there is a matrix $\{\mu_{k\ell}\}$ such that $E_k = \sum_{\ell\in K}\mu_{k\ell}O_\ell$. Since both ι and o consist of n elements, $\{\mu_{k\ell}\}$ is a $n\times n$ matrix. The range over which the vector $(x_k)_K$; $x_k = \mathrm{Tr}(\rho E_k)$ varies for

variable ρ is n–dimensional. Hence $\{\mu_{k\ell}\}$ is invertible. There is a matrix $\{\lambda_{k\ell}\}$ such that $O_k = \sum_{\ell\in K} \lambda_{k\ell} E_\ell$. Since $\mathfrak{x}$ is a PVM, $\{\lambda_{k\ell}\}$ has to have the properties of a non-ideality matrix if $\mathfrak{o}$ is to be an EVM.

□

Note that the EVM $\mathfrak{m}$ of def. III.14 for the joint non-ideal measurement of $\mathfrak{x}$ and $\mathfrak{y}$ [as defined in § III.2.6, incl. (III.82)], is also of the form (16). This can be seen by taking:

$$C^{(k\ell)}_{k'\ell'} = \tfrac{1}{2}\mathrm{Tr}(\rho_0 F_{[[\ell'-\ell]]} E_{[[k'-k]]}) \quad (\rho_0 = nM_{00})$$

PROOF OF THEOREM 27

⇒

This follows easily from the fact that:

$$\forall_{k\in K} \quad S_y E_k S_y^\dagger = E_k \ .$$

⇐

Every operator can be written as:

$$G = \sum_{a=0}^{n-1}\sum_{b=0}^{n-1} g_{ab}|x_a\rangle\langle x_b| \ ;$$

so that we can write:

$$S_y G S_y^\dagger = \sum_{a=0}^{n-1}\sum_{b=0}^{n-1} g_{ab}\,|x_a\rangle\langle x_b|\,\exp\left(j\tfrac{2\pi}{n}(a-b)\right) \ .$$

Combining these two equations gives $S_y G S_y^\dagger = G \iff g_{ab} = g'_a\,\delta_{ab}$. Hence G is diagonal in X–representation. If we now substitute O_ℓ for G, the fact that $\mathfrak{x}$ is a PVM implies that the coefficients involved constitute a non-ideality matrix.

□

PROOF OF THEOREM 28

This can be seen when one realizes that:

$$\exp\left[i\tfrac{2\pi}{n}[[X\otimes 1' + 1\otimes X']]\right] = \exp\left[i\tfrac{2\pi}{n}(X\otimes 1' + 1\otimes X')\right] =$$

$$= \exp(i\tfrac{2\pi}{n}X)\otimes\exp(i\tfrac{2\pi}{n}X')$$

and that:

$$\left[\exp(i\tfrac{2\pi}{n}X)\otimes\exp(i\tfrac{2\pi}{n}X') \,,\, \exp(i\tfrac{2\pi}{n}Y)\otimes\exp(i\tfrac{2\pi}{n}Y')\right]_- = 0 \quad ,$$

as follows from[7]

$$S_x^a S_y^b = S_y^b S_x^a \exp(-i\tfrac{2\pi}{n}ab) \quad (a,b \in \mathbb{I}). \tag{29}$$

□

PROOF OF THEOREM 29

We can write[8] $\rho_0 = \sum_{a,b} r_{ab} S_x^a S_y^b$. So:

$$\mathrm{Tr}(M_{k\ell} S_x^a S_y^b) = \tfrac{1}{n} \sum_{a',b'} r_{a'b'} \mathrm{Tr}(S_x^k S_y^\ell S_x^{a'} S_y^{b'} (S_x^k S_y^\ell)^\dagger S_x^a S_y^b) =$$

$$= \exp(i\tfrac{2\pi}{n}(-a\ell+kb-ab))\, r_{[[-a]][[-b]]}$$

Fourier transformation gives:

$$\sum_{k,\ell} \exp(i\tfrac{2\pi}{n}(ck+d\ell))\, \mathrm{Tr}(M_{k\ell} S_x^a S_y^b) =$$

$$= n^2 r_{[[-d]]c} \exp(i\tfrac{2\pi}{n}cd)\, \delta_{[[c+b]]0}\, \delta_{[[d-a]]0}$$

---

[7] J. Schwinger (1960): *Proc. Nat. Acad. Sc.* **46**, p. 570

[8] Schwinger, *op. cit.*

For an arbitrary ρ, write $\rho = \sum_{a,b} f_{ab} S_x^a S_y^b$ to get:

$$\sum_{k,\ell} \exp(i\tfrac{2\pi}{n}(ck+d\ell)) \operatorname{Tr}(M_{k\ell}\,\rho) \;=\; n^2\, r_{[[-d]]c} f_{d[[-c]]} \exp(i\tfrac{2\pi}{n}cd)$$

Thus, if the distribution $(\operatorname{Tr}[M_{k\ell}\,\rho])$ is known, $f_{d[[-c]]}$ can be calculated in the Fourier–domain iff $r_{[[-d]]c} \neq 0$ for all d,c.

□

Not surprisingly, the state separation condition of th. 29 is stronger than the condition for invertibility, which reads

(30) $$\forall_{a\in K} \;\; \operatorname{Tr}(\rho_0 S_x^a) \neq 0 \;\wedge\; \operatorname{Tr}(\rho_0 S_y^a) \neq 0 \quad ,$$

as can be proved analogously [ρ_0 as defined in (III.84)].

Samenvatting

Het onzekerheidsbeginsel is een van de meest karakteristieke punten van verschil tussen de quantummechanica en de klassieke mechanica. Dit beginsel wordt met de steeds verder voortschrijdende experimentele nauwkeurigheid relevant voor de praktijk. Daarnaast komen een aantal van de problemen van de interpretatie van de quantummechanica ook aan de orde in de interpretatie van het onzekerheidsbeginsel. Daarom is het feit dat in de gebruikelijke literatuur het beginsel wordt geïntroduceerd op een informele manier, die sinds 1930 nauwelijks nog ontwikkeling heeft vertoond, nogal onbevredigend. De gebruikelijke wijze van invoering, besproken in hoofdstuk I, is in feite afgeleid uit het werk van Niels Bohr. In hoofdstuk II is dat dan ook aan een nadere analyse onderworpen. Bohr's redeneringen, in tegenstelling tot echte berekeningen die het quantumformalisme op een niet-triviale manier gebruiken, zijn sterk intuïtief van aard. Dit heeft tot gevolg dat de draagwijdte van Bohr's visie op het onzekerheidsbeginsel niet ten volle tot uiting komt in de wiskundige relaties, zoals de Heisenberg-relaties, die uit de berekeningen volgen.

Bohr's interpretatie van de quantummechanica is niet de enig mogelijke. Daarenboven maken intuïtieve redeneringen niet ondubbelzinnig duidelijk welke typen van onzekerheidsrelaties er nu precies gelden. Daarom is verdere studie zinvol, en hoofdstuk III is daaraan gewijd. De verschillende kanten van het onzekerheidsbeginsel kunnen echter niet besproken worden binnen het kader van de quantummechanica zoals die in de gebruikelijke leerboeken gepresenteerd wordt, de quantummechanica die door von Neumann en Dirac is opgezet. We gebruiken dus een algemener formalisme: dat van Davies en Ludwig. Hierin komen *effect-waardige maten* als generalisatie van *Hermitische operatoren* voor om *observabelen* te representeren. In het bijzonder is aangetoond dat *incompatibele* observabelen alleen gelijktijdig meetbaar zijn met beperkte precisie: er is een *onnauwkeurigheidsrelatie* afgeleid uit het formalisme. Hierbij wordt een concretisatie van een 'niet-ideale' of 'onnauwkeurige' meting als partiële ordening op de verzameling van effect-waardige maten gebruikt.

Deze begrenzing van meetnauwkeurigheid vormt samen met de Heisenberg-relaties (*spreidings-relaties*) een *tweezijdig* onzekerheidsbeginsel. Dit dubbele beginsel stelt ons in staat een groot deel van de verwarring die rond de betekenis van het onzekerheidsbeginsel heerste, weg te nemen. Het representeert Bohr's onzekerheidsbeginsel afdoende, zoals aangetoond is in hoofdstuk IV: andere populaire versies van het beginsel, zoals de 'verstoringsinterpretatie', kunnen worden afgeleid. Het een en ander is verder geïllustreerd aan de hand van gedachtenexperimenten en uitvoerbare experimenten.

Dankwoord

Bij het werk waarvan in dit proefschrift verslag wordt gedaan, is natuurlijk de morele steun van vele mensen belangrijk geweest. In het bijzonder zou ik Johan Cuypers, Mariëtte v.d. Heuvel, Henk Stoof (zelfs op 7000 km afstand) en Leo Weegels willen noemen. Verder bedank ik Willem de Muynck voor zijn almaar groeiend enthousiasme, en Jan de Graaf en Frans Sluijter voor hun suggesties voor tekstuele verbeteringen. Jos Uffink heeft enige misverstanden mijnerzijds over de inhoud van § III.1 recht gezet.

Curriculum Vitae

1963	Geboren te Roermond;
1981	Diploma Atheneum B, Bisschoppelijk College Echt;
1987	Doctoraal Examen Technische Natuurkunde, TU Eindhoven (*cum laude*);
1988	In dienst van de Stichting Wijsgerig Onderzoek Nederland (SWON) als projectmedewerker.

Bibliography

O. Abu–Zeid (1987): *Phys. Lett. A* **125**, p. 167

L. Accardi, A. Frigerio & V. Gorini (eds.) (1984): *Quantum probability and applications to the quantum theory of irreversible processes* (Lecture Notes in Mathematics **1055**, Springer, Berlin)

S. Ali & E. Prugovečki (1977): *J. Math. Phys.* **18**, p. 219

S. Ali (1985): *Riv. Nuovo Cim.* **8**, p. 1

G. Allcock (1969): *Ann. of Phys.* **53**, p. 311

H. Araki & M. Yanase (1960): *Phys. Rev.* **120**, p. 622 [reprinted in Wheeler & Zurek (1983) p. 707]

E. Arthurs & J. Kelly (1965): *Bell Syst. Techn. J.* **44**, p. 725

E. Arthurs & M. Goodman (1988): *Phys. Rev. Lett.* **60**, p. 2447

A. Aspect & P. Grangier (1990): in Miller (1990), p. 45

L. Ballentine (1970): *Rev. Mod. Phys.* **42**, p. 358

W. Beckner (1975): *Ann. Math.* **102**, p. 159

V. Belavkin (1989): *Phys. Lett. A* **140**, p. 355, 359

E. Beltrametti & G. Casinelli (1981): *The Logic of Quantum Mechanics* (Addison-Wesley, Reading, Mass.)

E. Beltrametti, G. Casinelli & P. Lahti (1990): *J. Math. Phys.* **31**, p. 91

J. de Boer, E. Dal & O. Ulfbeck (eds.) (1986): *The Lesson of Quantum Theory* (North Holland, Amsterdam)

A. Böhm (1979): *Quantum Mechanics* (Springer, Berlin)

D. Bohm (1951): *Quantum Theory* (Prentice Hall, Englewood Cliffs, NJ)

N. Bohr, H. Kramers & J. Slater (1924): *Phil. Mag.* **47**, p. 785 [reprinted in van der Waerden (1968) p. 321 and in Bohr (1984), p. 99]

N. Bohr & L. Rosenfeld (1933): *Mat.-Fys. Medd. Dan. Vidensk. Selsk.* 12, no. 8 [reprinted in Wheeler & Zurek (1983) p. 479]

N. Bohr & L. Rosenfeld (1950): *Phys. Rev.* **78**, p. 794 [reprinted in Wheeler & Zurek (1983) p. 523]

N. Bohr (1927) [Como Lecture]: *Atti del Congresso Internazionale dei Fisici 1927, Como-Pavia-Roma* (Nicola Zanichelli, Bologna), p. 565 [reprinted in Bohr (1985) p. 109]

N. Bohr (1928) [Nature paper]: *Nature* (suppl.) **121**, p. 580 [reprinted in Wheeler & Zurek (1983) p. 85 and in Bohr (1985) p. 147 and Bohr (1934) p. 52]

N. Bohr (1929a) [Planck Paper]: *Naturwiss.* **17**, p. 483 [reprinted in Bohr (1985) p. 201 and in Bohr (1934) p. 92]

N. Bohr (1929b) [Scandinavian Natural Scientists Congress Lecture]: p. 102 of Bohr (1934) [reprinted in Bohr (1985) p. 219]

N. Bohr (1929c) [ATDN Introduction]: p. 1 of Bohr (1934) [reprinted in Bohr (1985) p. 257]

N. Bohr (1929d): "The Magnetic Electron", unpublished manuscript, p. 333 of Bohr (1985)

N. Bohr (1930) [Faraday Lecture]: *J. Chem Soc.* **26** (1932), p. 349 [reprinted in Bohr (1985) p. 371]

N. Bohr (1931a) [Maxwell Lecture]: *Nature* (suppl.) **128**, p. 691 [reprinted in Bohr (1985) p. 357]

N. Bohr (1931b) [Wills Lecture]: unpublished, p. 361 of Bohr (1985)

N. Bohr (1933) [Light and Life]: *Nature* **131**, p. 423 [reprinted in Bohr (1985) p. 147 and in Bohr (1958a) p. 3]

N. Bohr (1934): *Atomic Theory and the Description of Nature* (Cambridge University Press)

N. Bohr (1935) [EPR answer]: *Phys. Rev.* **48**, p. 696 [reprinted in Wheeler & Zurek (1983) p. 142]

N. Bohr (1939) [Warsaw Lecture]: *New Theories in Physics* (International Institute of Intellectual Co-operation, Paris) p. 11

N. Bohr (1948): *Dialectica* **2**, p. 312.

N. Bohr (1949) [Einstein essay]: in Schilpp (1949), p. 201 [reprinted in Wheeler & Zurek (1983) p. 5 and in Bohr (1958a) p. 32]

N. Bohr (1958a): *Atomic Physics and Human Knowledge* (Wiley, NY)

N. Bohr (1958b): in Klibansky (1958), p. 308 [reprinted in Bohr (1963) p. 1]

N. Bohr (1960) [Pharmaceutic Sciences Congress lecture]: in Bohr (1963) p. 17

N. Bohr (1962) [Light and Life revisited]: in Bohr (1963) p. 23

N. Bohr (1963): *Essays 1958-1962 on Atomic Physics and Human Knowledge* (Wiley, NY)

N. Bohr (1984): *Collected Works*, vol. 5 (ed. by J. Kalckar, North Holland, Amsterdam)

N. Bohr (1985): *ibid.*, vol. 6

R. Bondurant (1986): *Phys. Rev. A* **34**, p. 3927

M. Born, W. Heisenberg & P. Jordan (1926): *Zs. f. Phys.* **35**, p. 557 [reprinted in van de Waerden (1968) p. 321]

M. Born (1926a): *Zs. f. Phys.* **37**, p. 863 [reprinted in Wheeler & Zurek (1983) p. 52]

M. Born (1926b): *Zs. f. Phys.* **38**, p. 803

V. Braginskii (1988): *Sov. Phys. Usp.* **31**, p. 836

P. Busch & P. Lahti (1984): *Phys. Rev. D* **29**, p. 1634

P. Busch & P. Lahti (1985): *Phil Sc.* **52**, p. 64

P. Busch (1985a): *Int. J. Theor. Phys.* **24**, p. 63

P. Busch (1985b): *J. Phys. A* **18**, p. 3351

P. Busch (1987): *Found. Phys.* **17**, p. 905

A. Capri (1985): *Nonrelativistic Quantum Mechanics* (Benjamin–Cummings, Menlo Park, CA)

P. Carruthers & M. Nieto (1968): *Rev. Mod. Phys.* **40**, p. 411

H. Casimir (1986): in de Boer et al. (1986), p. 13

C. Caves (1981): *Phys. Rev. D* **23**, p. 1693

C. Caves (1985): *Phys. Rev. Lett.* **54**, p. 2465

D. Champernay (1973): *Fourier Transforms and their Physical Application* (Academic, NY)

K. Chung (1967): *Markov Chains with Stationary Transition Probabilities* (2nd ed., Springer, Berlin)

C. Cohen–Tannoudji, B. Diu & F. Laloë (1977): *Quantum Mechanics*, vol. I (Wiley, NY)

R. Colodny (ed.) (1972): *Paradigms and Paradoxes* (Pittsburgh University Press)

F. Combley (1979): *Rep. Progr. Phys.* **42**, p. 1887

O. Darrigol (1984): *Historical Studies in the Physical Sciences* **15**, no. 1, p. 38

E. Davies & J. Lewis (1969): *Comm. Math. Phys.* **17**, p. 239

E. Davies (1970): *J. Funct. Anal.* **6**, p. 318

E. Davies (1976): *Quantum Theory of Open Systems* (Academic, London)

P. Dirac (1925): *Proc. R. Soc. A* **109**, p. 642 [reprinted in van der Waerden (1968) p. 307]

P. Dirac (1927): *Proc. R. Soc. A* **113**, p. 621

P. Dirac (1930): *The Principles of Quantum Mechanics* (1st ed., Oxford Univ. Press)

P. Dirac (1958): *The Principles of Quantum Mechanics* (4th ed., Oxford Univ. Press)

V. Dodonov & V. Man'ko (1989): *Invariants and the Evolution of Nonstationary Quantum Systems* (ed. by M. Markov, Nova Science Publishers, Commack, NY), p. 3

N. Dunford & J. Schwartz (1957): *Linear Operators*, part I (Interscience, NY)

J. Eberly, L. Mandel & E. Wolf (eds.) (1990): *Coherence and Quantum Optics VI* (Plenum, NY)

W. Edelstein, J. Hough, J. Pugh & W. Martin (1978): *J. Phys. E* **11**, p. 710

A. Einstein, B. Podolsky & N. Rosen (1935): *Phys. Rev.* **47**, p. 777

A. Einstein (1948): *Dialectica* **2**, p. 320

A. Einstein (1949): Schilpp (1949), p. 665

R. Ellman (1982): *James Joyce* (rev. ed., Oxford University Press)

P. Feyerabend (1958): *Proc. Arist. Soc. (suppl. vol.)* **32**, p. 83.

P. Feyerabend (1981): *Realism, Rationalism and Scientific Method*, vol. 1 (Cambridge University Press)

R. Feynman, R. Leighton & M. Sands (1965): *Feynman lectures on Physics*, vol. 3 (Addison–Wesley, Reading, Mass.)

A. French & P. Kennedy (eds.) (1985): *Niels Bohr, a Centenary Volume* (Harvard University Press)

H. Groenewold (1946): *Physica* **12**, p. 405

A. Hartkämper & H.–J. Schmidt (eds.) (1981): *Structure and Approximation in Physical Theories* (Plenum, NY)

H. Haus & J. Mullen (1962): *Phys. Rev.* **128**, p. 2407 [reprinted in Wheeler & Zurek (1983) p. 725]

H. Heffner (1962): *Proc. IRE* **50**, p. 1604 [reprinted in Wheeler & Zurek (1983) p. 736]

W. Heisenberg (1925): *Zs. f. Phys.* **33**, p. 879 [reprinted in van der Waerden (1968) p. 261]

W. Heisenberg (1927): *Zs. f. Phys.* **43**, p. 172 [reprinted in Wheeler & Zurek (1983) p. 62]

W. Heisenberg (1967): in Rozental (1967), p. 94
W. Heisenberg (1969): *Der Teil und das Ganze* (Piper, Munich)
C. Helstrom (1974): *Found. Phys.* **4**, p. 453
C. Helstrom (1976): *Quantum Detection and Estimation Theory* (Academic, NY)
D. Hilbert, J. von Neumann & L. Nordheim (1927): *Math. Annal.* **98**, p. 1
J. Hilgevoord & J. Uffink (1988): in Selleri et al. (1988), p. 91
J. Hilgevoord & J. Uffink (1990): in Miller (1990), p. 121
M. Hillery, R. O'Connell, M. Scully & E. Wigner (1984): *Phys. Rep.* **106**, p. 123
A. Holevo (1972): *Trans. Moscow Math. Soc.* **26**, p. 133
A. Holevo (1982): *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam)
A. Holevo (1986): *Theor. & Math. Phys.* **65**, p. 1250
C. Hooker (1972): in Colodny (1972), p. 67
D. Howard (1989): in Miller (1990), p. 61
K. Husimi (1940): *Proc. Phys. Math. Soc. Japan* **22**, p. 264
N. Imoto, H. Haus & Y. Yamamoto (1985): *Phys. Rev. A* **32**, p. 2287
N. Imoto, S. Watkins & Y. Sasaki (1987): *Opt. Comm.* **61**, p. 159
G. Jameson (1970): *Ordered Linear Spaces* (Lecture Notes in Mathematics 141, Springer, Berlin)
M. Jammer (1974): *The Philosophy of Quantum Mechanics* (Wiley, NY)
M. Jammer (1985): in Lahti & Mittelstaedt (1985), p. 129
M. Jammer (1989): *The Conceptual Development of Quantum Mechanics* (2nd. ed., Tomash/American Institute of Physics)
J. Jauch (1968): *Foundations of Quantum Mechanics* (Addison–Wesley, Reading, Mass.)
P. Jordan (1927): *Zs. F. Phys.* **40**, p. 809
H. Kennard (1927): *Zs. f. Phys.* **44**, p. 326
V. Klee (1957): *Arch. Math.* **8**, p. 234
R. Klibansky (ed.) (1958): Philosophy in the Mid–Century (La Nuovo Italia, Florence)
K. Kraus & J. Schroeter (1983): *Int. J. Theor. Phys.* **7**, p. 431
K. Kraus (1983): *States, Effects and Operations* (Lecture Notes in Physics **190**, Springer, Berlin)
K. Kraus (1987): *Phys. Rev. D* **35**, p. 3070

P. Kruszynski & W. de Muynck (1987): *J. Math. Phys.* **28**, p. 1761

P. Lahti & P. Mittelstaedt (eds.) (1985): *Symposium on the Foundations of Modern Physics*, Joensuu, Finland (World Scientific, Singapore)

H. Landau & H. Pollak (1961): *Bell Syst. Techn. J.* **40**, p. 63

J.-M. Levy-Leblond (1972): *Am. J. Phys.* **40**, p. 899

J.-M. Levy-Leblond (1973): *Encart Pedagogique* 1 (suppl. au *Bull. Soc. Fra. Phys.* 14), p. 15

J.-M. Levy-Leblond (1976): *Ann. of Phys.* **101**, p. 319

R. Loudon (1983): *The Quantum Theory of Light* (2nd ed., Oxford University Press)

G. Lüders (1951): *Ann. der Phys.* **8**, p. 322

G. Ludwig (1981): in Hartkämper & Schmidt (1981) p. 7

G. Ludwig (1983): *Foundations of Quantum Mechanics*, 2 vols. (Springer, Berlin)

H. Maassen & J. Uffink (1988): *Phys. Rev. Lett.* **60**, p. 1103

L. Mandel'shtam & I. Tamm (1945): *Bull. Acad. Sci. USSR, Phys. Ser.* **9**, p. 249

H. Margenau (1936): *Phys. Rev.* **46**, p. 240

A. Marlow (ed.) (1978): *Mathematical Foundations of Quantum Theory* (Academic, NY)

H. Martens & W. de Muynck (1990a): *Found. Phys.* **20**, p. 257

H. Martens & W. de Muynck (1990b): *Found. Phys.* **20**, p. 355

H. Martens & W. de Muynck (1990c): submitted to *Phys. Lett. A*

H. Martens & W. de Muynck (1990d): submitted to *Found. Phys.*

H. Martens & W. de Muynck (1990e): paper presented at the International Workshop on *Quantum Aspects of Optical Communications*, Paris (France), to be published by Springer, Berlin.

H. Martens (1989): *Phys. Lett. A.* **137**, p. 155

R. McEliece (1977): *The Theory of Information and Coding* (Addison-Wesley, London)

E. McKinnon (1982): *Scientific Explanation and Atomic Physics* (University of Chicago Press)

E. McKinnon (1985): in French & Kennedy (1985), p. 101

E. McMullin (1954): *The Principle of Uncertainty* (PhD Thesis, Catholic University of Louvain, Belgium), unpublished

J. Mehra & H. Rechenberg (1982): *The Historical Development of Quantum Theory*, 6 vols. (Springer, Berlin)

A. Messiah (1955): *Quantum Mechanics*, vol. I (North Holland, Amsterdam)

A. Miller (ed.) (1990): *Sixty-two Years of Uncertainty* (NATO ASI B **226**, Plenum, NY)

P. Mittelstaedt, A. Prieur & R. Schieder (1987): *Found. Phys.* **17**, p. 891

N. Mott & H. Massey (1965): *The Theory of Atomic Collisions* (Clarendon Oxford) [Magnetic electron section reprinted in Wheeler & Zurek (1983) p. 701]

D. Murdoch (1987): *Niels Bohr's Philosophy of Physics* (Cambridge University Press)

W. de Muynck & J. van den Eijnde (1984): *Found. Phys.* **14**, p. 111

W. de Muynck & J. Koelman (1983): *Phys. Lett. A* **98**, p. 1

W. de Muynck & H. Martens (1990a): *Phys. Rev. A.* **42**, p. 5079

W. de Muynck & H. Martens (1990b): Proc. 3rd Int. Symp. *Foundations of Quantum Mechanics* (Physical Society of Japan, Tokyo), p. 171

W. de Muynck & H. Martens (1990c): in Eberly et al. (1990), p. 221

W. de Muynck, P. Janssen & A. Santman (1979): *Found. Phys.* **9**, p. 71.

W. de Muynck (1986): *Found. Phys.* **16**, p. 973

J. von Neumann (1932): *Mathematische Grundlagen der Quantenmechanik* (reprinted 1982, Springer, Berlin)

J. von Neumann (1961): *Collected Works*, vol. 1 (ed. by A. Taub, Pergamon, NY)

J. Ortega (1987): *Matrix Theory* (Plenum, NY)

M. Ozawa (1984): *J. Math. Phys.* **25**, p. 79

M. Ozawa (1989): in Tombesi & Pike (1989), p. 263

M. Ozawa (1990): "Does a Conservation Law Limit Position Measurement ?", Harvard University preprint HUTP–90/B002

J. Park & H. Margenau (1968): *Int. J. Theor. Phys.* **1**, p. 211

J. Park, W. Band & W. Yourgrau (1980): *Ann. der Phys.* **37**, p. 189

J. Park (1968): *Am. J. Phys.* **36**, p. 211

W. Pauli (1933): in *Handbuch der Physik* (2nd ed., ed. by H. Geiger & K. Scheel, Springer, Berlin), vol. 24, part I [reprinted as W. Pauli (1980): *General Principles of Quantum Mechanics* (Springer, Berlin)]

W. Pauli (1979): *Wissenschaftlicher Briefwechsel, 2 vols.* (ed. by A. Herman, K. von Meyenn & V. Weisskopf, Springer, Berlin)

S. Personick (1971): *Bell Syst. Techn. J.* **50**, p. 213

K. Popper (1972): *The Logic of Scientific Discovery* (6th rev. impr., Hutchinson, London)

K. Popper (1982): *Quantum Theory and the Schism in Physics* (ed. by W. Bartley, Rowman & Littlefield, Totowa, NJ)

H. Primas (1983): *Chemistry, Quantum Mechanics and Reductionism* (2nd ed., Springer, Berlin)

G. Prosperi (1984): in Accardi et al. (1984), p. 301

E. Prugovečki (1977): *J. Phys. A* **10**, p. 543

E. Prugovečki (1984): *Stochastic Quantum Mechanics and Quantum Space Time* (Reidel, Dordrecht)

H. Reichenbach (1944): *Philosophical Foundations of Quantum Mechanics* (Univ. of California Press, Berkeley)

H. Reichenbach (1980): *Selected writings 1909-1953*, vol. 2 (ed. by M. Reichenbach & R. Cohen, Reidel, Dordrecht)

H. Robertson (1929): *Phys. Rev.* **34**, p. 163

L. Rosenfeld (1971): *Arch. Hist. Exact Sci.* **7**, p. 69

S. Rozental (ed.) (1967): *Niels Bohr: his life and works as seen by his friends and colleagues* (North Holland, Amsterdam)

P. Schilpp (ed.) (1949): *Albert Einstein, Philosopher-Scientist* (Open Court, Evanston, IL)

F. Schroeck (1982): *Found. Phys.* **12**, p. 479

F. Schroeck (1982): *Found. Phys.* **12**, p. 825

F. Schroeck (1989): *Int. J. Theor. Phys.* **28**, p. 247

J. Schwinger (1960): *Proc. Nat. Acad. Sci.* **46**, p. 570

F. Selleri, A. van der Merwe & G. Tarozzi (eds.) (1988): Proceedings of the International Conference on *Microphysical Reality and Quantum Description*, Urbino, Italy (Reidel, Dordrecht)

C. Shannon (1948): *Bell Syst. Techn. J.* **27**, p. 379

C. Shannon (1958): *Inform. Control* **1**, p. 390

J. Shapiro & H. Yuen & J. Machado Mata (1979): *IEEE J. Inf. Th.* IT–25, p. 179

C. She & H. Heffner (1966): *Phys. Rev.* **152**, p. 1103

A. Shimony & H. Stein (1979): *Am. Math. Mon.* **86**, p. 292

D. Slepian & H. Pollak (1961): *Bell Syst. Techn. J.* **40**, p. 43

M. Srinivas (1985): *Pramana* **24**, p. 673

P. Suppes (1961): *Phil. of Sc.* **28**, p. 378

P. Tombesi & R. Pike (eds.) (1989): *Squeezed & Non-Classical Light* (NATO ASI B **190**, Plenum, NY)

S. Tsyplyaev (1989): *Sov. Phys. J.* **31**, p. 699

J. Uffink & J. Hilgevoord (1985): *Found. Phys.* **15**, p. 925

J. Uffink & J. Hilgevoord (1988): *Physica B* **151**, p. 309

J. Uffink (1990): *Measures of Uncertainty and the Uncertainty Principle* (PhD Thesis, University of Utrecht, Netherlands), unpublished

P. Vergilius: *Aeneis*

A. Vogt (1978): in Marlow (1978), p. 368

M. Vol'kenshtein (1988): *Sov. Phys. Usp.* **31**, p. 140

B. van der Waerden (ed.) (1968): *Sources of Quantum Mechanics* (Dover, NY)

H. Weyl (1928): *Gruppentheorie und Quantenmechanik* (Hirzel, Leipzig)

J. Wheeler & W. Zurek (eds.) (1983): *Quantum Theory and Measurement* (Princeton University Press)

E. Whittaker & G. Watson (1927): *A Course of Modern Analysis* (4th ed., Cambridge University Press)

E. Wigner (1952): *Zs. f. Phys.* **133**, p. 101

L. Wittgenstein: *Tractatus Logico-Philosophicus*

W. Wootters (1979): *Phys. Rev. D* **19**, p. 473

T.-Y. Wu (1986): *Quantum Mechanics* (World Scientific, Singapore)

Y. Yamamoto & H. Haus (1986): *Rev. Mod. Phys.* **58**, p. 1001

Y. Yamamoto, S. Machida, S. Saito, T. Yanagawa, M. Kitagawa & G. Björk (1990): *Progress in Opt.* (ed. by E. Wolf, North Holland, Amsterdam) **28**, p. 87

M. Yanase (1961): *Phys. Rev.* **123**, p. 666 [reprinted in Wheeler & Zurek (1983) p. 712]

H. Yuen & J. Shapiro (1978): *IEEE J. Inf. Th.* **IT-24**, p. 657

H. Yuen & J. Shapiro (1980): *IEEE J. Inf. Th.* **IT-26**, p. 78

H. Yuen (1982): *Phys. Lett. A* **91**, p. 101

H. Yuen (1983): *Phys. Rev. Lett.* **51**, p. 719

H. Yuen (1984): *Phys. Rev. Lett.* **52**, p. 1730

Stellingen

behorende bij het proefschrift van Hans Martens

1

De bewering van Hilgevoord & Uffink dat Heisenberg's γ-microscoop los staat van de problematiek rond gelijktijdige metingen, is onjuist.

J. Hilgevoord & J. Uffink (1990): in *Sixty-two Years of Uncertainty* (ed. by A. Miller, Plenum, NY), p. 121; dit proefschrift.

2

De rol van covariantie bij de bestudering van gelijktijdige metingen is zeer misleidend. Enerzijds vergemakkelijkt covariantie het bereiken van resultaten zeer, anderzijds zijn die resultaten in het geheel niet representatief. Daarom zijn algemene conclusies over eigenschappen van onnauwkeurigheidsrelaties op basis van covariante gelijktijdige metingen onmogelijk.

A. Holevo (1982): *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam); dit proefschrift.

3

De gebruikelijke definitie van (gegeneraliseerde) tijdmeting als een effectwaardige maat die covariant is ten opzichte van tijdverschuivingen, moet verzwakt worden tot de eis dat de effectwaardige maat compatibel is met tijdverschoven versies van zichzelf.

A. Holevo (1982).

4

De bewering van Ballentine (en anderen) dat in de "statistische" interpretatie een projectiepostulaat overbodig is, is onjuist. Wel is het zo dat vanuit een instrumentalistisch standpunt, de achterliggende filosofie van vele voorstanders van de "statistische" interpretatie, het meetprobleem niet opgelost *behoeft* te worden.

L. Ballentine (1970): *Rev. Mod. Phys.* 42, p. 358

5

De bewering van Beltrametti *et al.* dat 'objectivering' een essentieel kenmerk van meten is, is onjuist.

E. Beltrametti, G. Casinelli & P. Lahti (1990): *J. Math. Phys.* **31**, p. 91; dit proefschrift.

6

In het formalisme van de quantumoptica kunnen twee vormen van interferentie onderscheiden worden: klassieke (corresponderend met de coherente toestand $|\alpha+\beta\rangle$) en quantummechanische interferentie (corresponderend met de superpositie $|\alpha\rangle+|\beta\rangle$). Daarom is het twee-spletenexperiment (proef van Young) in het algemeen geen goede illustratie van specifiek quantummechanische effecten (incompatibiliteit, "golf-deeltje dualisme").

A. Aspect & P. Grangier (1990): *Sixty-two Years of Uncertainty* (ed. by A. Miller, Plenum, NY), p. 45.

7

De hoeveelheid energie die in neoklassieke "verklaringen" van de Bell-ongelijkheden wordt gestoken, in verhouding tot de energie die besteed wordt aan mechanische "verklaringen" van de Lorentz-contractie, zegt veel over de moeite die natuurkundigen hebben om te accepteren dat de realiteit op het microniveau van heel andere aard kan zijn dan die in onze dagelijkse ervaring.

8

In klassieke probabilistische theorieën treedt ook een meetprobleem op, als 'kans' in zo'n theorie ontisch geïnterpretreed wordt (bijvoorbeeld als propensity).

9

De proliferatie van eufemismen vormt een belangrijker bron van taalverloedering dan slechte spelling.

10

Het motto "mens agitat molem" blijkt bij juiste vertaling beter bij een theologische universiteit te passen, dan bij een technische.

Vergilius, *Aeneis* (boek VI).

11

De term 'zelfconsistente oplossing' is behalve een anglicisme ook een pleonasme: een niet in zichzelf consistente oplossing is geen oplossing. Ter aanduiding van een berekeningsmethode (in plaats van ter aanduiding van de oplossing) is de term 'successieve substitutie' aanzienlijk beter.