N95-13915

TWO NEW KINDS OF UNCERTAINTY RELATIONS

Jos Uffink

Department of History and Foundations of Mathematics and Science
University of Utrecht, P.O. Box 80.000, 3508 TA Utrecht, the Netherlands

Abstract

We review a statistical-geometrical and a generalized entropic approach to the uncertainty principle. Both approaches provide a strengthening and generalization of the standard Heisenberg uncertainty relations, but in different directions.

1 Introduction

The purpose of this note is to introduce two approaches to the uncertainty principle which have been developed recently, a statistical-geometrical approach and a generalized entropic approach. But before we go into this, let us consider why one would need a new approach at all. In other words, what is unsatisfactory with the traditional approach to the uncertainty principle? In the standard textbook approach the uncertainty principle for position and momentum is expressed by the inequality

$$\forall\psi: \quad \Delta_\psi P \Delta_\psi Q \geq \frac{\hbar}{2} \tag{1}$$

or more generally:

$$\forall\psi: \quad (\Delta_\psi A)^2(\Delta_\psi B)^2 \geq \frac{1}{4}\left(|\langle [A,B]_-\rangle_\psi|^2 + \langle [A-\langle A\rangle_\psi, B-\langle B\rangle_\psi]_+\rangle^2\right) \tag{2}$$

for arbitrary observables A and B. Here, $\Delta_\psi A$ etc. is defined as:

$$(\Delta_\psi A)^2 = \langle (A - \langle A\rangle_\psi)^2\rangle_\psi \tag{3}$$

There are three problems. First, uncertainty relations as (1) or (2) presuppose that all observables for which one wants to write down an uncertainty relation can be represented as self-adjoint (or at least normal [1]) operators. Unfortunately this is not always the case. Notorious examples are time and energy, and phase and photon number. Further, in relativistic quantum theory, even the status of the position observable becomes dubious. There is no self-adjoint position (vector) operator for photons [2].

Secondly, note that the right-hand side of (2) still depends on ψ. It may become zero, even if A and B do not commute. In fact, this always happens in an eigenstate of A or B. Then, taken as a general statement about $\Delta_\psi A \Delta_\psi B$, the inequality only says that this product is greater than zero for some states and equal to zero for others. That, however, is true also in classical physics. To read off more from (2), one needs to know the state. But then, when ψ is given, one can also calculate $\Delta_\psi A$ and $\Delta_\psi B$ directly, without using the inequality at all.

Even in the case where the right-hand side is always strictly greater than zero, as in relation (1), there are further problems, due to the properties of the standard deviations. In the definition (3) of the standard deviation, the probability density is integrated with a quadratic weight factor that puts most emphasis on the tails of the distribution. As a result, the standard deviation can become very high, even if 99% or more of the probability distribution is concentrated in a very small interval, and the remainder is located in long tails, as e.g. in a Breit-Wigner lineshape. Thus a large standard deviation does not necessarily prevent a probability distribution from being very sharply concentrated, and a bound on the product of standard deviations by itself does not prevent both observables from being as precisely determined as we please. In the next sections we ask whether there are more stringent inequalities that improve on the above aspects.

2 Statistical-geometrical approach

It is usual to assume in quantum theory that the state of the system is given. But in this section we consider an inverse problem. Suppose we don't know the state of the system. Our problem is to make a statistical inference about this state from given measurement results.

For definiteness, let us assume that some partial information about the state is given: it belongs to a given set of (pure) states labeled by an index parameter θ. To be more specific, it is asumed that these states are generated by some unitary group:

$$|\psi_\theta\rangle = e^{i\theta A/\hbar}|\psi\rangle \tag{4}$$

where A is a self-adjoint operator. We can think of this set of states as describing a curve in state space. The problem of statistical inference is now equivalent to that of estimating the value of θ.

It is clear that a detailed discussion of this estimation problem should involve the kind of measurements performed, the results obtained, and criteria distinguishing 'good' from 'bad' estimates. However, even without going into details of statistical theory [3, 4], it can be made plausible that a fundamental bound for the estimation accuracy is obtained by considering the overlap $|\langle\psi_\theta|\psi_{\theta+\delta\theta}\rangle|$. If this overlap is high the states resemble each other much and a typical measurement result which would be probable or improbable in one state would likewise be probable or improbable in the other. Then one cannot expect to discriminate the states by any measurement procedure. It is only when the overlap begins to fall off that there are observables whose probability distributions for the states ψ_θ and $\psi_{\theta+\delta\theta}$ differ enough to allow for accurate discrimination.

This suggests the following definition. Choose some fixed value $\beta < 1$ and define the estimation inaccuracy $\delta_\psi\theta$ as the smallest value of $\delta\theta$ for which

$$|\langle\psi_\theta|\psi_{\theta+\delta\theta}\rangle| = \beta$$

Due to the particular choice (4), this overlap does not depend on the value of θ. One can then show: [5, 6, 7, 8]

$$\forall\psi: \quad \delta_\psi\theta\Delta_\psi A \geq 2\hbar \arccos\beta \tag{5}$$

This then represents a useful uncertainty relation. It says that an unlimited increase in the estimation accuracy of θ is only possible at the expense of an increased spread in A.

Several remarks are called for. First, these relations are applicable to any one-parameter unitary group. Obvious examples are the translations in time, represented by the evolution operators

$U(t) = \exp(-itH/\hbar)$ or the translations in space, $\exp(ixP/\hbar)$, where P is the momentum operator; or we can consider an angle of rotation and angular momentum, or phaseshifts and photon number. In short, we find statistical uncertainty relations of the type (5) in every case where A is the generator of a unitary group and θ the group parameter. This approach is in fact ideally suited for a relativistical treatment of quantum theory in which one starts from the construction of unitary groups from the symmetries of the system.

Secondly, relation (5) is asymmetrical; $\delta_\psi \theta$ is an inaccuracy of estimation of a parameter (i.e. a *c-number*). $\Delta_\psi A$ on the other hand is the r.m.s. spread of a quantum mechanical *observable* (self-adjoint operator). Since one does not need a pair of operators to obtain relation (5) there are no problems when such a pair does not exist.

However if one does exist, e.g. in the case of non-relativistic position and momentum, it is possible to take advantage of that fact. Then there is a second, independent, uncertainty relation for the spread in position and the estimation accuracy of the parameter in the group of kicks $\{U(p) = \exp(-ipQ/\hbar)\}$, i.e. shifts in momentum. This rèstores symmetry between position and momentum. More importantly, we note that the position operator mimicks the parameter x of the translation group in the sense that

$$\langle \psi_x | Q | \psi_x \rangle = x \tag{6}$$

(assuming $\langle \psi | Q | \psi | \rangle = 0$) i.e. the position operator acts as an unbiassed estimator of the location parameter. From this it follows: [9]

$$|\langle \psi_\theta | \psi_{\theta+\delta\theta} \rangle|^2 \leq \left(1 + (\frac{\delta\theta}{2\Delta_\psi Q})^2\right)^{-1}$$

combined with (5), where A is interpreted as momentum, this result implies the standard uncertainty relation (1). In fact, as a bonus, we obtain (1) not only for the position operator proper, but for any other operator acting as unbiassed estimator of x as well.

There is only one problem of those mentioned in the previous section that is not solved by the relations (5): they still rely on one standard deviation, and thus become useless for states in which this diverges. To fix this problem, the standard deviation can be replaced by an interquantile range, i.e. the smallest size $W_\psi(A)$ of an interval W on which a fraction $\alpha < 1$ of the total probability distribution for A is concentrated: $\int_W |\langle \psi | a \rangle|^2 da = \alpha$. A variation of the proof of (5) gives [6, 8]

$$\delta_\psi \theta W_\psi(A) \geq \hbar \arccos \frac{1 + \beta - \alpha}{\alpha} \quad \text{if } \beta \geq 2\alpha - 1 \tag{7}$$

Finally we note that this concept of estimation inaccuracy fits into a general geometrical approach to statistical inference on the basis of the Fisher information metric [4]. Let it suffice here to note that this metric equips Hilbert space with a statistical distance between states which equals

$$d(\psi, \psi') = \arccos |\langle \psi | \psi' \rangle|$$

and that the geometrical background of (5) is the simple fact that the distance between ψ_θ and $\psi_{\theta'}$ (i.e. the right hand side of (5)) is less than the length of the curve (4) connecting these points. This also point the way to how the relations are to be generalized in cases where the curve is not generated by a unitary group.

3 Generalized entropic approach

For a discrete probability distribution $p = (p_1, \ldots p_n)$, with $p_i \geq 0$, $\sum_i p_i = 1$, the Shannon entropy

$$H(p) := -\sum_i p_i \log p_i \tag{8}$$

represents, roughly speaking, a measure of whether the distribution is 'spread out' or 'peaked'. If A and B are quantum observables for which an uncertainty principle holds, it is natural to ask for a lower bound of the sum of the Shannon entropies for the probabilities $|\langle\psi|a_i\rangle|^2$ and $\langle\psi|b_j\rangle|^2$ [10, 11]. Here we assume a discrete spectrum and $|a_i\rangle$ and $|b_j\rangle$ denote the eigenstates of A and B. It turns out that

$$H(\psi, A) + H(\psi, B) := H(|\langle\psi|a_i\rangle|^2) + H(|\langle\psi|b_j\rangle|^2) \geq -2 \log \sup_{ij} |\langle a_i|b_j\rangle| \tag{9}$$

This entropic uncertainty relation limits the concentration of both probability distributions by a bound which is independent of ψ.

There is a class of expressions that share many properties with the Shannon entropy, and also represent useful measures of 'peakedness' or concentration:

$$M_r(p) = (\sum_i p_i^{1+r})^{1/r} \tag{10}$$

Their properties have been studied extensively by Hardy, Littlewood and Pólya and by Renyi [12]. ($-\log M_r$ is known as Renyi entropy.) Special cases are:

$$M_\infty(p) = \sup_i p_i \quad , \quad M_0(p) = e^{-H(p)} \quad , \quad M_{-1}(p) = (\#\{i \; : \; p_i > 0\})^{-1} \tag{11}$$

Where # counts the number of elements in a set. The generalized entropic uncertainty relation then reads: [11, 13]

$$M_r(|\langle\psi|a_i\rangle|^2) M_s(|\langle\psi|b_j\rangle|^2) \leq \sup_{ij} |\langle a_i|b_j\rangle|^2 \quad \text{for } r = -s/(2s+1) \;\; r, s \geq -1/2 \tag{12}$$

which contains the relations (9) as the special case with $r = s = 0$.

Remarks: The above inequalities apply to any pair of discrete observables and yield a non-trivial bound iff these observables do not share an eigenstate. (A condition which is slightly stronger than mere non-commutativity.) In the case of a two-dimensional Hilbert space, the most restrictive bounds are obtained by the choice $r = -1/2, s = \infty$ or v.v.

Secondly, in the proof of (12) it is not necessary to assume that observables are represented by self-adjoint operators. It is sufficient to demand the existence of the sets of "eigenstates" $\{|a_i\rangle\}$ and $\{|b_j\rangle\}$, possibly non-orthogonal, such that

$$\sum_i |a_i\rangle\langle a_i| = \mathbb{I}, \quad \sum_i |b_i\rangle\langle b_i| = \mathbb{I} \tag{13}$$

Thus, the approach of this section is also applicable if one accepts unitary operators [14] or even more generally, positive-operator-valued measures (POVM's) [15] as *bona fide* representations of observables. As an example, we mention a phase observable below.

In the case of continuous observables, it appears necessary to replace (8) and (10) by a relative notion of entropy:

$$M_r(p|\mu) := \left(\int (\frac{\partial p}{\partial \mu})^{1+r} d\mu\right)^{1/r} \quad , \quad H(p|\mu) = -\log M_0(p|\mu) = -\int \frac{\partial p}{\partial \mu} \log \frac{\partial p}{\partial \mu} d\mu \tag{14}$$

where $\partial p/\partial \mu$ is the Radon-Nikodym derivative of a probability measure p with respect to a 'background measure' μ. These expressions reflect whether the probability distribution p is concentrated in comparison with μ. In the discrete case, the absolute entropies are recovered by choosing μ to be the counting measure $\#$. In the case of continuous observables it is natural to take for μ the Lebesgue measure, and $\partial p/\partial \mu$ becomes an ordinary probability density.

For continuous nondegenerate operators A and B a theorem of Haussdorf and Young analogous to (12) leads to [17, 13]

$$M_r(\psi, A|\mu) M_s(\langle\psi, Q, |\mu) := \left(\int |\langle\psi|a\rangle|^{2(1+r)} da\right)^{1/r} \left(\int |\langle\psi|b\rangle|^{2(1+s)} db\right)^{1/s} \leq \sup |\langle a|b\rangle|^2$$

(Still assuming $r = -s/(2s+1)$.) For position and momentum a slightly stronger inequality $M_r(\psi, P|\mu) M_s(\psi, Q|\mu) \geq 2(1+r)(1+2r)^{-(1+2r)/(2r)}(2\pi\hbar)^{-1}$ is obtained by a theorem of Beckner. These inequalities are all sharp for Gaussian 'minimum uncertainty' states and strictly imply the standard uncertainty relations (1).

Let N be the photon number operator of a single electromagnetic field mode, with eigenstates $|n\rangle$ and $\sum |n\rangle\langle n| = I\!I$. A description of a phase observable by means of a POVM Φ was constructed by Lévy-Leblond and by Susskind and Glogower [14, 16] from the non-orthogonal improper "eigenstates"

$$|\phi\rangle = (2\pi\hbar)^{-1/2} \sum_n e^{i\phi n} |n\rangle$$

This yield a resolution of identity $\int_0^{2\pi} d\phi |\phi\rangle\langle\phi| = I\!I$ in analogy with (13). Here one finds an analogous uncertainty relation

$$M_r(\psi, \Phi) M_s(\psi, N) \leq \frac{1}{2\pi\hbar} \tag{15}$$

where

$$M_r(\psi, \Phi) = \left(\int |\langle\psi|\phi\rangle|^{2(1+r)} d\phi\right)^{1/r} \quad , \quad M_s(\psi, N) = \left(\sum_{n=0}^{\infty} |\langle\psi|n\rangle|^{2(1+s)}\right)^{1/s}$$

Let us finally compare the two different approaches. Both improve on the standard approach in the sense that they yield bounds which are state independent and strictly imply $\Delta_\psi P \Delta_\psi Q \geq \hbar/2$. The statistical/geometrical approach is restricted to conjugate pairs of quantities: time/energy, location/momentum, angle/angular momentum etc. All such pairs are treated, on the same footing, as consisting of a parameter and an observable. It is essential that the observable is represented as a self-adjoint operator because of its role in (4). The approach is relativistically invariant. The entropic approach, on the other hand, is applicable to pairs of arbitrary observables, not necessarily conjugate pairs or even self-adjoint operators. It is sufficient that they do not share eigenstates. However, these uncertainty relations do not seem to have a relativistic generalization. Also, the last problem mentioned in the introduction is not completely overcome in this approach. IN the example of the Breit-Wigner distribution one still has the result $M_r(p|\mu) = 0$ for $r \leq 0$ and the inequality is not very restrictive.

References

[1] M. Bersohn, Am. J. Phys. **34**, 62 (1966).

[2] A.S. Wightman, Rev. Mod. Phys. **34**, 845 (1962); K. Kraus, in *Uncertainty Principle and Foundations of Quantum Mechanics*, W.C. Price and S.S. Chissick (eds), (Wiley, London, 1977).

[3] A. Bhattacharyya, Bull. Calcutta Math. Soc. **35**, 99 (1943); C.R. Rao, Bull. Calcutta Math. Soc. **37**, (1945).

[4] W.K. Wootters, Phys. Rev. D **23**, 357 (1981). J. Hilgevoord and J. Uffink, Found. Phys. **21**, 323 (1991).

[5] L. Mandelstam and I. Tamm, J. Phys. (USSR) **9**, 249 (1945); G.N. Fleming, Nuovo Cimento A **16**, 263 (1973); K. Bhattacharyya, J. Phys. A **16**, 2993 (1983).

[6] J. Uffink and J. Hilgevoord, Found. Phys. **15**, 925 (1985); J. Hilgevoord and J. Uffink in *Microphysical Reality and Quantum Formalism*, A. van der Merwe et al. (eds) (Kluwer, Dordrecht, 1988) and in *Sixty-two years of Uncertainty*, A. Miller (ed), (Plenum, N.Y., 1990).

[7] J. Anandan and Y. Aharonov, Phys. Rev. Lett. **65**, 1697 (1991); L. Vaidman, Am. J. Phys. **60**, 182 (1992); A. Uhlmann, Phys. Lett. A **161**, 329 (1992).

[8] J. Uffink, Am. J. Phys. (to appear).

[9] J-M. Lévy-Leblond, *Phys. Lett. A* **111**, 353 (1985).

[10] D. Deutsch, *Phys. Rev. Lett.* **50**, 631 (1983); I.Białynicki-Birula, in*Quantum Probability and Applications II*, L. Accardi and W. von Waldenfels (eds.), (Springer, Berlin, 1984); K. Kraus, *Phys. Rev. D* **35**, 3070 (1987).

[11] H. Maassen and J. Uffink, *Phys. Rev. Lett.* **60** 1103 (1988).

[12] G.H. Hardy, J.E. Littlewood and G. Pólya, *Inequalities*, (Cambridge U.P., Cambridge, 1934); A.Renyi, *Wahrscheinlichkeitsrechnung*, (Deutscher Verlag der Wissenschaften, 1962).

[13] J. Uffink, *Measures of Uncertainty and the Uncertainty Principle*, Ph. D. thesis, University of Utrecht (1990).

[14] J-M. Lévy-Leblond, Ann. Phys. **101**, 319 (1976).

[15] E.B. Davies, *Quantum Theory of open Systems* (Academic Press, London, 1976); A.S. Holevo *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).

[16] L. Susskind and J. Glogower, Physics **1**, 49 (1964).

[17] I.I. Hirschman, Am. J. Math. **79**, 152 (1957); I. Białynicki-Birula and J. Micielski, Comm. Math. Phys. **44**, 129 (1975); W. Beckner, Ann. Math. **102**, 159 (1975).