SLAC-PUB-3100
April 1983
(T/E)

ENTROPIC FORMULATION OF UNCERTAINTY FOR QUANTUM MEASUREMENTS*

M. HOSSEIN PARTOVI
Stanford Linear Accelerator Center
Stanford University, Stanford, California 94305

ABSTRACT

A quantitative formulation of the uncertainty principle, based on Deutsch's idea of using entropy as a measure of uncertainty to remove the inadequacies of the standard treatment, is presented. This uncertainty in general involves the resolution of the measuring device.

Submitted to Physical Review Letters

---

*Work supported by the Department of Energy, contract DE-AC03-76SF00515.

In a recent letter,[1] Deutsch presented a compelling criticism of the standard *quantitative* formulation of the uncertainty principle (usually expressed in terms of variances and based on the Heisenberg inequality and its generalizations[2]). He then proposed the natural and appealing alternative of using entropy as a measure of uncertainty, and presented arguments to demonstrate the viability of this entropic definition of uncertainty. However, his formulation, while essentially correct, is incomplete, and the required modification turns out to be of a rather fundamental nature, as will be seen below. The purpose of this note is to present this modified structure and to demonstrate that the emerging formulation is a complete and satisfactory expression of the uncertainty principle according to the criteria of Ref. 1.

Given a pair of observables represented by (self-adjoint) operators $\hat{A}$ and $\hat{B}$ with *discrete* spectra, Ref. 1 defines the uncertainty in the simultaneous measurement of the pair in the state $|\psi\rangle$ to be

$$U(\hat{A},\hat{B}\,|\psi) = S^A(\psi) + S^B(\psi) \ ,$$

where the entropy is defined by

$$S^A(\psi) = -\sum_a |\langle a|\psi\rangle|^2 \ell n |\langle a|\psi\rangle|^2 \ , \tag{1}$$

and where $\{|a\rangle\}$ is the set of eigenstates of $\hat{A}$ (see Ref. 1 for details). It is then demonstrated that U is never less than $-2\ \ell n[\frac{1}{2}(1 + sup\ |\langle a|b\rangle|)]$, thus showing that U is a satisfactory measure of uncertainty. Although it is further noted that a generalization of the above definition to the continuous case is inappropriate, this is dismissed as a technicality on the strength of the statement that actual measurements always involve a countable set of outcomes. While this assertion is correct, it does not reduce the continuous case to the discrete one, and it therefore leaves one with a severely limited definition (e.g., the cases of position-momentum and angle-angular momentum are excluded). In fact even the purely discrete case is not dealt with in a completely satisfactory manner. For example, in the case of a discrete spectrum with a limit point (such as the bound sates of the Coulomb potential accumulating near the ionization limit), the fact that any measuring device has finite resolution and cannot resolve the entire set of eigenvalues is not accounted for by the above definition of entropy. The resolution of the above difficulties lies, not unexpectedly, in a formulation more intimately anchored in the details of the measurement process. To wit, the relevant characteristics of the measuring device must be included in a correct formulation. This inclusion, as the sequel will show, has non-trivial consequences.

A measuring device D^A, used to measure the observable $\hat{A}$, in general corresponds to a partitioning of the spectrum of $\hat{A}$ into a collection of subsets α_i, and the assignment to a state $|\psi\rangle$ of a corresponding set of probabilities $P_i^A(\psi|D^A)$. These numbers (to be abbreviated P_i^A) express the probability of finding the outcome of the measurement to have a value in the subset α_i. In symbols,

$$P_i^A = \langle\psi|\ \hat{\pi}_i^A\ |\psi\rangle/\langle\psi|\psi\rangle\ , \tag{2}$$

where $\hat{\pi}_i^A$ is the projection onto the subspace spanned by the states corresponding to the points of α_i.[3] Note that while the whole spectrum is characterized by $\hat{A}$, the manner of its partitioning is a property of the measuring device. Often the α_i are a collection of intervals, which we may descriptively refer to as "bins". The existence of the $\hat{\pi}_i^A$ is assured by the spectral theorem, valid for any self-adjoint operator[3], and the completeness property is given by the operator statement $\Sigma_i\, \hat{\pi}_i^A = 1$.

The entropy associated with the measurement by means of the device D^A is now defined as

$$S(\psi|D^A) = -\sum_i P_i^A(\psi|D^A)\ell n\, P_i^A(\psi|D^A) \;, \tag{3}$$

which is an inherently non-negative quantity. In the rather special case that each α_i includes one point of the spectrum (degeneracies being ignored for simplicity), namely an eigenvalue a_i, one has $\hat{\pi}_i^A = |a_i\rangle\langle a_i|$, and the above definition reduces to that of Ref. 1. In realistic situations, however, at least some of the α_i will be infinite-dimensional (e.g., continuous spectra, discrete spectra with limit points), which is the essential reason necessitating the present formulation.

It will now be shown that the uncertainty in the measurement of two observables $\hat{A}$ and $\hat{B}$,

$$U(D^A, D^B|\psi) = S(\psi|D^A) + S(\psi|D^B) \;, \tag{4}$$

possesses a lower bound which in general depends on the measuring devices, but not on $|\psi\rangle$. Since the P_i^A are probabilities, one can write (following Ref. 1)

$$U(D^A, D^B|\psi) = -\sum_{i,j} P_i^A P_j^B \;\ell n(P_i^A P_j^B) \;,$$

and proceed to find the infimum of $-\ell n(P_i^A P_j^B)$ for a given pair (i,j) and all $|\psi\rangle$. To that end, let $||\hat{\pi}||$ denote the norm of $\hat{\pi}$. Then for any $|\psi\rangle$ with $\langle\psi|\psi\rangle = 1$,

$$\begin{aligned} ||\hat{\pi}_i^A + \hat{\pi}_j^B||^2 &\geq \langle\psi|(\hat{\pi}_i^A + \hat{\pi}_j^B)^2|\psi\rangle \geq \langle\psi|(\hat{\pi}_i^A + \hat{\pi}_j^B)|\psi\rangle^2 \\ &= \langle\psi|\,\hat{\pi}_i^A\,|\psi\rangle^2 + \langle\psi|\,\hat{\pi}_j^B\,|\psi\rangle^2 + 2\langle\psi|\,\hat{\pi}_i^A\,|\psi\rangle\langle\psi|\,\hat{\pi}_j^B\,|\psi\rangle \;, \end{aligned}$$

implying that $P_i^A P_j^B \leq \frac{1}{4}||\hat{\pi}_i^A + \hat{\pi}_j^B||^2$. Hence

$$U(D^A, D^B|\psi) \geq -\sum_{i,j} p_i^A p_j^B \ell n\,(\|\hat{\pi}_i^A + \hat{\pi}_j^B\|^2/4)$$
$$\geq 2\,\ell n\,(2/\sup_{i,j}\|\hat{\pi}_i^A + \hat{\pi}_j^B\|)\ . \tag{5}$$

Since the $\hat{\pi}_i$ are projections, it follows that $1 \leq \|\hat{\pi}_i^A + \hat{\pi}_j^B\| \leq 2$, where the upper bound is attained if and only if $\hat{\pi}_i^A$ and $\hat{\pi}_j^B$ have (at least) one eigenvector in common.[4] Clearly Eq. (5) is a satisfactory quantitative expression of the uncertainty principle according to the criteria of Ref. 1.[5]

The physical consequences of the present formulation will now be illustrated by means of two examples. The first is the celebrated angle-angular momentum case,[2] $\hat{A} = \hat{\varphi}$, $\hat{B} = \hat{L}_z$. However, our formulation requires further detailes of the measuring devices, namely their bin assignments. As a simple arrangement, we shall assume that the device measuring φ is organized in bins of size $\Delta\varphi$ ($2\pi/\Delta\varphi =$ some integer), and that the device measuring L_z can resolve down to a single value of the angular momentum. For a given angular bin $(\varphi_i, \varphi_{i+1})$ and a value $\hat{L}_z = m$, the corresponding projections are defined by (Θ is the step function)

$$\hat{\pi}_i^{\varphi}\,\psi(\varphi) = \Theta(\varphi - \varphi_i)\Theta(\varphi_{i+1} - \varphi)\psi(\varphi)\ ,$$

$$\hat{\pi}_m^{L_z}\,\psi(\varphi) = (2\pi)^{-1}\int_0^{2\pi} d\varphi'\ \exp[im(\varphi - \varphi')]\psi(\varphi')\ ,$$

where $\psi(\varphi)$ is the φ representation of $|\psi\rangle$ (note a slight change of notation, $j \to m$). A simple calculation shows that the extremal values of $p_i^{\varphi} p_m^{L_z}$ are attained by the eigensolutions of the (bounded, self-adjoint) operator $\hat{\pi}_i^{\varphi} + \hat{\pi}_m^{L_z}$,

$$(\hat{\pi}_i^{\varphi} + \hat{\pi}_m^{L_z})|v\rangle = \lambda|v\rangle\ ,$$

so that $\max(P_i^{\varphi} P_m^{L_z}) = \frac{1}{4}\lambda_{max}^2$, in accordance with the general inequality derived before. The relevant solution is found to be $|v_{max}\rangle = [1 + (\Delta\varphi/2\pi)^{-1/2}\,\hat{\pi}_i^{\varphi}]|m\rangle$ with $\lambda_{max} = 1 + (\Delta\varphi/2\pi)^{1/2}$, where $\hat{L}_z\,|m\rangle = m|m\rangle$. Therefore

$$U(D^{\varphi}, D^{L_z}|\psi) \geq 2\,\ell n\,\frac{2}{1 + (\Delta\varphi/2\pi)^{1/2}}\,. \tag{6}$$

Note that $\Delta\varphi$ is not a ***variance***, but effectively the ***resolution*** of the measuring device. The criteria of Ref. 1 are fulfilled by (6); see remark (c) below.

Finally, we consider the archtypal example, the case of position and momentum (restricted to one dimension for simplicity). Here, we shall assume that position and momentum are measured in bins of size Δx and Δp, respectively. The projection $\hat{\pi}_i^x$ is then an exact analogue of $\hat{\pi}_i^{\varphi}$ considered above, and

$$\hat{\pi}_j^p\,\psi(x) = (2\pi)^{-1}\int_{p_i}^{p_{i+1}} dp \int_{-\infty}^{+\infty} dx'\exp\left[ip(x - x')\right]\psi(x')\,.$$

Just as in the previous example, $\max(P_i^x P_j^p)$ is obtained from the solution of an eigenvalue equation for the (bounded, self-adjoint) operator $\hat{\pi}_i^x + \hat{\pi}_j^p$. After some computation, one transforms the above equation to the form

$$\int_0^1 d\xi'\,w(\xi')(\xi - \xi')^{-1}\sin\,k\pi(\xi - \xi') = \pi\mu^2 w(\xi)\ \ (0 \leq \xi \leq 1)\,,$$

with $\max(P_i^x P_j^p) = \frac{1}{4}(1+\mu_{max})^2$, where $\pi\mu_{max}^2$ is the largest (positive) eigenvalue of the above integral equation (guaranteed to exist by the compactness of the kernel[4]), and $k = (\Delta x)(\Delta p)/(2\pi)$. It suffices for our purposes to find the behavior of μ_{max} for the limiting values of k. One finds from the integral equation that $\mu_{max}^2 \to k(1)$ for $k \to 0(\infty)$. Assembling the above information, we finally arrive at

$$U(D^x,D^p|\psi) \geq 2\,\ell n\;\frac{2}{1+\mu_{max}}\;,$$
$$\mu_{max} \to [(\Delta x)(\Delta p)/2\pi]^{1/2}\;,(\Delta x)(\Delta p)/2\pi \ll 1\;, \tag{7}$$
$$\mu_{max} \to 1\;,(\Delta x)(\Delta p)/2\pi \gg 1\;,$$

with μ_{max} restricted to $[0,1]$. It is worth recalling that the results given in Eqs. (6) and (7) crucially depend on the introduction of the resolutions of the measuring devices into the basic definition of uncertainty.

We conclude with a few remarks. (a) The contrast between the standard and the entropic measures of uncertainty is highlighted in the example of $\hat{L}_z$ and $|\psi\rangle = |-m\rangle + |+m\rangle$. While the variance of $\hat{L}_z$ in this state is $|m|$ and unbounded, the entropy is independent of m and equal to $\ell n\,2$. (b) The lower bounds given in Eqs. (6) and (7) are not necessarily optimal (i.e., they are not the infima), even though those corresponding to $-\ell n\; p_i^A p_j^B$ are. (c) In both examples considered above, the minimum of the lower bound given for U is attained for the case of minimum resolution, i.e., $(\Delta\varphi)/2\pi = 1$ and $(\Delta x)(\Delta p)/2\pi = \infty$, respectively. This is exactly as it should be, since, given these resolutions, one is already in possession of the maximum accessible information about φ, respectively x and/or p, without the need to perform the corresponding measurement. (d) $0 \leq S(D^A|\psi) \leq \ell n(No.\ of\ bins\ of\ D^A) \leq \ell n(rank\ of\ \hat{A})$. For a given $|\psi\rangle$, one can (in principle) design measuring devices such that either bound is approached as closely as desired. For a given device (e.g., fixed resolution), on the other hand, the bounds may not necessarily be attainable, as can be seen in the examples.

I wish to thank Professor R. Blankenbecler for a reading of the manuscript and Professor S. Drell for hospitality and support at the Stanford Linear Accelerator Center. This work was supported by the Department of Energy under contract No. DE-AC03-76SF00515.

REFERENCES

1. D. Deutsch, Phys. Rev. Lett. 50, 631 (1983).
2. See Ref. 1 for original papers. A further aspect of the difficulties of the standard formulation, and one which is associated with restricted domains of definition, was discussed extensively in an earlier period; see P. Carruthers and M. Nieto, Rev. Mod. Phys. 40, 411 (1968).
3. Generally, α_i can be any Borel subset of the real line; see J. M. Jauch, "Foundations of Quantum Mechanics" (Addison-Wesley Publishing Company, Reading, Massachusetts, 1968) for this and other mathematical aspects.
4. In *general*, there is no guarantee for the existence of a state for which the norm $||\hat{\pi}_i^A + \hat{\pi}_j^B||$ is reached (a sufficient condition is the compactness of the operator). However, for the examples treated in the text, this is the case.
5. For the special case of Ref. 1, i.e., $\hat{\pi}_i^A = |a_i\rangle\langle a_i|$, $\hat{\pi}_j^B = |bj\rangle\langle bj|$, one calculates $||\hat{\pi}_i^A + \hat{\pi}_j^B|| = 1 + |\langle a_i|b_j\rangle|$, whence follows Eq. (13) therein.