PHYSICAL REVIEW LETTERS

VOLUME 60 | 21 MARCH 1988 | NUMBER 12

Generalized Entropic Uncertainty Relations

Hans Maassen
Institute for Theoretical Physics, University of Nijmegen, 6525 ED Nijmegen, The Netherlands

and

J. B. M. Uffink[(a)]
Institute for Theoretical Physics, University of Amsterdam, 1018 XE Amsterdam, The Netherlands
(Received 18 November 1987)

A new class of uncertainty relations is derived for pairs of observables in a finite-dimensional Hilbert space which do not have any common eigenvector. This class contains an "entropic" uncertainty relation which improves a previous result of Deutsch and confirms a recent conjecture by Kraus. Some comments are made on the extension of these relations to the case where the Hilbert space is infinite dimensional.

PACS numbers 03.65.Bz

Let A and B denote two Hermitean operators representing physical observables in an N-dimensional Hilbert space, and let $\{|a_j\rangle\}$ and $\{|b_j\rangle\}$ with $j=1,\ldots,N$ be the corresponding complete sets of normalized eigenvectors. The uncertainty principle states, broadly speaking, that for any quantum state ψ the two probability distributions $p=(p_1,\ldots,p_N)$ and $q=(q_1,\ldots,q_N)$, defined by

$$p_j=|\langle a_j|\psi\rangle|^2, \quad q_j=|\langle b_j|\psi\rangle|^2, \tag{1}$$

cannot both be arbitrarily peaked, provided that A and B are sufficiently noncommuting. In most text books this principle is expressed by the Robertson relation[1]:

$$\Delta_\psi A\,\Delta_\psi B \geqq \tfrac{1}{2}\,|\langle[A,B]\rangle_\psi|, \tag{2}$$

where $\Delta_\psi A$ and $\Delta_\psi B$ denote the standard deviations of the distributions (1):

$$(\Delta_\psi A)^2=\langle A^2\rangle_\psi-(\langle A\rangle_\psi)^2,$$

$$(\Delta_\psi B)^2=\langle B^2\rangle_\psi-(\langle B\rangle_\psi)^2.$$

This formulation of the uncertainty principle has recently been criticized[2,3] on the grounds that the right-hand side of (2) is not a fixed lower bound, but depends on the state ψ. For example, when ψ is an eigenstate of A, one has $\Delta_\psi A=0$, and $\langle[A,B]\rangle_\psi=0$, so that no restriction on $\Delta_\psi B$ is imposed by relation (2).

To improve on this situation, so-called "entropic" uncertainty relations have been proposed[3,4] which rely on the Shannon entropy H as a measure of uncertainty. For a general probability distribution $P=(P_1,\ldots,P_N)$, $P_i\geqq 0$, $\sum_i P_i=1$, on a set of N possible outcomes, the Shannon entropy is defined as

$$H(P)=-\textstyle\sum_j P_j\ln P_j. \tag{3}$$

Applying this notion to the probability distributions p and q introduced in (1), Deutsch[3] has shown the relation

$$H(p)+H(q)\geqq -2\ln\tfrac{1}{2}(1+c), \tag{4}$$

where

$$c=\max_{j,k}|\langle a_j|b_k\rangle|. \tag{5}$$

More recently, Kraus[5] has conjectured that this relation may be improved to

$$H(p)+H(q)\geqq -2\ln c. \tag{6}$$

The advantage of these relations over (2) is that they have a right-hand side which is independent of the state ψ. Thus, they yield nontrivial information on the probability distributions p and q as long as $c<1$, that is when

A and B do not share any common eigenvector.

In this paper, we shall prove relation (6). Furthermore, we will show that relation (6) is just one member of a general class of inequalities, all of which may be said to express the uncertainty principle in the sense that they put bounds on the extent to which the distributions p and q can be simultaneously peaked. Finally we shall make some remarks on the extension of these relations to the case where the Hilbert space is not finite dimensional.

Let P denote an arbitrary probability distribution over a set of N possible outcomes, and consider the expression

$$M_r(P)=\left[\sum_j (P_j)^{1+r}\right]^{1/r} \quad (-1<r<0 \text{ or } 0<r). \tag{7}$$

Expressions of this type were studied in detail by Hardy, Littlewood, and Polya.[6] We list a few of their properties.

(a) $M_r(P)$ is invariant under a relabeling of the set of possible outcomes.

(b) $M_r(P)$ is convex in P. That is, when $P^{(1)}$ and $P^{(2)}$ are two probability distributions over the same set, and P is defined by $P_j=aP_j^{(1)}+(1-a)P_j^{(2)}$, for $0\leqq a\leqq 1$, we have

$$M_r(P)\leqq aM_r(P^{(1)})+(1-a)M_r(P^{(2)}).$$

(c) When $P=(P_1,\ldots,P_N)$ and $Q=(Q_1,\ldots,Q_M)$ are two probability distributions and R is their independent product, $R_{ij}=P_iQ_j$, one has

$$M_r(R)=M_r(P)M_r(Q).$$

(d) $M_r(P)$ is a continuous nondecreasing function of r for $-1\leqq r\leqq\infty$ provided one defines the following limiting values:

(e) $M_0(P)=\exp[-H(P)]$.

(f) $M_{-1}(P)=1/N'$, where N' denotes the number of possible outcomes with a probability $P_j>0$.

(g) $M_\infty(P)=\mathrm{Max}_j P_j$.

In view of these properties we may regard the expressions $M_r(P)$ for $-1\leqq r\leqq\infty$ as measures of the "average peakedness" of P. Indeed, considered as a function of P, $M_r(P)$ attains its maximal value (unity) only if $P_j=\delta_{jj_0}$, where δ denotes the Kronecker δ, and for $r>-1$, its minimal value $1/N$ is reached only when the probability is uniformly distributed over all possible outcomes, i.e., when $P_j=1/N$ for all j. Alternatively, the class of expressions $-\ln M_r(P)$ may be seen as a measure of the amount of uncertainty associated with P, providing a natural generalization of the Shannon entropy.

Let us now apply these notions to the problem of expressing the uncertainty principle. Consider the product $M_r(p)M_s(q)$, where p and q are defined by (1). From the uncertainty principle one expects that when A and B do not share any eigenstate, this product cannot be arbitrarily close to unity. This intuition proves to be correct. In fact we may quote a result of Landau and Pollak[7] which in the present notation reads

$$\arccos c\leqq \arccos M_\infty(p)+\arccos M_\infty(q). \tag{8}$$

Maximizing the product $M_\infty(p)M_\infty(q)$ under this condition, one obtains

$$M_\infty(p)M_\infty(q)\leqq \tfrac{1}{4}(1+c)^2, \tag{9}$$

a result that, by virtue of the properties (d) and (e) above, already implies the Deutsch relation (4).

However, one may obtain a better bound on the product $M_r(p)M_s(q)$ for certain combinations of finite r and s. The key to this is the following theorem due to Riesz.

Riesz Theorem: Let $x=(x_1,\ldots,x_N)$ denote a sequence of complex numbers and T_{jk} a linear transformation matrix, $(Tx)_j=\sum_k T_{jk}x_k$, which obeys $\sum_j |(Tx)_j|^2=\sum_k |x_k|^2$ for all x, and let $c=\mathrm{Max}_{j,k}|T_{jk}|$. Then

$$c^{1/a'}\left[\sum_j |(Tx)_j|^{a'}\right]^{1/a'} \leqq c^{1/a}\left(\sum_k |x_k|^a\right)^{1/a}, \tag{10}$$

for $1\leqq a\leqq 2$, $1/a+1/a'=1$. An elementary proof of this relation may be found in Riesz.[8] For a more general version of the theorem, see Reed and Simon.[9] The conditions assumed in this theorem are clearly fulfilled for $x_k=\langle a_k|\psi\rangle$, $T_{jk}=\langle b_j|a_k\rangle$, and $(Tx)_j=\langle b_j|\psi\rangle$. Then, putting $a=2(1+r)$, $a'=2(1+s)$, we may rewrite (10) as

$$M_r(p)M_s(q)\leqq c^2 \tag{11}$$

for

$$s\geq 0,\quad r=-s/(2s+1). \tag{12}$$

Of course, the role of the operators A and B may be interchanged in this derivation. This does not alter the value of c, so that under the conditions (12) we also have

$$M_r(q)M_s(p)\leqq c^2. \tag{13}$$

The relations (11) and (13) form the general class of inequalities alluded to above. Taking $r=s=0$ in either of these relations, we arrive at inequality (6).

A natural question at this stage is whether the general set of inequalities (11) and (13) is more informative than the special case (6). Consider a simple example. Let ψ be the two-dimensional state of a spin-$\frac{1}{2}$ particle, and A and B be spin components in orthogonal directions. In this case, we have $c=2^{-1/2}$. Suppose that the probabilities of the two possible outcomes of the A measurement are given, say $p_1=\frac{3}{4}$, $p_2=\frac{1}{4}$. What restriction does this put on the probabilities q_1, $1-q_1$ for the two outcomes of the B measurement? Using relation

(6), we obtain

$$H(q) \geqq 0.1309,$$

corresponding to a value of q_1 between 0.03 and 0.97. Using the general set (11) and adopting the optimal choice for r and s, $r=-\frac{1}{2}$, $s=\infty$, one finds the stronger result $0.066 \leqq q_1 \leqq 0.933$, which almost doubles the lower bound on the entropy:

$$H(q) \geqq 0.2458.$$

In contrast to this, we note that no restriction on q_1 can be derived from the uncertainty relations (2) or (4).

Further, one may ask whether the inequalities (11) or (13) remain valid for mixed states. For a general mixed state,

$$W=\sum_n \alpha_n |\psi^{(n)}\rangle\langle\psi^{(n)}|, \quad \alpha_n \geqq 0, \quad \sum_n \alpha_n = 1,$$

we may write the probability distributions corresponding to the observables A and B as

$$\bar{p}_j = \sum_n \alpha_n p_j^{(n)}, \quad \bar{q}_j = \sum_n \alpha_n q_j^{(n)},$$

and the question is then whether the relation

$$M_r(\bar{p}) M_s(\bar{q}) \leqq c^2 \tag{14}$$

holds. For the case $r=s=0$ this inequality follows immediately from the fact that the Shannon entropy $H(p) = -\ln M_0(p)$ is a concave function of p. However, for arbitrary values of r, $-\ln M_r$ is not always concave, and the validity of (14) is not trivial. Nevertheless, for the choice (12) one obtains

$$\left\{\sum_j \left[\sum_n \alpha_n p_j^{(n)}\right]^{1+r}\right\}^{1/(1+r)} \geqq \sum_n \alpha_n \left\{\sum_j (p_j^{(n)})^{1+r}\right\}^{1/(1+r)} \geqq c^{2r/(r+1)} \sum_n \alpha_n \left\{\sum_j (q_j^{(n)})^{1+s}\right\}^{1/(1+s)}$$
$$\geqq c^{2r/(r+1)} \left\{\sum_j \left[\sum_n \alpha_n q_j^{(n)}\right]^{1+s}\right\}^{1/(1+s)}. \tag{15}$$

In this chain, the inequality in the middle is derived from the fact that for all values of n the relation (11) holds for the pair $p^{(n)}$ and $q^{(n)}$, pertaining to the pure states $\psi^{(n)}$; the first and last inequality signs in (15) are obtained from an application of the Minkowski inequality.[10] Comparing the first and last expressions in (15) we arrive at (14).

A next question is, whether the above results can be extended to the case where the Hilbert space is not of finite dimension. In fact the properties (a)–(g) remain valid when $N=\infty$, with one exception: for $r \leqq 0$, M_r need no longer be continuous, but is only continuous from the right.[11] Furthermore, the Riesz theorem remains valid, assuming the convergence of the expressions in (10). Somewhat more difficult is the case where A and B possess an unbounded continuous spectrum. We will not go into the mathematical details, but as an illustration only mention the corresponding results for position and momentum in one spatial dimension. The analog of definition (7) for continuous variables is

$$M_r(P) = \left[\int \{P(x)\}^{1+r} dx\right]^{1/r}, \tag{16}$$

for a probability density $P(x)$. Unlike its counterpart (7) in the discrete case, this expression is not bounded from above for any value of r. Accordingly, the entropy

$$H(P) = -\ln M_0(P) = -\int P(x) \ln P(x)\, dx \tag{17}$$

may attain negative values.

Now let $\psi(x)$ be a normalized position wave function, and

$$\phi(k) = \int \langle k|x\rangle \psi(x)\, dx \tag{18}$$

be the associated wave function in momentum representation where $\langle k|x\rangle = (2\pi)^{-1/2} e^{ikx}$. The analog of the Riesz theorem in this case is the Hausdorff-Young inequality[12]:

$$c^{1/a'} \left[\int |\phi(k)|^{a'} dk\right]^{1/a'} \leqq c^{1/a} \left[\int |\psi(x)|^{a} dx\right]^{1/a} \tag{19}$$

for $c=(2\pi)^{-1/2}$ and $1 \leqq a \leqq 2$, $1/a + 1/a' = 1$. Again, putting $a=2(1+r), a'=2(1+s)$, we obtain

$$M_r(|\psi|^2) M_s(|\phi|^2) \leqq (2\pi)^{-1}. \tag{20}$$

For $r=s=0$ this reduces to the entropic relation for position and momentum,

$$H(|\psi|^2) + H(|\phi|^2) \geqq \ln 2\pi, \tag{21}$$

derived by Hirschman[13] and later improved by Beckner.[14] In a similar vein, one may recover entropic relations for angle and angular momentum obtained by Bialynicki-Birula and Mycielski.[4] Thus the set of inequalities (11) and (13) share a feature with the Robertson uncertainty relation (2), namely that they possess a natural extension to canonical variables, although for position and momentum they do not yield the best possible bound.

However, before such extensions can be regarded as satisfactory expressions of an uncertainty principle, one must check that their interpretation is preserved. Unfortunately, this is not the case for all of the expressions M_r. Obviously, a desirable property of any measure $U(P)$ for the amount of uncertainty associated with a probability distribution P is that $U(P)$ should approach its minimal value whenever the distribution P approaches a δ distribution, where the total probability is concentrated on a single point. The reason for this continuity requirement is the following. In writing an uncertainty relation which puts a bound on expressions of the form $U(P)$, we want to express that two probability distributions cannot simultaneously "resemble" a δ distribution. Clearly, when the above continuity requirement is not fulfilled, we have not succeeded in this goal, since then it may still be possible that one or both of the distributions are sharply concentrated, without violating the bound on their uncertainties.

Let us try to make this idea more concrete. We confine the discussion to probability distributions on an infinite discrete set, $P=(P_1,P_2,\ldots,)$ with $\sum_j P_j=1$. A series of such distributions $P^{(n)}$ may be said to converge to the distribution P when

$$\sum_j |P_j^{(n)}-P_j| \to 0 \text{ for } n\to\infty. \qquad (22)$$

With this sense of convergence, it can be easily shown that the Shannon entropy does not fulfill the continuity requirement mentioned above, nor does any of the expressions $-\ln M_r(P)$ for $r\leqq 0$. A simple counterexample will suffice. Take $P^{(\epsilon,M)}$ as

$$P_1^{(\epsilon,M)}=1-\epsilon,$$
$$P_j^{(\epsilon,M)}=\epsilon/M, \text{ for } j=2,\ldots,M+1, \qquad (23)$$
$$P_j^{(\epsilon,M)}=0, \text{ otherwise.}$$

Obviously, $P^{(\epsilon,M)}$ approaches the distribution $(1,0,0,0,\ldots)$ when $\epsilon\to 0$, $M\to\infty$, irrespective of the route along which this simultaneous limit is taken. However, the value of $H(P^{(\epsilon,M)})$ can be made to approach any value (including infinity) depending on the chosen route. Similar remarks may be made about $-\ln M_r(P^{(\epsilon,M)})$ for $r\leqq 0$.

On the other hand, for $r>0$, $-\ln M_r(P)$ does satisfy the continuity requirement. To show this, observe that the function $f(x)=x^{1+r}$ has a differential quotient $[f(x)-f(y)]/(x-y)$, which for $r>0$, $0\leqq x,y\leqq 1$, never exceeds $(df/dx)_{x=1}=1+r$. It follows that

$$|(P_j^{(n)})^{1+r}-(P_j)^{1+r}|\leqq(1+r)|P_j^{(n)}-P_j|. \qquad (24)$$

Summing over j and assuming the convergence (22), we may establish the convergence of $M_r(P^{(n)})$ to $M_r(P)$.

Thus, the general set of inequalities (11) and (13) possess one more advantage over the entropic uncertainty relation (6) in the case where the observables have an unbounded discrete spectrum, namely that they connect two measures of uncertainty of which at least one satisfies the continuity requirement.

One of the authors (H.M.) acknowledges support from the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

[a]Present address: Institute for History and Foundations of Science, Physical Laboratory, P. O. Box 80.000, 3508 TA Utrecht, The Netherlands.

[1]H. P. Robertson, Phys. Rev. **34**, 163 (1929).

[2]J. Hilgevoord and J. B. M. Uffink, in *Proceedings of the International Conference on Microphysical Reality and Quantum Description, Urbino, Italy, 1985,* edited by A. van der Merwe, F. Selleri, and G. Tarozzi (Reidel, Dordrecht, 1988).

[3]D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).

[4]I. Bialynicki-Birula and J. Mycielski, Commun. Math. Phys. **44**, 129 (1975).

[5]K. Kraus, Phys. Rev. D **35**, 3070 (1987).

[6]G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities* (Cambridge Univ. Press, Cambridge, 1952), 2nd ed, Ch. 2.

[7]H. J. Landau and H. O. Pollak, Bell Syst. Tech. J. **40**, 65 (1961).

[8]M. Riesz, Acta Math. **49**, 465 (1926).

[9]M. Reed and B. Simon, in *Fourier Analysis, Self-Adjointness,* Methods of Modern Mathematical Physics Vol. 2 (Academic, New York, 1975), p. 27.

[10]Hardy, Littlewood, and Polya, Ref. 6, p. 31.

[11]Hardy, Littlewood, and Polya, Ref. 6, p. 122.

[12]Reed and Simon, Ref. 9, p. 11.

[13]I. I. Hirschman, Am. J. Math. **79**, 152 (1957).

[14]W. Beckner, Ann. Math. **102**, 159 (1975).