Pramāṇa, Vol. 24, No. 5, May 1985, pp. 673–684. © Printed in India.

Entropic formulation of uncertainty relations for successive measurements

M D SRINIVAS
Department of Theoretical Physics, University of Madras, Guindy Campus, Madras 600 025, India

MS received 21 November 1984

Abstract. An entropic formulation of uncertainty relations is obtained for the case of successive measurements. The lower bound on the overall uncertainty, that is obtained for the case of successive measurements, is shown to be larger than the recently derived Deutsch-Partovi lower bound on the overall uncertainty in the case of distinct measurements.

Keywords. Uncertainty; variance; information theoretic entropy; successive measurements; entropic formulation.

PACS No. 03·65

1. Entropic formulation of uncertainty relations

Recent investigations have focussed attention on some of the serious inadequacies of the usual (text book) formulations of the uncertainty relations inequality (Heisenberg 1927) and its generalisations (Robertson 1929; Schrödinger 1930)) which employ the variance of an observable as a measure of 'uncertainty'. As is well known, the variance $(\Delta^\rho A)^2$ of an observable A in state ρ is given by

$$(\Delta^\rho A)^2 = \langle A^2 \rangle_\rho - \langle A \rangle_\rho^2 = \mathrm{Tr}(\rho A^2) - (\mathrm{Tr}\rho A)^2, \quad (1)$$

and is to some extent a measure of the 'spread' in the probability distribution of A in state ρ. For instance, $(\Delta^\rho A)^2$ is non-negative and vanishes only if ρ is a mixture of eigenstates of A, all associated with the same eigenvalue. However, the standard formulation of uncertainty relations in terms of variances, *viz*,

$$(\Delta^\rho A)^2 (\Delta^\rho B)^2 \geq \tfrac{1}{4} |\langle [A, B] \rangle|^2 \quad (2)$$

does not always provide us with an absolute lower bound on the uncertainty in one variable (say B) given the uncertainty in the other (A). This is because the right side of (2) depends in general on the state ρ of the system. Only when the commutator $[A, B]$ becomes a constant multiple of the identity operator (as it happens in the case of canonically conjugate observables) does the right side of (2) become independent of ρ; and it is in such cases alone that we have a lower bound on the uncertainty in B given the uncertainty in A. This basic deficiency of (2) cannot be overcome by taking the infimum of the right side over all the states ρ; as this infimum invariably vanishes even when one of the observables A, B has just one discrete eigenvalue, or equivalently a normalisable eigenvector. Thus, except for the case of canonically conjugate observables, the standard formulation of uncertainty relation seems to be totally ineffective in providing an estimate of the uncertainty in one observable, given the uncertainty in the other.

Recently Deutsch (1983) highlighted the above and other inadequacies of the standard formulation (2) of the uncertainty relation. He has further argued that for an observable with a purely discrete spectrum the variance is not an appropriate measure of the uncertainty or the 'spread' in its probability distribution. If A is an observable with a purely discrete spectrum and has the following spectral resolution

$$A = \sum_i a_i P^A(a_i), \tag{3}$$

where $\{a_i\}$ are the eigenvalues and $\{P^A(a_i)\}$ the associated eigenprojectors, then the variance given by

$$(\Delta^\rho A)^2 = \sum_i a_i^2 \operatorname{Tr}(\rho P^A(a_i)) - \left\{ \sum_i a_i \operatorname{Tr}(\rho P^A(a_i)) \right\}^2, \tag{4}$$

depends also on the eigenvalues $\{a_i\}$ apart from the probability distribution of A in state ρ given by

$$\Pr{}_A^\rho(a_i) = \operatorname{Tr}[\rho P^A(a_i)]. \tag{5}$$

Deutsch therefore argued that the appropriate measure of the 'spread' in the probability distribution such as (5) is not the variance (4) (which depends also on irrelevant factors such as the eigenvalues) but the well-known information-theoretic entropy (Khinchin 1957) of the distribution (5), defined by

$$S^\rho(A) = -\sum_i \Pr{}_A^\rho(a_i) \log \Pr{}_A^\rho(a_i), \tag{6}$$

where it is always assumed that $0 \log 0 = 0$. As is well known $S^\rho(A)$ is non-negative and vanishes only if the probability distribution of A in state ρ reduces to the deterministic case

$$\Pr{}_A^\rho(a_i) = \delta_{ij} \tag{7}$$

for some j, which happens only if ρ is a mixture of eigenstates of A all associated with the same eigenvalue. In fact it is a basic result of information theory (Khinchin 1957) that whenever we have a discrete probability distribution, the entropy $S^\rho(A)$ given by (6) is an appropriate measure of the 'spread' in the probability distribution (5) of A in state ρ and thus of the uncertainty in the outcomes of an A-measurement performed on an ensemble of systems in state ρ. This (information–theoretic) entropy $S^\rho(A)$ of A in state ρ is sometimes referred to as the A-entropy (Ingarden 1976; Grabowski 1978a, b) perhaps to distinguish it from the more familiar Von Neumann or the thermodynamic entropy $S(\rho)$ of state ρ (Von Neumann 1955; Wehrl 1978; see also Lindblad 1973) given by

$$S(\rho) = -\operatorname{Tr}(\rho \log \rho). \tag{8}$$

As is well known $S(\rho)$ (as contrasted with $S^\rho(A)$) is a characteristic of the state ρ alone, and is in fact a measure of the extent to which ρ is 'mixed' or 'chaotic' (Wehrl 1974).

Deutsch has argued that instead of (2) a more appropriate formulation of the uncertainty relation should be sought in the form

$$S^\rho(A) + S^\rho(B) \geq u(A, B), \tag{9}$$

where $u(A, B)$ is a non-negative number *independent* of the state ρ. Clearly a relation of the form (9) would always provide us with a lower bound on the uncertainty (or now the entropy) of B given the uncertainty (or entropy) of A. Deutsch also stipulated that $u(A, B)$ should vanish only if the observables A, B are such that both the entropies $S^{\rho}(A)$ and $S^{\rho}(B)$ can be made arbitrarily small at the same time, which happens essentially only when A and B have a common eigenvector. Deutsch also succeeded in deriving such an uncertainty relation in the 'entropic form' (9) for the case when the observables A, B have a purely discrete and non-degenerate spectrum. This was extended by Partovi (1983) so as to include cases with degeneracies also.

In order to state the Deutsch-Partovi uncertainty relation, let us consider two observables A, B with purely discrete spectra, and the following spectral resolutions

$$A = \sum_i a_i P^A(a_i), \tag{10a}$$

$$B = \sum_j b_j P^B(b_j). \tag{10b}$$

The associated entropies are given by

$$S^{\rho}(A) = -\sum_i \mathrm{Tr}(\rho P^A(a_i)) \log \mathrm{Tr}(\rho P^A(a_i)), \tag{11a}$$

$$S^{\rho}(B) = -\sum_j \mathrm{Tr}(\rho P^B(b_j)) \log \mathrm{Tr}(\rho P^B(b_j)). \tag{11b}$$

Then we have the Deutsch-Partovi uncertainty relation

$$S^{\rho}(A) + S^{\rho}(B) \geqslant 2 \log \frac{2}{\sup\limits_{i,j} \| P^A(a_i) + P^B(b_j) \|} \tag{12}$$

where $\| \quad \|$ denotes the operator norm. Since,

$$\|P^A(a_i) + P^B(b_j)\| \leqslant 2$$

the right side of (12) is clearly non-negative and vanishes only if

$$\sup_{i,j} \|P^A(a_i) + P^B(b_j)\| = 2,$$

which happens essentially only when A, B have a common eigenvector. For the particular case when the spectra of A, B are totally non-degenerate, so that (10a, b) become

$$A = \sum_i a_i |a_i\rangle \langle a_i|, \tag{13a}$$

$$B = \sum_j b_j |b_j\rangle \langle b_j|, \tag{13b}$$

the uncertainty relation (12) reduces to the following original form derived by Deutsch

$$S^{\rho}(A) + S^{\rho}(B) \geqslant 2 \log \frac{2}{1 + \sup\limits_{i,j} |\langle a_i | b_j \rangle|}. \tag{14}$$

It is indeed curious to note that while the Deutsch-Partovi relation (12) which has been recently derived, is valid only for observables with purely discrete spectra, an entropic formulation of the position-momentum (and even angle-angular momentum) uncertainty relation is already available in literature—though it has not been generally taken note of and has thus been overlooked even by Deutsch and Partovi. For a particle in one-dimension characterised by the wave function $\psi(x)$, this entropic form of the position-momentum (or more correctly the position- wave number) uncertainty relation is usually written as

$$-\int_{-\infty}^{\infty} |\psi(x)|^2 \log|\psi(x)|^2 \mathrm{d}x - \int_{-\infty}^{\infty} |\tilde{\psi}(k)|^2 \log|\tilde{\psi}(k)|^2 \mathrm{d}k \geqslant 1 + \log\pi \tag{15}$$

where

$$\tilde{\psi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-ikx)\psi(x)\mathrm{d}x \tag{16}$$

so that $|\tilde{\psi}(k)|^2$ is the probability density that the particle has wave number k. It has recently been pointed out by Bialynicki-Birula (1984) that (15) was conjectured by Everett in his well known thesis of 1957 (Everett 1973) as also by Hirschman (1957) in the same year. It was proved in 1975 by Beckner (1975b) and independently by Bialynicki-Birula and Mycielski (1975) based on the estimation of the so-called (p, q)-norm of the Fourier transformation due to Beckner (1975a).

While (15) is mathematically valid, it is not physically meaningful because the quantities on the left side in (15) have both undefined physical dismensions. This relatively unsatisfactory feature is in fact a general feature common to all the usual definitions of entropy associated with a continuous probability distribution. If a continuous random quantity X has (physical) dimension D then the associated probability density $p(x)$ is not a pure number, but has dimension $1/D$. Thus the entropy of X which is usually defined in literature on information theory (McElice 1977) by

$$S(X) = -\int p(x)\log p(x)\mathrm{d}x \tag{17}$$

is not physically meaningful, as it has the inadmissible 'dimension' $\log D$. This unsatisfactory feature of the entropy of a continuous random variable persists even if one uses some of the other measures of entropy discussed in literature such as the one due to Renyi (1961). However we can easily arrive at a physically meaningful measure of the uncertainty of a continuous random variable by considering the exponential of the entropy $S(X)$. This quantity called the exponential entropy $E(X)$ (Padma 1984), and given by

$$E(X) = \exp S(X) = \exp\{-\int p(x)\log p(x)\mathrm{d}x\} \tag{18}$$

is clearly physically meaningful and has the same physical dimension D as the random quantity X. Further since $E(X)$ is a monotonic function of $S(X)$, it is thus as good a measure of the uncertainty in X as $S(X)$ is taken to be. The quantity $E(X)$ can be defined as the exponential of $S(X)$ for a discrete random variable also, but in this case both $S(X)$ and $E(X)$ are pure dimensionless numbers. It may also be noted that while we have $E(X) \geqslant 1$ for a discrete random variable (because $S(X) \geqslant 0$), we only have $E(X) \geqslant 0$ for a continuous random variable, since $S(X)$ given by (17) is not non-negative in general.

We are now in a position to give a physically meaningful formulation of the position-momentum uncertainty relation in the (exponential) entropic form. If $E^{\rho}(Q)$, $E^{\rho}(K)$ and

$E^\rho(P)$ are the exponential entropies (in state ρ) of position, wavenumber and momentum respectively, which are defined in terms of the appropriate probability densities as in (18), then from (15) we can clearly deduce that

$$E^\rho(Q)E^\rho(K) \geqslant \pi e \tag{19}$$

and since it can easily be seen that

$$E^\rho(P) = \hbar E^\rho(K) \tag{20}$$

we finally get the required uncertainty relation (Padma 1984):

$$E^\rho(Q)E^\rho(P) \geqslant \hbar\pi e. \tag{21}$$

Unlike (15), the relations (19)–(21) are all physically meaningful as the quantities $E^\rho(Q)$, $E^\rho(K)$ and $E^\rho(P)$ have the dimensions of Q, K and P respectively, and both sides of these equations have meaningful and matching physical dimensions.

Since all that goes into the proof of (21) or of (15) are certain basic results of the Fourier transform theory, the uncertainty relations (21) can indeed be shown to be valid for any pair of canonically conjugate observables (satisfying the Weyl form of CCR). An equally important result, which was demonstrated by Everett in 1957 itself (Everett 1973), is that (21) is stronger than the conventional Heisenberg relation

$$(\Delta^\rho Q)(\Delta^\rho P) \geqslant \hbar/2 \tag{22}$$

This can be easily seen by considering the inequalities

$$\Delta^\rho Q \geqslant \frac{1}{(2\pi e)^{1/2}} E^\rho(Q) \tag{23a}$$

$$\Delta^\rho P \geqslant \frac{1}{(2\pi e)^{1/2}} E^\rho(P) \tag{23b}$$

which are in fact true of any continuous probability density defined on the whole real line (McElice 1977). From (23a, b) and (21) we get the inequality

$$(\Delta^\rho Q)(\Delta^\rho P) \geqslant \frac{1}{2\pi e} E^\rho(Q)E^\rho(P) \geqslant \hbar/2 \tag{24}$$

thus showing that the entropic form of the uncertainty relation for canonically conjugate variables (21), is indeed stronger than the conventional variance form (22).

We thus see that an entropic formulation of the uncertainty relation is available for any arbitrary pair of observables with discrete spectra and also for canonically conjugate observables; and in both cases the entropic form of the uncertainty relation is superior to the standard form involving variances. The problem that still remains is one of extending the entropic formulation to the case of arbitrary self-adjoint operators, and this does not appear straightforward. In fact for the case of observables with continuous spectra, the notion of entropy itself is easily definable only if the spectrum is absolutely continuous (Grabowski 1978a). However it must be mentioned that the entropic formulation appears to be extendable even to those situations where we cannot characterise an observable by a self-adjoint operator but can still associate a probability distribution with it, as has been demonstrated by Bialynicki-Birula and Mycielski (1975) with their formulation of the angle-angular momentum uncertainty relation.

2. Uncertainty relation for successive measurements

One important feature which is common to the uncertainty relations in the variance form (2) and those in the entropic form (9), (12) is that they both refer to distinct measurements of A, B performed on different (though identically prepared) ensembles of systems in state ρ. This is clear from the fact that both the variance $(\Delta^\rho A)^2$ and the entropy $S^\rho(A)$ refer to an experimental situation where an ensemble of systems in state ρ is subjected to a measurement of observable A with no other observation carried out prior to that. In the same way $(\Delta^\rho B)$ and $S^\rho(B)$ refer to an experimental situation where an ensemble of systems in state ρ is subjected to a B-measurement with no other observation carried out prior to that.* Thus the uncertainty relations either in the form (2) or in the form (9), (12) clearly refer to distinct measurements of A, B performed on *different* ensembles of systems in state ρ and hence they are sometimes referred to as the uncertainty relations for distinct measurements (Gnanapragasam and Srinivas 1979).

In order to give content to the various remarks which are often made in the context of uncertainty relations concerning the interference of a measurement of one observable A on the outcomes of another measurement of observable B, one will have to consider an entirely different experimental situation where the *same* ensemble of systems is subjected to successive measurements, that of A followed by that of B. An uncertainty relation in the variance form, for such successive measurements was derived by Gnanapragasam and Srinivas (1979) and may be briefly recalled here. If the observables A, B have pure discrete spectra and the associated spectral resolutions are given by (10a, b), then the joint probability $\Pr^\rho_{A,B}(a_i, b_j)$ that when the sequence of measurements A, B are performed on an ensemble systems in state ρ, the value a_i results in the A-measurement and the value b_j results in the succeeding B-measurement, is given by the Wigner formula (Wigner 1963; Srinivas 1975)

$$\Pr^\rho_{A,B}(a_i, b_j) = \mathrm{Tr}(P^B(b_j)P^A(a_i)\rho P^A(a_i)P^B(b_j)). \tag{25}$$

The above expression for joint probability is a direct consequence of the collapse postulate or the projection postulate due to Von Neumann (1955) and Lüders (1951) which fixes the state of a system after the measurement of an observable with a purely discrete spectrum, such as A in (10a). It is the above joint probability (25) which will have to be employed in evaluating the variances $(\Delta^\rho A)^2_{A,B}$ and $(\Delta^\rho B)^2_{A,B}$ of the A-measurement and the B-measurement respectively, when an ensemble of systems in state ρ is subjected to a sequence of measurements A, B in that order. Thus we have

$$(\Delta^\rho A)^2_{A,B} = \sum_{i,j} a_i^2 \Pr^\rho_{A,B}(a_i, b_j) - \left\{\sum_{i,j} a_i \Pr^\rho_{A,B}(a_i, b_j)\right\}^2 \tag{26a}$$

$$(\Delta^\rho B)^2_{A,B} = \sum_{i,j} b_j^2 \Pr^\rho_{A,B}(a_i, b_j) - \left\{\sum_{i,j} b_j \Pr^\rho_{A,B}(a_i, b_j)\right\}^2 \tag{26b}$$

While the variance $(\Delta^\rho A)^2_{A,B}$ given by (26a) turns out to be the same as $(\Delta^\rho A)^2$ given by (1), the variance $(\Delta^\rho B)^2_{A,B}$ is in general different from $(\Delta^\rho B)^2$ given by

$$(\Delta^\rho B)^2 = \mathrm{Tr}(\rho B^2) - \{\mathrm{Tr}(\rho B)\}^2 \tag{29}$$

* Here and in what follows we take the observables A, B as being represented in the Heisenberg picture so that all time evolution, in time intervals where no measurements are made, is carried by the observables themselves.

In fact it can easily be seen from (26b) that

$$(\Delta^\rho B)^2_{A,B} = \mathrm{Tr}\,[\varepsilon^A(\rho)B^2] - [\mathrm{Tr}\,(\varepsilon^A(\rho)B)]^2 \tag{30}$$

where the density operator $\varepsilon^A(\rho)$ is given by

$$\varepsilon^A(\rho) = \sum_i P^A(a_i)\,\rho\, P^A(a_i). \tag{31}$$

This difference between $(\Delta^\rho B)^2_{A,B}$ and $(\Delta^\rho B)^2$ is merely a manifestation of the well known 'quantum interference of probabilities' (de Broglie 1948; Srinivas 1975) as the former $[(\Delta^\rho B)^2_{A,B}]$ refers to the variance in the outcomes of B-measurement when an ensemble of systems in state ρ is subjected to the sequence of measurements A, B, while the latter $[(\Delta^\rho B)^2]$ refers to the variance in the outcomes of B-measurement performed on an ensemble of systems in state ρ when no other observations are performed prior to the B-measurement.

The uncertainty relation for successive measurements in the usual variance form can be derived straightaway by considering the following standard inequality

$$\left\{\sum_{i,j} a_i^2 p_{ij} - \left(\sum_{i,j} a_i p_{ij}\right)^2\right\} \times \left\{\sum_{i,j} b_j^2 p_{ij} - \left(\sum_{i,j} b_j p_{ij}\right)^2\right\}$$
$$\geqslant \left|\left\{\sum_{i,j} a_i b_j p_{ij} - \left(\sum_{i,j} a_i p_{ij}\right)\left(\sum_{i,j} b_j p_{ij}\right)\right\}\right|^2 \tag{32}$$

which is valid for any probability distribution* p_{ij} (*i.e.* satisfying $0 \leqslant p_{ij} \leqslant 1$ and $\sum_{i,j} p_{ij} = 1$). If we now substitute the joint probability (25) in place of p_{ij} in (32) we get the following uncertainty relation (Gnanapragasam and Srinivas 1979):

$$(\Delta^\rho A)^2_{A,B}(\Delta^\rho B)^2_{A,B} \geqslant |\langle A\varepsilon^A(B)\rangle_\rho - \langle A\rangle_\rho \langle \varepsilon^A(B)\rangle_\rho|^2 \tag{33}$$

where

$$\varepsilon^A(B) = \sum_i P^A(a_i)\,B\,P^A(a_i). \tag{34}$$

The above uncertainty relation for successive measurements (33), being formulated in terms of variances, suffers from the same limitations that were noted in §1 in connection with the standard uncertainty relation for distinct measurements (2). Further, since we are dealing with only observables with purely discrete spectra, the inadequacy of (33) is much more obvious and was noted in the same paper (Gnanapragasam and Srinivas 1979) where this relation was derived. The main problem with (33) is again that it does not provide a lower bound on the variance of B given the variance of A (in the case of successive measurements) as the infinum of the right side of (33) taken over all states ρ, vanishes *always* as the operators A and $\varepsilon^A(B)$ commute. This is all the more unfortunate as one would have expected to learn from an uncertainty relation for successive measurements how a prior measurement of one observable say A influences the uncertainty in the outcomes of a following measurement of another observable B. As we shall see, the fault entirely lies with the choice of variance as the measure of uncertainty in (33); in fact there is an entropic formulation of the uncertainty

* The inequality (32) is nothing but the standard inequality of classical probability theory, variance X variance $Y \geqslant |\text{covariance}\,(X, Y)|^2$, valid for any pair of random variables X, Y.

relation for successive measurements which is free from the above defect and provides a clear estimate of the interference of one measurement on the uncertainty in the outcomes of the other.

We now introduce the entropies $S^{\rho}_{A,B}(A)$, $S^{\rho}_{A,B}(B)$ which give an appropriate measure of the uncertainty in the outcomes of an A-measurement and a B-measurement when an ensemble of systems in state ρ is subjected to the sequence of measurements A, B. Clearly these entropies (like the variances (26a, b)) have to be evaluated in terms of the joint probability (25) which is appropriate for the given experimental situation. We therefore have

$$S^{\rho}_{A,B}(A) = -\sum_i \left\{ \sum_j \Pr^{\rho}_{A,B}(a_i, b_j) \log \sum_j \Pr^{\rho}_{A,B}(a_i, b_j) \right\} \tag{35a}$$

$$S^{\rho}_{A,B}(B) = -\sum_j \left\{ \sum_i \Pr^{\rho}_{A,B}(a_i, b_j) \log \sum_i \Pr^{\rho}_{A,B}(a_i, b_j) \right\}. \tag{35b}$$

From the basic properties of the spectral projectors $P^A(a_i)$, *viz*,

$$P^A(a_i)P^A(a_j) = \delta_{ij} P^A(a_i) \tag{36a}$$

$$\sum_i P^A(a_i) = I \tag{36b}$$

where I is the identity operator, we can easily obtain the relation

$$S^{\rho}_{A,B}(A) = S^{\rho}(A) = S^{\varepsilon^A(\rho)}(A) \tag{37}$$

where $S^{\rho}(A)$ and $\varepsilon^A(\rho)$ are given by (11a) and (31) respectively. However, again because of the quantum interference of probabilities, the entropy $S^{\rho}_{A,B}(B)$ is in general different from the entropy $S^{\rho}(B)$ given by (11b), as they refer to different experimental situations. In fact from (35b) we can easily show that

$$S^{\rho}_{A,B}(B) = S^{\varepsilon^A(\rho)}(B) \tag{38}$$

so that $S^{\rho}_{A,B}(B)$ and $S^{\rho}(B)$ will coincide only for those states ρ which satisfy $\varepsilon^A(\rho) = \rho$.

At this stage, we can employ (37) and (38) together with the Deutsch-Partovi inequality (12) to obtain

$$S^{\rho}_{A,B}(A) + S^{\rho}_{A,B}(B) = S^{\varepsilon^A(\rho)}(A) + S^{\varepsilon^A(\rho)}(B) \geqslant 2 \log \frac{2}{\sup_{i,j} \| P^A(a_i) + P^B(b_j) \|}. \tag{39}$$

This shows that the Deutsch-Partovi lower bound for the sum of uncertainties in the case of distinct measurements, is valid also for the case of successive measurements. However, while in the case of distinct measurements (of observables with discrete spectra) the lower bound given by the right side of (12) is indeed optimal, it is not so for the case of successive measurements. This should be obvious from the fact that unlike in (12), in the left side of (39) only states of the form $\varepsilon^A(\rho)$ are involved. In fact we shall now obtain a much stronger inequality than (39) for the case of successive measurements thus showing that the interference of one measurement on the other does indeed contribute to an 'overall increase in uncertainty'.

For this purpose we consider the joint entropy $S^{\rho}_{A,B}(A, B)$ of the observables A, B, when they are successively measured, in that order on an ensemble of systems in state ρ.

Clearly $S^{\rho}_{A,B}(A, B)$ should be defined in terms of the joint probability (25) as (Srinivas 1978)

$$S^{\rho}_{A,B}(A, B) = -\sum_{i,j} \Pr{}^{\rho}_{A,B}(a_i, b_j) \log \Pr{}^{\rho}_{A,B}(a_i, b_j) \tag{40}$$

If we now employ the standard inequality

$$-\sum_i \left(\sum_j p_{ij} \log \sum_j p_{ij} \right) - \sum_j \left(\sum_i p_{ij} \log \sum_i p_{ij} \right) \geqslant -\sum_{i,j} p_{ij} \log p_{ij} \tag{41}$$

which is valid for any probability distribution* p_{ij}, and if we substitute the joint probability (25) in place of p_{ij} in (41), then we get the inequality

$$S^{\rho}_{A,B}(A) + S^{\rho}_{A,B}(B) \geqslant S^{\rho}_{A,B}(A, B) \tag{42}$$

which was noted sometime ago in the context of quantum information theory (Srinivas 1978). Further if we note that

$$\Pr{}^{\rho}_{A,B}(a_i, b_j) = \operatorname{Tr}(\rho P^A(a_i) P^B(b_j) P^A(a_i)) \leqslant \| P^A(a_i) P^B(b_j) P^A(a_i) \| \tag{43}$$

which follows from the fact that the density operator is of trace-norm one, then we can deduce from (41)–(43) the inequality

$$S^{\rho}_{A,B}(A) + S^{\rho}_{A,B}(B) \geqslant \log \left\{ \sup_{i,j} \| P^A(a_i) P^B(b_j) P^A(a_i) \| \right\}^{-1}. \tag{44}$$

Relation (44) is the desired uncertainty relation for successive measurements in the entropic form. To see that (44) is indeed stronger than (39) let us note the following inequality

$$4 \| PQP \| \leqslant \| P + Q \|^2 \tag{45}$$

which is valid for any pair of projection operators** P, Q. We therefore have

$$\log \frac{1}{\| PQP \|} \geqslant 2 \log \frac{2}{\| P + Q \|} \tag{46}$$

so that

$$\begin{aligned} S^{\rho}_{A,B}(A) + S^{\rho}_{A,B}(B) &\geqslant \log \left\{ \frac{1}{\sup_{i,j} \| P^A(a_i) P^B(b_j) P^A(a_i) \|} \right\} \\ &\geqslant 2 \log \left\{ \frac{2}{\sup_{i,j} \| P^A(a_i) + P^B(b_j) \|} \right\} \end{aligned} \tag{47}$$

* The inequality (41) is nothing but the standard inequalitty of classical information theory $S(X) + S(Y) \geqslant S(X, Y)$, which is valid for any pair of discrete random variables X, Y.

** To derive (45) we need to proceed as follows:

$$\begin{aligned} 2\|PQ\| &= \|PQ + PQ\| = \|P(Q + PQ)\| \\ &\leqslant \|Q + PQ\| = \|(Q + P)Q\| \\ &\leqslant \|P + Q\|. \end{aligned}$$

Therefore

$$4\|PQP\| \leqslant 4\|PQ\|^2 \leqslant \|P + Q\|^2.$$

thus showing that inequality (44) is stronger than (39). In the process we have also clearly shown that the lower bound for the sum of uncertainties is indeed greater for the case of successive measurements than for the case of distinct measurements. This of course confirms the conventional wisdom that the interference of one measurement on the other should contribute to a larger overall uncertainty in the case of successive measurements. Actually an estimate of the interference caused by an earlier A-measurement on the uncertainty in the outcome of a later B-measurement can be had from the following inequality

$$S^{\rho}_{A,B}(B) \geqslant \log\left\{\frac{1}{\sup_j \|\varepsilon^A(P^B(b_j))\|}\right\} \tag{48}$$

where

$$\varepsilon_A(P^B(b_j)) = \sum_i P^A(a_i)P^B(b_j)P^A(a_i). \tag{49}$$

The inequality (48) follows directly from (38) and reveals the following very important fact that while the lower bound of $S^{\rho}(B)$ is always zero (which is attained for eigenstates of B), the lower bound of $S^{\rho}_{A,B}(B)$ can be strictly non-negative. In other words if

$$\sup_j \|\varepsilon^A(P^B(b_j))\| < 1 \tag{50}$$

then whatever be the initial state ρ of the system, the outcome of a B-measurement which follows an A-measurement is *always* uncertain. The inequality (48) appears to be the clearest statement of the fact that an earlier A-measurement in itself can induce a 'non-zero uncertainty' in the outcome of a later B-measurement, whatever be the initial state of the system. We should however note that the lower bound derived for the sum of uncertainties $S^{\rho}_{A,B}(A) + S^{\rho}_{A,B}(B)$ in (44), is in general greater than the lower bound obtained in (48) for $S^{\rho}_{A,B}(B)$ alone. This follows from the relation

$$P^A(a_i)P^B(b_j)P^A(a_i) = P^A(a_i)\,\varepsilon^A(P^B(b_j))\,P^A(a_i)$$

which gives rise to the inequality

$$\sup_{i,j} \|P^A(a_i)P^B(b_j)P^A(a_i)\| \leqslant \sup_j \|\varepsilon^A(P^B(b_j))\|. \tag{51}$$

Thus the lower bound on the sum of uncertainties $S^{\rho}_{A,B}(A) + S^{\rho}_{A,B}(B)$ as given by (44) does not always arise solely from the interference of the earlier A-measurement.*

We may now note some of the salient features of the uncertainty relation (44). Clearly

---

*We may however note that when both A and B have a purely non-degenerate spectra as given by (13a, b), then the inequality (51) will reduce to an equality because then

$$\|\varepsilon^A(P^B(b_j))\| = \sup_i |\langle a_i|b_j\rangle|^2$$

so that

$$\sup_j \|\varepsilon^A(P^B(b_j))\| = \sup_{i,j} |\langle a_i|b_j\rangle|^2$$
$$= \sup_{i,j} \|P^A(a_i)P^B(b_j)P^A(a_i)\|.$$

the right side of(44) is non-negative and vanishes only when the observables A, B are such that

$$\sup_{i,j} \|P^A(a_i)P^B(b_j)P^A(a_i)\| = 1. \tag{52}$$

It is a well-known result (Rehder 1979) that for any two projectors P, Q and any vector ψ, the relation

$$\|PQP\psi\| = 1$$

will be satisfied if and only if

$$\psi \in \text{Range } P \cap \text{Range } Q.$$

From the above it follows that (52) will be satisfied (*i.e.* the right side of (44) will vanish) essentially only when the observables A, B have a joint eigenvector. This could have also been seen directly from the inequality (47) and our remarks concerning the limiting case of the Deutsch-Partovi relation. Also if observables A, B have a purely non-degenerate spectrum as given by (13a, b), then (44) and (47) reduce to

$$S^{\rho}_{A,B}(A) + S^{\rho}_{A,B}(B) \geqslant \log \frac{1}{\sup_{i,j} |\langle a_i|b_j\rangle|^2} \geqslant 2 \log \frac{2}{1 + \sup_{i,j} |\langle a_i|b_j\rangle|}. \tag{53}$$

Another curious feature of the uncertainty relation for successive measurements (44) follows from the following relation (Rehder 1979)

$$\|PQP\| = \|QPQ\| \tag{54}$$

valid for all projectors P, Q. From (54) it follows that the lower bound on the sum $S^{\rho}_{A,B}(A) + S^{\rho}_{A,B}(B)$ is the same as that for the sum $S^{\rho}_{B,A}(A) + S^{\rho}_{B,A}(B)$. In other words the lower bound on the sum of uncertainties is the same if the B-measurement *immediately* follows the A-measurement or vice-versa.

Finally we may note that many of the results obtained in this section for the case of a sequence of two measurements can be easily generalised to any arbitrary sequence of measurements, provided we still restrict ourselves to a consideration of observables with discrete spectra only. For instance if C is an observable with spectral resolution

$$C = \sum_K C_K P^C(c_k) \tag{55}$$

and if we consider the sequence of measurements A, B, C performed on an ensemble of systems originally prepared in state ρ, then our results (42), (44) and (48) can be generalised (with obvious extension of the notation) to

$$S^{\rho}_{A,B,C}(A) + S^{\rho}_{A,B,C}(B) + S^{\rho}_{A,B,C}(C) \geqslant S^{\rho}_{A,B,C}(A/B, C)$$
$$\geqslant \log \left\{ \sup_{i,r,k} \|P^A(a_i)P^B(b_r)P^C(c_k)P^B(b_r)P^A(a_i)\| \right\}^{-1} \tag{56}$$

$$S^{\rho}_{A,B,C}(C) \geqslant \log \frac{1}{\sup_k \|\varepsilon^A \varepsilon^B P^C(c_k)\|} \tag{57}$$

What however appears to be a formidable problem is the extension of these results on uncertainties in successive observations to the case of observables with a continuous spectrum. This is because of the general no-go theorem proved recently (Srinivas 1980) that any physically meaningful extension of the Von Neumann-Lüders collapse postulate to the case of observables with a continuous spectrum will have to necessarily allow measurement transformations which take density-operator states into the class of 'non-normal states', so that the joint probabilities associated with successive measurements will no longer be always σ-additive, but only finitely additive. There is very little chance for notions such as variance or entropy to make sense under such 'pathological' conditions.

Acknowledgements

The author is grateful to R Padma and his colleagues for helpful discussions.

References

Beckner W 1975a *Ann. Math.* **102** 159
Beckner W 1975b *Proc. Natl. Acad. Sci. USA* **72** 638
Bialynicki-Birula I 1984 *Comment on entropic uncertainty relations*, Warsaw Preprint
Bialynicki-Birula I and Mycielski J 1975 *Commun. Math. Phys.* **44** 129
de Broglie L 1948 La Revue Scientifique No 3292 Fasc 5, 87, 259
Deutsch D 1983 *Phys. Rev. Lett.* **50** 631
Everett H 1973 in *The many-world interpretation of quantum mechanics.* (eds) B S DeWitt and N Graham (Princeton: University Press)
Gnanapragasam B and Srinivas M D 1979 *Pramana* **12** 699
Grabowski M 1978a *Rep. Math. Phys.* **14** 377
Grabowski M 1978b *Int. J. Theor. Phys.* **17** 635
Heisenberg W 1927 *Z. Phys.* **43** 172
Hirschman I I 1957 *Am. J. Math.* **79** 152
Ingarden R S 1976 *Rep. Math. Phys.* **10** 43, 131
Khinchin A I 1957 *Mathematical foundations of information theory* (New York: Dover)
Lindblad 1973 *Commun. Math. Phys.* **33** 305
Lüders G 1951 *Ann. Phys.* **8** 322
McElice R J 1977 *The theory of information and coding* (Boston: Addisson Wesley)
Padma R 1984 *Entropic formulation of uncertainty relations*, MSc Thesis (unpublished) Madras University, Madras
Partovi M H 1983 *Phys. Rev. Lett.* **50** 1883
Rehder W 1979 *Intl. J. Theor. Phys.* **18** 791
Renyi A 1961 in *Proceedings of 4th Berkeley Symp. on Math. Statistics and Probability* (Berkeley: University Press) Vol. I p. 547–561
Robertson H P 1929 *Phys. Rev.* **34** 163
Schrödinger E 1930 *Ber. Bericht.* 296
Srinivas M D 1975 *J. Math. Phys.* **16** 1676
Srinivas M D 1978 *J. Math. Phys.* **19** 1952
Srinivas M D 1980 *Commun. Math. Phys.* **71** 131
Von Neumann J 1955 *Mathematical foundations of quantum mechanics* (Princeton: University Press)
Wehrl A 1974 *Rep. Math. Phys.* **6** 15
Wehrl A 1978 *Rev. Mod. Phys.* **50** 221
Wigner E P 1963 *Am. J. Phys.* **31** 6