

Quantum-mechanical histories and the uncertainty principle: Information-theoretic inequalities

J. J. Halliwell*

Theory Group, Blackett Laboratory, Imperial College, London SW7 2BZ, United Kingdom

(Received 3 May 1993)

This paper is generally concerned with understanding how the uncertainty principle arises in formulations of quantum mechanics, such as the decoherent histories approach, whose central goal is the assignment of probabilities to histories. We first consider histories characterized by position or momentum projections at two moments of time. Both exact and approximate (Gaussian) projections are studied. Shannon's information is used as a measure of the uncertainty expressed in the probabilities for these histories. We derive a number of inequalities in which the uncertainty principle is expressed as a lower bound on the information of phase space distributions derived from the probabilities for two-time histories. We go on to consider histories characterized by position samplings at n moments of time. We derive a lower bound on the information of the joint probability for n position samplings. Similar bounds are derived for histories characterized by samplings of other variables. All lower bounds on the information of histories have the general form $\ln(V_H/V_S)$, where V_H is a volume element of history space, which we define, and V_S is the volume of that space probed by the projections. We thus obtain a concise and general form of the uncertainty principle referring directly to the histories description of the system, and making no reference to notions of phase space.

PACS number(s): 03.65.Bz, 02.50.Cw, 05.20.-y

I. INTRODUCTION

A quantum-mechanical history is defined by an initial quantum state at some time t_0 and by a sequence of propositions at a succession of times t_1, t_2, \dots, t_n . The initial state is represented by a density matrix ρ . Each proposition is represented by a set of projection operators P_α . These are positive Hermitian operators that are both exclusive and exhaustive:

$$P_\alpha P_\beta = \delta_{\alpha\beta} P_\alpha, \quad (1.1)$$

$$\sum_\alpha P_\alpha = 1. \quad (1.2)$$

Evolution between each projection is described by the unitary evolution operator $e^{-iHt/\hbar}$. The probability for histories described in this way is given by the expression

$$p(\alpha_1, \alpha_2, \dots, \alpha_n) = \text{Tr}[P_{\alpha_n}^n(t_n) \cdots P_{\alpha_1}^1(t_1) \rho P_{\alpha_1}^1(t_1) \cdots P_{\alpha_n}^n(t_n)], \quad (1.3)$$

where

$$P_{\alpha_k}^k(t_k) = e^{iH(t_k - t_0)/\hbar} P_{\alpha_k}^k e^{-iH(t_k - t_0)/\hbar}. \quad (1.4)$$

Equation (1.3) is central to any formulation of quantum mechanics whose aim is the assignment of probabilities to histories. One particular such approach is the decoherent histories approach [1–5]. In that approach the central aim is to find, for a given Hamiltonian and initial state, the sets of histories of closed quantum systems for which the probabilities (1.3) satisfy the so-called

“probability sum rules.” Loosely, these are the rules obtained by demanding that the probability of a composite history be the sum of the probabilities of the more elementary histories of which it is comprised. An example of such a sum rule (of which there are many) is

$$p(\cdots \alpha_{k-1}, \alpha_{k+1} \cdots) = \sum_{\alpha_k} p(\cdots \alpha_{k-1}, \alpha_k, \alpha_{k+1} \cdots). \quad (1.5)$$

Histories satisfying these rules are said to be “consistent” or “decoherent,” and it is solely in terms of such histories that predictions may be made. Quantum-mechanical interference means that these rules are generally not satisfied, and demonstrating consistency is typically non-trivial.

Formula (1.3) also arises in a different context. It is a concise summary of the Copenhagen approach to the quantum mechanics of measured subsystems. It incorporates both the unitary evolution of states together with the “collapse of the wave function” incurred as a result of measurement by an external agency, modeled by the projection operators [6,7].

Irrespective of which interpretational scheme one is concerned with, the mathematical properties of the expression (1.3) are of interest. This paper is concerned with exploring those properties.

Our particular concern is the question of how the uncertainty principle arises in (1.3). The usual form

$$\Delta p \Delta q \geq \frac{\hbar}{2} \quad (1.6)$$

is a simple consequence of the Fourier transform of the wave function of the system at a fixed moment of time. However, in formulations that give a central role to (1.3),

*E-mail address: j_halliwell@vax1.physics.imperial.ac.uk

the state of the system at a fixed moment of time does not enter in a fundamental way. Instead, all physically meaningful notions must be expressed through the probabilities (1.3). It therefore becomes an important issue to understand how these probabilities recognize the uncertainty principle. It is not difficult to see that it will arise as a limitation on the degree to which (1.3) may be peaked about a particular history. This is because the probability (1.3) is a distribution over quantities that are generally noncommuting, and so one would not expect it to become arbitrarily peaked. The aim of this paper is to establish the detailed form this limitation takes.

As a measure of the degree to which (1.3) is peaked, we shall use the Shannon information

$$I = - \sum_{\alpha_1, \dots, \alpha_n} p(\alpha_1, \dots, \alpha_n) \ln p(\alpha_1, \dots, \alpha_n) . \quad (1.7)$$

This measure, for histories, is in many ways more natural and easier to use than the variances, employed in (1.6). We shall show that the uncertainty principle generally arises as a lower bound on the information (1.7). In particular, for the case in which the alternatives α_k are discrete, the probabilities (1.3) have an upper bound p_{\max} , over all initial states ρ and over all possible values of the alternatives. If there is a restriction on the degree to which (1.3) is peaked, as one would expect when the projections do not commute, then $p_{\max} < 1$. The information (1.7) then has a nontrivial lower bound

$$I \geq \ln \left(\frac{1}{p_{\max}} \right) . \quad (1.8)$$

We begin in Sec. II with a brief review of some properties of Shannon information. We then go on in Sec. III to discuss information-theoretic measures of uncertainty in quantum mechanics. We review earlier work on information-theoretic versions of the uncertainty principle, expressed in terms of the state of the system at a fixed moment of time.

In Secs. IV and V we discuss quantum-mechanical histories of the form (1.3) characterized by position and/or momentum projections at two moments of time. We consider the case of both exact and approximate (Gaussian) projection operators. The general idea is to use the two-time histories to derive imprecise samplings of phase space and then compute lower bounds on the information of the quantum-mechanical phase space distributions. In regimes where they are nontrivial, we find that all of the bounds have the approximate form

$$I(K, X) \geq \ln \left(\frac{2\pi\hbar}{\sigma_x \sigma_k} \right) , \quad (1.9)$$

where $I(K, X)$ is the information of the phase space distributions and $\sigma_x \sigma_k$ is the volume of phase space probed by the projections.

In Sec. VI we go on to study histories characterized by position samplings at n moments of time. We show that the uncertainty principle arises as a restriction on the information of the approximate form

$$I(X_1, X_2, \dots, X_n) \geq \ln \left(\frac{V_H}{\sigma_1 \sigma_2^2 \cdots \sigma_{n-1}^2 \sigma_n} \right) \quad (1.10)$$

in the regime where it is nontrivial. Here σ_i is the width of the position sampling i and V_H^{-1} is a ‘‘density of paths’’ factor. We argue that V_H thus has the interpretation as the ‘‘fundamental volume of history space,’’ analogous to the factor of $2\pi\hbar$ in (1.9). We derive a result identical in form for histories characterized by other types of projections. We thus obtain a form of the uncertainty principle which is both concise and general, and is phrased entirely in the language of histories, without reference to phase space. We summarize and discuss in Sec. VII.

Some words are in order concerning the use of Shannon information for the probabilities (1.3). Since the quantities defined by (1.3) generally do not in fact satisfy the probability sum rules, such as (1.5), they cannot strictly be regarded as probabilities. Use of the Shannon information (1.7) therefore requires some qualification. Although they do not obey the probability sum rules, the (candidate) probabilities (1.3) are non-negative and normalized, and thus the information (1.7) is a well-defined quantity and may be used as a measure of the degree of spread of the candidate probability. The important point is that at no stage are the probabilities sum rules assumed, and thus no inconsistencies arise.

It is of course an interesting question, from the perspective of the decoherent histories approach, to extend the considerations of the present paper to the case in which the candidate probabilities (1.3) do obey the probability sum rules. Decoherence may be achieved, for example, by coupling the system of interest to an environment. Modifications of the uncertainty relations (1.9) and (1.10) due to environmentally induced (e.g., thermal) fluctuations can then be expected. This is considered in Refs. [8,9]. The information-theoretic inequalities considered here then become conditions that such decohering probabilities must satisfy in the limit that the coupling to the environment goes to zero.

II. INFORMATION THEORY

In this section we briefly review some results from information theory. This section solely concerns generic probability distributions and makes no reference to quantum mechanics.

Let p_i be the probabilities for a data set S consisting of a discrete set of alternatives labeled by i , $i = 1, 2, \dots, N$. One has $0 \leq p_i \leq 1$ and $\sum_i p_i = 1$. The *information* of the data set S is defined to be

$$I(S) = - \sum_{i=1}^N p_i \ln p_i . \quad (2.1)$$

Here \ln is the logarithm to base e . $I(S)$ satisfies the inequalities

$$0 \leq I(S) \leq \ln N . \quad (2.2)$$

It reaches its minimum if and only if $p_i = 1$ for one particular value of i , and so $p_i = 0$ for all the other values. It reaches its maximum when $p_i = 1/N$ for all i . The infor-

mation of a probability distribution is therefore a measure of how strongly peaked it is about a given alternative. For this reason $I(S)$ is sometimes referred to as *uncertainty*, being large for spread out distributions and small for concentrated ones. $I(S)$ is sometimes also referred to as the *entropy* of the distribution, but we shall not use that nomenclature here.

Base 2 is often used in the definition (2.1). In this case $I(S)$ has the interpretation as the average number of bits required to specify an alternative given that alternative i occurs with probability p_i .

Information may also be defined for continuous probability distributions. Let X be a random variable with probability density $p(x)$. Then $\int dx p(x) = 1$. The information of X is defined to be

$$I(X) = - \int dx p(x) \ln p(x) . \quad (2.3)$$

Unlike the discrete case, $I(X)$ is no longer positive, since $p(x)$ is not a probability, but a probability density, and so may be greater than 1. However, it retains its utility as a measure of uncertainty. This is exemplified by a Gaussian distribution of variance Δx ,

$$p(x) = \frac{1}{[2\pi(\Delta x)^2]^{1/2}} \exp \left[- \frac{(x-x_0)^2}{2(\Delta x)^2} \right] . \quad (2.4)$$

It has the information

$$I(X) = \ln[2\pi e (\Delta x)^2]^{1/2} . \quad (2.5)$$

From this we see that $I(X)$ may be unbounded from below and, indeed, approaches $-\infty$ as $\Delta x \rightarrow 0$ and $p(x)$ approaches a δ function. $I(X)$ is also unbounded from above, as may be seen by taking the width Δx to be very large. However, if the variance is fixed, then a straightforward variational calculation shows that $I(X)$ is maximized by the Gaussian distribution (2.4). Equation (2.5) therefore represents an upper bound on the information of probability distributions with variance Δx :

$$I(X) \leq \ln[2\pi e (\Delta x)^2]^{1/2} , \quad (2.6)$$

with equality if and only if $p(x)$ is a Gaussian.

The literature contains a vast number of results about information. We will record only one, since it will be needed later. Suppose from a probability distribution $p(x)$ one constructs a ‘‘coarser-grained’’ probability distribution

$$q(\bar{x}) = \int dx f(\bar{x}, x) p(x) \quad (2.7)$$

for some smearing function $f(\bar{x}, x)$ satisfying $\int d\bar{x} f(\bar{x}, x) = 1$. Then, if we denote the information of $q(\bar{x})$ by $I(\bar{X})$, it may be shown that

$$I(\bar{X}) \geq I(X) . \quad (2.8)$$

This inequality expresses the intuitive idea that smearing or coarse graining a probability distribution increases the amount of uncertainty it expresses. A corresponding result also holds for the discrete case. The result, for both the continuous and discrete case, follows readily from the convexity of the function $x \ln x$, and so we shall refer to

this result as the convexity property.

For further details on information theory, see Refs. [10,11].

III. INFORMATION-THEORETIC UNCERTAINTY RELATIONS

We now describe a number of information-theoretic expressions of the uncertainty principle. We begin by describing the projection operators used to sample position and momentum.

A. Samplings of position and momentum

Approximate samplings of position may be carried out using projection operators. The projection operators effect a partition of the real line into regions (or ‘‘bins’’) of size σ_x . Explicitly, they take the form

$$P_\alpha^x = \int dx \Upsilon(x - \bar{x}_\alpha) |x\rangle \langle x| , \quad (3.1)$$

where $\Upsilon(x - \bar{x}_\alpha)$ is a sampling function. The most appropriate choice is to take it to be

$$\Upsilon(x - \bar{x}_\alpha) = \theta \left[\frac{x - \bar{x}_\alpha + \frac{1}{2}\sigma_x}{\sigma_x} \right] \theta \left[\frac{-x + \bar{x}_\alpha + \frac{1}{2}\sigma_x}{\sigma_x} \right] . \quad (3.2)$$

It is equal to 1 in an interval of size σ_x centered around \bar{x}_α and zero otherwise, where $\bar{x}_\alpha = \alpha\sigma_x$ and α is an integer. We will generally use a bar to denote coarse-grained variables. The sampling function satisfies the relations

$$\sum_\alpha \Upsilon(x - \bar{x}_\alpha) = 1 , \quad (3.3)$$

$$\int dx \Upsilon(x - \bar{x}_\alpha) = \sigma_x . \quad (3.4)$$

Equation (3.3) ensures that the projections are exhaustive. They are exclusive because Υ vanishes outside a unit interval.

Another choice for Υ which is sometimes convenient is a Gaussian of width σ_x :

$$\Upsilon(x - \bar{x}_\alpha) = \frac{1}{(2\pi)^{1/2}} \exp \left[- \frac{(x - \bar{x}_\alpha)^2}{2\sigma_x^2} \right] . \quad (3.5)$$

Again, $\bar{x}_\alpha = \alpha\sigma_x$, but α is now a continuous label. The properties (3.3) and (3.4) still hold, given the convention that the summation over α is now an integration. With this choice of Υ , the projections are only approximately exclusive. This means that the label α , although continuous, really has significance only up to order 1.

The case of precise samplings, $P_{\bar{x}} = |\bar{x}\rangle \langle \bar{x}|$, is obtained by writing $P_{\bar{x}} = \sigma_x^{-1} P_\alpha^x$, and letting $\sigma_x \rightarrow 0$, and one has

$$\sigma_x^{-1} \Upsilon(x - \bar{x}_\alpha) \rightarrow \delta(x - \bar{x}) , \quad \sigma_x \sum_\alpha \rightarrow \int d\bar{x} . \quad (3.6)$$

In a similar manner, one may construct projections for samplings of momentum,

$$P_\beta^k = \int dk \Gamma(k - \bar{k}_\beta) |k\rangle \langle k| \quad (3.7)$$

for some sampling function $\Gamma(k - \bar{k}_\beta)$, of width σ_k , where $\bar{k}_\beta = \beta\sigma_k$.

B. Samplings of two ensembles

The first result we shall describe envisages a situation in which one has two ensembles, prepared in an identical state. Samplings of position are made on the first ensemble, and samplings of momentum are made on the second.

Consider a position sampling of a system described by a density matrix ρ . The probability that the result lies in the region labeled by α is

$$\begin{aligned} p^x(\alpha) &= \text{Tr}(P_\alpha^x \rho) \\ &= \int dx \Upsilon(x - \bar{x}_\alpha) \langle x | \rho | x \rangle. \end{aligned} \quad (3.8)$$

We wish to use the information $I_\rho(\bar{X})$ as a measure of uncertainty in the probability distribution $p^x(\alpha)$. By the convexity property (2.8), one has

$$\begin{aligned} I_\rho(\bar{X}) &\equiv - \sum_\alpha p^x(\alpha) \ln p^x(\alpha) \\ &\geq - \int dx \langle x | \rho | x \rangle \ln \langle x | \rho | x \rangle - \ln \sigma_x \\ &\equiv I_\rho(X) - \ln \sigma_x. \end{aligned} \quad (3.9)$$

The $\ln \sigma_x$ term arises because of (3.4).

In a similar manner, we can consider a momentum sampling on an identically prepared system, giving the probability distribution

$$p^k(\beta) = \text{Tr}(P_\beta^k \rho). \quad (3.10)$$

One may compute its information $I_\rho(\bar{K})$, and one has

$$I_\rho(\bar{K}) \geq I_\rho(K) - \ln \sigma_k. \quad (3.11)$$

A general density operator ρ may be written

$$\rho = \sum_i c_i |\psi_i\rangle \langle \psi_i| \quad (3.12)$$

for some set of states $|\psi_i\rangle$. Again, using the convexity property (2.8), one has

$$I_\rho(X) \geq \sum_i c_i I_{\psi_i}(X), \quad (3.13)$$

where $I_{\psi_i}(X)$ denotes the information of the probability distribution obtained from precise sampling of the pure state $|\psi_i\rangle$. A similar result holds for momentum samplings. It follows that there exists a pure state $|\psi\rangle$ such that

$$I_\rho(\bar{X}) + I_\rho(\bar{K}) \geq I_\psi(X) + I_\psi(K) - \ln(\sigma_x \sigma_k), \quad (3.14)$$

with equality for precise samplings and $\rho = |\psi\rangle \langle \psi|$. [Note that the $\ln(\sigma_x \sigma_k)$ term disappears in the limit of precise samplings, since it is taken into the projections on the left-hand side, as described above.]

$I_\psi(K)$ and $I_\psi(X)$ are individually unbounded from below, since one can always find states which are arbitrarily peaked in either position or in momentum. However, a state strongly peaked in position, and hence with large negative $I_\psi(X)$, will be very spread out in momen-

tum, and thus $I_\psi(K)$ will be large and positive. It is therefore plausible that the uncertainty principle will express itself as a lower bound on the sum $I_\psi(X) + I_\psi(K)$. The usual inequality expressing the uncertainty principle,

$$\Delta x \Delta k \geq \frac{\hbar}{2}, \quad (3.15)$$

achieves equality for the minimum uncertainty wave packets. Since they are Gaussians, we immediately have, from (2.5),

$$\begin{aligned} I_\psi(X) + I_\psi(K) &= \ln(2\pi e \Delta x \Delta k) \\ &= \ln(\pi e \hbar) \end{aligned} \quad (3.16)$$

for the minimum uncertainty wave packets (coherent states). It was therefore conjectured by Everett [12] that the uncertainty principle may be expressed in information-theoretic terms as

$$I_\psi(X) + I_\psi(K) \geq \ln(\pi e \hbar). \quad (3.17)$$

He also noted that this inequality implies the usual form of the uncertainty principle. To see this, recall from Sec. II that the information of a probability distribution is bounded from above by the information of a Gaussian of the same variance. This means that for any state $|\psi\rangle$, with variances Δx and Δk , one has

$$\ln(2\pi e \Delta x \Delta k) \geq I_\psi(X) + I_\psi(K). \quad (3.18)$$

The usual uncertainty principle then follows immediately by comparing (3.17) and (3.18). The inequality (3.17) was proved by Beckner [13], Bialynicki-Birula and Mycielski [14], and Hirschmann [15], using the Hausdorff-Young inequalities from Fourier analysis.

Combining all of the above results, we have

$$I_\rho(\bar{X}) + I_\rho(\bar{K}) \geq 1 + \ln \left[\frac{\pi \hbar}{\sigma_x \sigma_k} \right], \quad (3.19)$$

with equality for precise samplings and ρ a minimum uncertainty wave packet. Equation (3.19) represents a very modest generalization of (3.17) to the case of imprecise samplings of position and momentum.

C. Samplings of a single ensemble

Of greater interest for our purposes is the situation in which the samplings of position and momentum are made on the *same* system. There are a number of ways of doing this, and we shall consider them in turn. Perhaps the simplest is to carry out simultaneous but imprecise samplings of both position and momentum. These may be effected using the coherent state projectors, which we now describe.

The (canonical) coherent states [16] may be defined to be the states

$$|z\rangle = |p, q\rangle = U(p, q)|0\rangle, \quad (3.20)$$

where $|0\rangle$ is the ground state of the harmonic oscillator. $U(p, q)$ is the unitary Weyl operator

$$U(p, q) = \exp \left[\frac{i}{\hbar} (p\hat{Q} - q\hat{P}) \right], \quad (3.21)$$

where \hat{Q} and \hat{P} denote the position and momentum operators. In the position representation, the coherent states are given by

$$\langle x|p, q\rangle = \frac{1}{(2\pi\sigma_q^2)^{1/4}} \left[-\frac{(x-q)^2}{4\sigma_q^2} + \frac{i}{\hbar} px \right]. \quad (3.22)$$

Their most important property is the completeness relation

$$\int \frac{dp dq}{2\pi\hbar} |p, q\rangle \langle p, q| = 1. \quad (3.23)$$

They are, however, only approximately orthogonal:

$$\begin{aligned} \langle p, q|p', q'\rangle = \exp \left[\frac{i}{2\hbar} (p'q - q'p) \right. \\ \left. - \frac{1}{4} \left[\frac{(p-p')^2}{\sigma_p^2} + \frac{(q-q')^2}{\sigma_q^2} \right] \right], \end{aligned} \quad (3.24)$$

where $\sigma_p\sigma_q = \hbar/2$. These properties suggest that we may regard the operator

$$P_z = |p, q\rangle \langle p, q| \quad (3.25)$$

as an approximate projection operator affecting approximate simultaneous samplings of position and momentum. The approximate orthogonality property (3.24) means that the labels p and q are coarse-grained momentum and position, having significance only up to the widths σ_p and σ_q , respectively.

If the state of the system is described by a density operator ρ , the probability distribution of approximate position \bar{x} and approximate momentum \bar{k} is therefore

$$p(\bar{k}, \bar{x}) = \text{Tr}(P_z \rho) = \langle \bar{k}, \bar{x} | \rho | \bar{k}, \bar{x} \rangle. \quad (3.26)$$

This probability is normalized in the measure $d\bar{k} d\bar{x} / 2\pi\hbar$. Consider the information of this distribution:

$$I_\rho(K, X) = - \int \frac{d\bar{k} d\bar{x}}{2\pi\hbar} p(\bar{k}, \bar{x}) \ln p(\bar{k}, \bar{x}). \quad (3.27)$$

If $p(\bar{k}, \bar{x})$ were a classical phase space distribution, then (3.27) would be the usual entropy in statistical mechanics. The entropy would be unbounded from below because in classical mechanics, the phase space distribution may be arbitrarily concentrated about a particular region of phase space. In quantum mechanics, by contrast, phase space distributions concentrated on regions smaller in size than \hbar would violate the uncertainty principle. We therefore expect a lower bound on (3.27).

A reasonable guess as to what this lower bound should be is obtained by evaluating (3.27) with ρ a coherent state, since the coherent states are normally thought of as being the states most concentrated in phase space. Writing $\rho = |z\rangle \langle z|$, one finds

$$I_{|z\rangle}(K, X) = 1. \quad (3.28)$$

For these reasons it was conjectured by Wehrl [17] that

$$I_\rho(K, X) \geq 1, \quad (3.29)$$

with equality if and only if ρ is a coherent state. This was subsequently proved by Lieb [18], again using some inequalities from Fourier analysis (best constants in the Hausdorff-Young and Young inequalities).

A simple but important generalization of this result was noted by Grabowski [19]. This is that the inequality (3.29) continues to hold for projections constructed from a class of *generalized* coherent states, namely, those of the form

$$|\psi_{\bar{x}\bar{k}}\rangle = U(\bar{x}, \bar{k}) |\psi\rangle, \quad (3.30)$$

where $|\psi\rangle$ is an arbitrary state. The point is that they share with the usual coherent states the completeness relation (3.23), and it is this property that is exploited in Lieb's proof.

This generalization also permits a connection with the usual uncertainty relation to be made. One has

$$\begin{aligned} I_\rho(K, X) &\leq I_\rho(X) + I_\rho(K) \\ &\leq 1 + \ln \left[\frac{\Delta x \Delta k}{\hbar} \right], \end{aligned} \quad (3.31)$$

where Δx and Δk are the variances of x and k in the probability distribution (3.26), but with the generalized coherent states (3.30). The first inequality is a standard property of information; the second is the inequality (2.6) used twice (up to a factor of $2\pi\hbar$, because of our choice of phase space measure). Together with (3.24), (3.31) implies that

$$\Delta x \Delta k \geq \hbar. \quad (3.32)$$

This is not the usual uncertainty relation (no factor of $\frac{1}{2}$), because the variances express not only the uncertainty in the initial state, but also the uncertainty in the projections, which are imprecise. Indeed, one has

$$(\Delta x)^2 = (\Delta_\rho x)^2 + (\Delta_\psi x)^2, \quad (3.33)$$

$$(\Delta k)^2 = (\Delta_\rho k)^2 + (\Delta_\psi k)^2. \quad (3.34)$$

The first term on the right-hand side of each relation is the variance in the initial state; the second is the variance in the generalized coherent state projection with fiducial state $|\psi\rangle$. Choosing ρ to be the pure state $|\psi\rangle \langle \psi|$, one thus obtains the usual uncertainty relation

$$\Delta_\psi x \Delta_\psi k \geq \frac{\hbar}{2}. \quad (3.35)$$

An alternative method of connecting (3.29) with the usual uncertainty relations may be found in Ref. [8].

Results similar to (3.17) and (3.19) have been obtained by Deutsch [20], Partovi [21], and Maassen and Uffink [22]. Equation (3.17) has been generalized to include thermal fluctuations at thermal equilibrium by Abe and Suzuki [23]. Anderson and Halliwell have generalized (3.29) to include thermal fluctuations in a class of non-

equilibrium systems [8] (see also Ref. [9]). For an alternative approach to unsharp samplings of noncommuting observables using positive-operator-valued measures, see Schroeck [24]. For other related results on information-theoretic uncertainty relations, not directly relevant to the present paper, see Refs. [25–32].

**IV. TWO-TIME HISTORIES:
APPROXIMATE PROJECTORS**

We now show how to obtain information-theoretic uncertainty relations for histories characterized by projections at two moments of time. The projections will be onto position at two moments of time or onto momentum and position. The important feature is that the time-dependent projections $P_{\alpha}^k(t_k)$ do not commute, and so one would not expect their probability distributions to be arbitrarily peaked. We therefore expect to derive lower bounds on the information, in analogy with (3.29).

The case of position and momentum samplings by exact projections, such as Eq. (3.2), is quite different from the case of approximate projections, such as Eq. (3.5), and each case needs to be treated separately. The approximate projection case is a direct extension of the results of Sec. III C, and we consider this case first. The case of exact projections will be treated in the next section.

A. Lower bound on the information

In brief, the idea is as follows. The probabilities for histories are most generally given by an expression of the form

$$p(\underline{\alpha}) = \text{Tr}(C_{\underline{\alpha}}^{\dagger} C_{\underline{\alpha}} \rho), \tag{4.1}$$

where $C_{\underline{\alpha}}$ denotes a string of time-dependent projection operators,

$$C_{\underline{\alpha}} = P_{\alpha_n}^n(t_n) \cdots P_{\alpha_1}^1(t_1),$$

and we use the notation $\underline{\alpha}$ to denote a string of α 's. The burden of the results described below will be to show that for the case of two-time histories considered here, the operator $C_{\underline{\alpha}}^{\dagger} C_{\underline{\alpha}}$ may be written in the form

$$C_{\underline{\alpha}}^{\dagger} C_{\underline{\alpha}} = U(\bar{k}, \bar{x}) \Omega U^{\dagger}(\bar{k}, \bar{x}) \tag{4.2}$$

for some operator Ω . The point is that the dependence on the sampled positions and momenta \bar{x} and \bar{k} resides entirely in the unitary Weyl operator $U(\bar{k}, \bar{x})$. Now $C_{\underline{\alpha}}^{\dagger} C_{\underline{\alpha}}$ is a positive Hermitian operator, and thus Ω is also. It may therefore be written

$$\Omega = \sum_a \lambda_a |a\rangle \langle a|, \tag{4.3}$$

where the coefficients λ_a are positive. The probabilities (4.1) for our two-time histories may now be written

$$p(\alpha, \beta) = \sum_a \lambda_a \langle a | U(\bar{k}, \bar{x}) \rho U^{\dagger}(\bar{k}, \bar{x}) | a \rangle. \tag{4.4}$$

Here we have introduced, as earlier, the continuous bin labels α and β , defined in terms of the sampled positions

and momenta by $\bar{x} = \alpha \sigma_x$, $\bar{k} = \beta \sigma_k$, where σ_x and σ_k are their respective widths. The right-hand side of (4.4) involves the expectation value of ρ in the generalized coherent states, $|a_{\bar{k}\bar{x}}\rangle = U^{\dagger}(\bar{k}, \bar{x}) |a\rangle$. By the convexity property (2.8), the information of (4.4) satisfies the inequality

$$\begin{aligned} I(K, X) &\equiv - \int d\alpha d\beta p(\alpha, \beta) \ln p(\alpha, \beta) \\ &\geq - \int d\bar{x} d\bar{k} \langle a_{\bar{k}\bar{x}} | \rho | a_{\bar{k}\bar{x}} \rangle \ln \langle a_{\bar{k}\bar{x}} | \rho | a_{\bar{k}\bar{x}} \rangle \\ &\quad - \ln(\sigma_x \sigma_k). \end{aligned} \tag{4.5}$$

The factor of $\ln(\sigma_x \sigma_k)$ arises from the change of variables from α, β , to \bar{x}, \bar{k} . From the previous section [Eqs. (3.29) and (3.30)], we thus deduce the inequality

$$I(K, X) \geq 1 + \ln \left[\frac{2\pi\hbar}{\sigma_x \sigma_k} \right]. \tag{4.6}$$

The factor of $2\pi\hbar$ appears because of the difference in phase space measures used in (3.27) and (4.5). The factor of 2 difference between (4.6) and (3.19) is due to the fact that at equality, (3.19) measures the uncertainty in the state alone, whereas (4.6) also includes the uncertainty in the coherent state projector.

Equation (4.6) is an intuitively appealing result. The argument of the logarithm is the inverse of the number of elementary cells of phase space sampled. If that number is large, i.e., $\sigma_x \sigma_k \gg 2\pi\hbar$, then the lower bound approaches $-\infty$, and thus the uncertainty principle imposes little restriction on samplings of phase space large compared to the fundamental cell. On the other hand, the bound becomes significant when $\sigma_x \sigma_k$ is of order $2\pi\hbar$ or smaller, in agreement with the expectation that the uncertainty principle imposes limitations on samplings comparable to the size of the fundamental cell.

Everything up to Eq. (4.5) is also true for the discrete case (with the integral over α, β replaced by a discrete sum), but it is not possible to deduce the inequality (4.6), since this holds only for the continuous case.

We now need to show that the projections satisfy the condition (4.2) for the two-time histories of interest. It is also necessary to calculate Ω to determine the conditions under which the inequality becomes equality. Before that, we need to describe some mathematical tools.

B. Weyl calculus

The analysis of (4.1) is conveniently carried out with the aid of a set of mathematical tools referred to as the Weyl calculus. This in turn is part of a larger area of mathematics called microlocal analysis [33]. The basic idea is to define a one-to-one correspondence between every self-adjoint operator, \hat{A} , say, on the Hilbert space, and a real function $A(p, q)$ defined in a phase space, referred to as the Weyl symbol of \hat{A} . A particular example of how this correspondence may be obtained is through the Wigner transform

$$W_A(p, q) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} d\xi \left\langle q + \frac{\xi}{2} \left| \hat{A} \right| q - \frac{\xi}{2} \right\rangle e^{-ip\xi/\hbar}. \tag{4.7}$$

When the operator \hat{A} is the density operator $\hat{\rho}$, W_ρ is called the Wigner function. It shares many properties of classical phase space distributions, although it is often not positive. It has been used extensively in discussions of the classical limit [1–36]. We shall make use of the Wigner transform (4.7) to analyze (4.1).

An alternative form of (4.7) that we shall find more useful is

$$W_A(p, q) = \text{Tr}[\hat{\Delta}(p, q) \hat{A}], \quad (4.8)$$

where

$$\hat{\Delta}(p, q) = \frac{\hbar}{2\pi} \int du dv e^{ipu + iqv} e^{-iu\hat{P} - iv\hat{Q}}. \quad (4.9)$$

Here \hat{Q} and \hat{P} are the usual position and momentum operators, satisfying $[\hat{Q}, \hat{P}] = i\hbar$.

We record and prove some useful properties of the Wigner transform. First, one has

$$\text{Tr}(\hat{A}\hat{B}) = \frac{1}{2\pi\hbar} \int dp dq W_A(p, q) W_B(p, q). \quad (4.10)$$

This follows readily from inserting the explicit form (4.7) into the right-hand side of (4.10).

Next, we discuss the properties of the Weyl symbol under shifts of its arguments. Introduce the unitary Weyl operator

$$U(P, q) = \exp\left[\frac{i}{\hbar} p \hat{Q} - \frac{i}{\hbar} q \hat{P}\right]. \quad (4.11)$$

It has the properties

$$U^\dagger(\bar{p}, \bar{q}) \hat{Q} U(\bar{p}, \bar{q}) = \hat{Q} + \bar{q}, \quad (4.12)$$

$$U^\dagger(\bar{p}, \bar{q}) \hat{P} U(\bar{p}, \bar{q}) = \hat{P} + \bar{p}. \quad (4.13)$$

The Baker-Campbell-Hausdorff relation is

$$e^{\hat{A} + \hat{B}} = e^{\hat{A}} e^{\hat{B}} e^{[\hat{A}, \hat{B}]/2} \quad (4.14)$$

if $[\hat{A}, \hat{B}]$ commutes with \hat{A} and \hat{B} . It follows that

$$U^\dagger(\bar{p}, \bar{q}) e^{-iu\hat{P} - iv\hat{Q}} U(\bar{p}, \bar{q}) = e^{i\bar{p}u + i\bar{q}v} e^{-iu\hat{P} - iv\hat{Q}} \quad (4.15)$$

and thus

$$U^\dagger(\bar{p}, \bar{q}) \hat{\Delta}(p, q) U(\bar{p}, \bar{q}) = \hat{\Delta}(p + \bar{p}, q + \bar{q}). \quad (4.16)$$

From this we see that

$$\begin{aligned} W_A(p + \bar{p}, q + \bar{q}) &= \text{Tr}[\hat{\Delta}(p + \bar{p}, q + \bar{q}) \hat{A}] \\ &= \text{Tr}[\hat{\Delta}(p, q) \hat{A}'] \\ &= W_{A'}(p, q), \end{aligned} \quad (4.17)$$

where $\hat{A}' = U(\bar{p}, \bar{q}) \hat{A} U^\dagger(\bar{p}, \bar{q})$. That is, translating the coordinates and momenta of the Weyl symbol is equivalent to a unitary transformation under the Weyl operator of the original operator.

From (4.1) and (4.10), it follows that the probabilities for histories characterized by the chain operator C_α are given by

$$p(\underline{\alpha}) = \frac{1}{2\pi\hbar} \int dp dq W_{C^\dagger C}(p, q) W_\rho(p, q). \quad (4.18)$$

C. Position and direct momentum samplings

The first type of history we shall consider is one characterized by an imprecise position sampling at time zero and an imprecise momentum sampling at time t . The probability for this history is given by

$$p(\alpha, \beta, t) = \text{Tr}[P_\beta^k e^{-iHt} P_\alpha^x \rho P_\alpha^x e^{iHt}]. \quad (4.19)$$

In the short time limit, employed here, evolution is described by the free Hamiltonian. This clearly commutes with the momentum projections, and thus t drops out in the short time limit. One thus has

$$C_\alpha^\dagger C_\alpha = P_\alpha^x P_\beta^k P_\alpha^x. \quad (4.20)$$

The Weyl symbol of this operator is

$$W_{C^\dagger C}(p, q) = \frac{1}{2\pi\hbar} \int d\xi \left\langle q + \frac{\xi}{2} \left| P_\alpha^x P_\beta^k P_\alpha^x \right| q - \frac{\xi}{2} \right\rangle e^{-ip\xi/\hbar}. \quad (4.21)$$

Inserting the explicit forms for the projection operators, one obtains

$$\begin{aligned} W_{C^\dagger C}(p, q) &= \frac{1}{2\pi\hbar} \int d\xi dk e^{i\xi(k-p)/\hbar} \Upsilon(q + \frac{1}{2}\xi - \bar{x}) \\ &\quad \times \Upsilon(q - \frac{1}{2}\xi - \bar{x}) \Gamma(k - \bar{k}). \end{aligned} \quad (4.22)$$

Letting $k \rightarrow k + \bar{k}$, it is readily seen that one has

$$W_{C^\dagger C}(p, q) = W_\Omega(p - \bar{k}, q - \bar{x}). \quad (4.23)$$

Here Ω is the operator whose Weyl symbol is (4.22), but with $\bar{k} = 0$ and $\bar{x} = 0$. Ω is therefore equal to $P_\alpha^x P_\beta^k P_\alpha^x$ at $\bar{k} = 0$ and $\bar{x} = 0$, that is,

$$\begin{aligned} \Omega &= \frac{1}{2\pi\hbar} \int dx dy dk \Upsilon(x) \Upsilon(y) \Gamma(k) \\ &\quad \times e^{ik(x-y)/\hbar} |x\rangle \langle y|. \end{aligned} \quad (4.24)$$

From (4.17) and (4.23), we therefore have a result of the general form (4.2):

$$P_\alpha^x P_\beta^k P_\alpha^x = U^\dagger(\bar{k}, \bar{x}) \Omega U(\bar{k}, \bar{x}). \quad (4.25)$$

From the above it therefore follows that the information of the phase space distribution (4.19) obeys the inequality (4.6).

Consider now the conditions for equality. Equality is obtained if and only if both ρ and Ω are of the form $|z\rangle \langle z|$, where $|z\rangle$ is a canonical coherent state [Eq. (3.20)]. From (4.24) one can see that Ω will be of that form, if and only if $\Gamma(k) = \delta(k)$ and $\Upsilon(x)$ is a Gaussian. That is, the first projection is a Gaussian projection onto position, and the second is an infinitely precise sampling of momentum.

D. Position and time-of-flight momentum samplings

We now consider a history characterized by imprecise position samplings at times 0 and t . The probability is

$$p(\alpha_1, \alpha_2, t) = \text{Tr}[P_{\alpha_2}^x e^{-iHt} P_{\alpha_1}^x \rho P_{\alpha_1}^x e^{iHt}]. \quad (4.26)$$

From this one can construct a phase space probability $p(\alpha_1, \beta, t)$, where $\beta\sigma_k = \bar{k}$ and $\bar{k} = m(\bar{x}_2 - \bar{x}_1)/t$, for small t . We have $\bar{x}_1 = \sigma_x \alpha_1$, $\bar{x}_2 = \sigma_x \alpha_2$; thus, $\beta = \alpha_2 - \alpha_1$ and $\sigma_k = m\sigma_x/t$. One has

$$C_{\alpha}^{\dagger} C_{\alpha} = P_{\alpha_1}^x e^{iHt} P_{\alpha_2}^x e^{-iHt} P_{\alpha_1}^x. \quad (4.27)$$

We will analyze this case for small times t . It is readily shown that the Weyl symbol is

$$W_{C^{\dagger}C}(p, q) = \frac{1}{2\pi\hbar} \int d\xi dx_2 e^{-ip\xi/\hbar} \Upsilon(q + \frac{1}{2}\xi - \bar{x}_1) \Upsilon(q - \frac{1}{2}\xi - \bar{x}_1) \Upsilon(x_2 - \bar{x}_2) \langle x_2, t | q - \frac{1}{2}\xi, 0 \rangle \langle x_2, t | q + \frac{1}{2}\xi, 0 \rangle^*. \quad (4.28)$$

Now, in the short time limit, the propagator is given by

$$\langle x_2, t | q - \frac{1}{2}\xi, 0 \rangle = \left[\frac{m}{2\pi\hbar it} \right]^{1/2} \exp \left[\frac{im}{2\hbar t} (x_2 - q + \frac{1}{2}\xi)^2 \right]. \quad (4.29)$$

Inserting this in (4.28) and performing the shift $x_2 \rightarrow x_2 + \bar{x}_2$, one finds that the answer may be written in the form

$$W_{C^{\dagger}C}(p, q) = \frac{1}{2\pi\hbar} \int d\xi dx_2 \Upsilon(q + \frac{1}{2}\xi - \bar{x}_1) \Upsilon(q - \frac{1}{2}\xi - \bar{x}_1) \Upsilon(x_2) \frac{m}{2\pi\hbar t} \times \exp \left[-\frac{i}{\hbar} \left[p - \frac{m}{t}(\bar{x}_2 - \bar{x}_1) \right] \xi + i \frac{m}{\hbar t} (x_2 - q + \bar{x}_1) \xi \right]. \quad (4.30)$$

One therefore has

$$W_{C^{\dagger}C}(p, q) = W_{\Omega}(p - \bar{k}, q - \bar{x}_1). \quad (4.31)$$

Ω is the operator whose Wigner transform is (4.30) with $\bar{x}_1 = 0$ and $\bar{x}_2 = 0$. Explicitly,

$$\Omega = \int dx dy dx_2 \langle x_2, t | y, 0 \rangle \langle x_2, t | x, 0 \rangle^* \Upsilon(x) \times \Upsilon(y) \Upsilon(x_2) |x\rangle \langle y|. \quad (4.32)$$

From (4.17) and (4.31), we now have the result

$$P_{\alpha_1}^x e^{iHt} P_{\alpha_2}^x e^{-iHt} P_{\alpha_1}^x = U^{\dagger}(\bar{k}, \bar{x}) \Omega U(\bar{k}, \bar{x}). \quad (4.33)$$

We therefore again have the inequality (4.6) for the information of the phase space distribution constructed from (4.26).

Consider the conditions for equality. Again, this is achieved when both ρ and Ω are of the form $|z\rangle \langle z|$. This means that Eq. (4.30) must be the Wigner transform of a coherent state, i.e., a product of Gaussians in p and q . This can be achieved by letting the width of the sampling function at t go to a δ function, setting the sampling function at $t=0$ to a Gaussian, and then letting $t \rightarrow \infty$. (This may be seen explicitly in Ref. [36].) We are, however, working in the short time approximation, and so this procedure can be carried out only for the free particle case, for which the short time approximation is exact.

It is also possible to deduce a lower bound on the information of the joint probability for position samplings [Eq. (4.26)]. The information of $p(\alpha_1, \alpha_2, t)$ is in fact equal to that of $p(\alpha_1, \beta, t)$, because the Jacobean of the transformation between these variables is unity. One

thus has the following bound on the information of (4.26):

$$I(X_1, X_2) \equiv - \int d\alpha_1 d\alpha_2 p(\alpha_1, \alpha_2) \ln p(\alpha_1, \alpha_2) \geq 1 + \ln \left[\frac{2\pi\hbar t}{m\sigma_x^2} \right]. \quad (4.34)$$

This is, strictly speaking, a trivial rewriting of (4.6). We record the result because it will be generalized to an arbitrary number of position samplings in Sec. VI.

E. Position samplings at arbitrary time separations

The previous case concerned position samplings for any Hamiltonian, but in the limit of small time separations. For the case of linear systems, we may extend this analysis to arbitrary time separations. We now outline how this is done.

The propagator for linear systems is given by

$$\langle x'', t'' | x', t' \rangle = \Delta(t'', t') \exp \left[\frac{i}{\hbar} S(x'', t'' | x', t') \right], \quad (4.35)$$

where S is the action of the classical solution connecting initial and final points, and is quadratic in the x 's. The prefactor Δ is independent of the x 's and is given by

$$\Delta(t'', t') = \left[-\frac{1}{2\pi i \hbar} \frac{\partial^2 S(x'', t'' | x', t')}{\partial x'' \partial x'} \right]^{1/2}. \quad (4.36)$$

Repeating the analysis of the previous subsection, Eq. (4.28) thus has the form

$$W_{C^{\dagger}C}(p, q) = \frac{1}{2\pi\hbar} \int d\xi dx_2 \Upsilon(q + \frac{1}{2}\xi - \bar{x}_1) \Upsilon(q - \frac{1}{2}\xi - \bar{x}_1) \Upsilon(x_2 - \bar{x}_2) |\Delta|^2 \times \exp \left[-\frac{i}{\hbar} p\xi + \frac{i}{\hbar} S(x_2, t | q - \frac{1}{2}\xi, 0) - \frac{i}{\hbar} S(x_2, t | q + \frac{1}{2}\xi, 0) \right]. \quad (4.37)$$

Now, letting $x_2 \rightarrow x_2 + \bar{x}_2$ and using the fact that S is quadratic, (4.37) may be written

$$W_{C^\dagger C}(p, q) = \frac{1}{2\pi\hbar} \int d\xi dx_2 \Upsilon(q + \frac{1}{2}\xi - \bar{x}_1) \Upsilon(q - \frac{1}{2}\xi - \bar{x}_1) \Upsilon(x_2) |\Delta|^2 \exp \left[-\frac{i}{\hbar} \xi(p - \bar{k}) - \frac{i}{\hbar} \xi \frac{\partial S}{\partial q}(x_2, t | q - \bar{x}_1, 0) \right], \quad (4.38)$$

where we have introduced

$$\bar{k} = -\frac{\partial S}{\partial \bar{x}_1}(\bar{x}_2, t | \bar{x}_1, 0). \quad (4.39)$$

From Hamilton-Jacobi theory, \bar{k} is the initial momentum for the classical path between \bar{x}_1 and \bar{x}_2 . Now the point is that (4.38) depends on \bar{x}_2 and \bar{x}_1 only through the combinations $p - \bar{k}$ and $q - \bar{x}_1$, and we again have a result of the form (4.33), but this time with \bar{k} given by (4.39). We therefore again deduce the inequality (4.6) for the information of the corresponding phase space distribution.

What is perhaps more interesting in this case is to derive the generalization of (4.34). Since \bar{k} is linear in \bar{x}_2, \bar{x}_1 in (4.39), we have

$$\bar{k} = \frac{\partial \bar{k}}{\partial \bar{x}_2} \bar{x}_2 + \frac{\partial \bar{k}}{\partial \bar{x}_1} \bar{x}_1 \quad (4.40)$$

and thus

$$\beta = \frac{\sigma_x}{\sigma_k} \left[\frac{\partial \bar{k}}{\partial \bar{x}_2} \alpha_2 + \frac{\partial \bar{k}}{\partial \bar{x}_1} \alpha_1 \right]. \quad (4.41)$$

Here, as before, $\beta \sigma_k = \bar{k}$, $\alpha_1 \sigma_x = \bar{x}_1$, and $\alpha_2 \sigma_x = \bar{x}_2$. Unlike the case of short times, the transformation from α_1, β to α_1, α_2 has a nontrivial Jacobian. It follows that

$$I(X_1, X_2) = I(K, X) - \ln \left[\frac{\sigma_x}{\sigma_k} \left| \frac{\partial \bar{k}}{\partial \bar{x}_2} \right| \right]. \quad (4.42)$$

Finally, using the bound (4.6) on $I(K, X)$ and noting that

$$\frac{\partial \bar{k}}{\partial \bar{x}_2} = -\frac{\partial^2 S}{\partial \bar{x}_1 \partial \bar{x}_2}(\bar{x}_2, t | \bar{x}_1, 0), \quad (4.43)$$

we derive the following bound on the information of position samplings at arbitrary time separations:

$$I(X_1, X_2) \geq 1 + \ln \left[\frac{2\pi\hbar}{\sigma_x^2} \left| \frac{\partial^2 S}{\partial \bar{x}_1 \partial \bar{x}_2} \right|^{-1} \right]. \quad (4.44)$$

We will generalize this result and discuss it further in Sec. VI.

V. TWO-TIME HISTORIES: EXACT PROJECTORS

As stated in the previous section, the case of exact projections is rather different to the case of approximate ones and needs to be treated separately. In this section we show how this is done.

We are again interested in an expression for the probability of a two-time history of the form (4.1), where $C_\alpha^\dagger C_\alpha$ is of the form (4.20) or (4.27). In each case it is again possible to show that $C_\alpha^\dagger C_\alpha$ may be written in the form (4.2), although note that now $\bar{x}_\alpha, \bar{k}_\beta$ are discrete rather than

continuous variables. We can go on to use steps (4.3)–(4.5), except that the integral in (4.5) becomes a discrete sum, and it is at this point that we can go no further. Of course, if the bin sizes σ_x, σ_k are very small, then the discrete sum may be approximated by the continuous integral (4.5), and we deduce the inequality (4.6). But more generally a different method is needed.

Very generally, probabilities for histories are given by an expression of the form (4.1). If the projections contained in the chain operators C_α are exact projections and either fine-grained projections onto discrete variables (e.g., spins) or coarse-grained projections onto continuous variables [e.g., as in Eq. (3.2)], then the variables α labeling the alternatives form a discrete set, and so there is a discrete (although possibly infinite) set of probabilities $p(\alpha)$. This means that they possess an upper bound $p(\alpha) \leq p_{\max} \leq 1$, and a lower bound on the information follows trivially:

$$I = -\sum_\alpha p(\alpha) \ln p(\alpha) \geq \ln \left[\frac{1}{p_{\max}} \right]. \quad (5.1)$$

[Note that this is not true of the information of continuous variables. There, the $p(\alpha)$'s are not probabilities, but probability densities, and so need not be bounded from above.] The upper bound p_{\max} may be computed by studying the spectrum of the operator $C_\alpha^\dagger C_\alpha$. In particular, the bound (5.1) will be nontrivial, i.e., $p_{\max} < 1$, if at least one pair of the time-dependent projections $P_{\alpha_k}^k(t_k)$ do not commute [37].

Now consider the case of two-time histories. As stated, everything in Sec. IV from (4.1) to (4.5) also holds in the case of exact projections. Suppose we obtain the spectrum of the operator Ω [Eq. (4.3)] and we look for the largest eigenvalue λ_{\max} ; thus, $\lambda_\alpha \leq \lambda_{\max}$. It follows from (4.4) that

$$p(\alpha, \beta) \leq \lambda_{\max} \sum_\alpha \langle a | U(\bar{k}, \bar{x}) \rho U^\dagger(\bar{k}, \bar{x}) | a \rangle = \lambda_{\max} \quad (5.2)$$

and thus $p_{\max} = \lambda_{\max}$. Position and momentum projections at the same time or position projections at different times do not commute; thus, the bound (5.1) will be nontrivial.

A. Position and direct momentum samplings

Consider first the case of a position followed by momentum sampling; thus, we have (4.20), but with exact projections. We again deduce (4.25), and so we are interested in the spectrum of the operator Ω , given by (4.24). Now write

$$\Omega |u\rangle = \lambda |u\rangle. \quad (5.3)$$

Inserting the explicit form of Ω and performing the k in-

tegration, one obtains the eigenvalue equation

$$\theta(\xi)\theta(1-\xi)\int_0^1 d\xi' \frac{\sin[\pi U(\xi-\xi')]}{\pi(\xi-\xi')} w(\xi') = \lambda w(\xi), \quad (5.4)$$

where $U = \sigma_k \sigma_x / 2\pi\hbar$, $x = \sigma_x(\xi - \frac{1}{2})$, and $w(\xi) = \langle x|u \rangle$. Apart from θ functions on the left-hand side, this equation is identical to an eigenvalue equation written down by Partovi in his study of the analogous question for the case of samplings of two ensembles, as in Sec. III B [21]. It is not clear whether it can be solved exactly, but it is straightforward to extract the relevant information in the regimes of interest. For $U \ll 1$ the kernel on the left-hand side is approximately equal to U . The spectrum is degenerate with $\lambda \approx U$, and $w(\xi)$ a constant on the interval $[0,1]$ and zero elsewhere. For $U \gg 1$ the kernel becomes a delta function $\delta(\xi - \xi')$. The eigenvalue equation is then satisfied by any function with support only in the interval $[0,1]$ (up to normalization), and the spectrum is again degenerate with $\lambda \approx 1$. The following bound on the information is thus obtained:

$$I(K, X) \geq I_{\min} \approx \begin{cases} 0 & \text{if } \sigma_k \sigma_x \gg 2\pi\hbar, \\ \ln \left[\frac{2\pi\hbar}{\sigma_k \sigma_x} \right] & \text{if } \sigma_k \sigma_x \ll 2\pi\hbar. \end{cases} \quad (5.5)$$

As in the continuous case [Eq. (4.6)], the result is intuitively appealing. The lower bound is nontrivial for probes of phase space comparable to or smaller than the fundamental cell. On the other hand, there is no restriction when the probe is much larger than the fundamental cell, and the lower bound is essentially zero. (It is not $-\infty$, as in the continuous case, because information is non-negative for discrete distributions.)

Note that the bounds (5.5) and (4.6) approximately coincide for the case $\sigma_k \sigma_x \ll 2\pi\hbar$. This is to be expected since, as stated above, this is the condition that the discrete and continuous versions of (4.5) coincide.

B. Position and time-of-flight momentum samplings

In the case of time-of-flight momentum samplings, we study (4.26) with exact position projections. We again have (4.33), and we thus need to find the largest eigenvalue of the operator Ω , in this case given by (4.32). It is straightforward to show that the eigenvalue equation is

$$\theta(\xi)\theta(1-\xi)\int_0^1 d\xi' \exp[-i\pi U(\xi-\xi')(\xi+\xi'+1)] \times \frac{\sin[\pi U(\xi-\xi')]}{\pi(\xi-\xi')} w(\xi') = \lambda w(\xi), \quad (5.6)$$

where the various quantities are all the same as in (5.4), recalling that $\sigma_k = m\sigma_x/t$, as in Sec. IV C. It is not difficult to see that the presence of the exponential factor in (5.6) in comparison to (5.4) actually makes no difference to the leading order asymptotic solutions in the regions $U \gg 1$ and $U \ll 1$. We thus once again obtain the result (5.5).

As in Eq. (4.34), one can again use this result to obtain a bound on the information of the joint probability of position samplings. In this case it is

$$I(X_1, X_2) \geq I_{\min} \approx \ln \left[\frac{2\pi\hbar t}{m\sigma_x^2} \right] \quad (5.7)$$

in the regime $m\sigma_x^2 \ll 2\pi\hbar t$. Similarly, we expect to be able to derive a result of the form (4.42), for linear systems, in the exact projections case, although we do not describe this in detail.

VI. GENERAL HISTORIES

We have studied the uncertainty principle for histories characterized by position and momentum projections at two moments of time. We now go on to study the more general case of histories characterized by position projections at an arbitrary number of times [38]. On general grounds and inspired by specific calculations [5], we expect the probability for a sequence of position samplings to be in some sense peaked about sets of solutions to the classical field equations, with a weight depending on the initial state. The precise sense in which this is true is discussed in another paper [39] (see also Ref. [40]). One expects the uncertainty principle to impose a limitation on the degree of peaking. Here we derive an information-theoretic inequality expressing this limitation for histories characterized by an arbitrary number of position samplings. This is a generalization of the results (4.34), (4.42), and (5.7). We then obtain the form of the uncertainty principle for histories characterized by other types of projections.

As in Sec. V, if some of the projections in the chain operators C_{α} do not commute, then the spectrum of the operator $C_{\alpha} C_{\alpha}^{\dagger}$ is strictly less than 1 and likewise the probabilities $p(\alpha)$. A lower bound on the information of the form (5.1) is thus obtained. Let us apply this rationale to strings of imprecise position projections, with sampling functions of the form (3.2). Our aim is to obtain a lower bound on the information

$$I(X_1, \dots, X_n) = - \sum_{\alpha_1} \dots \sum_{\alpha_n} p(\alpha_1, \dots, \alpha_n) \times \ln p(\alpha_1, \dots, \alpha_n). \quad (6.1)$$

The expression (1.3) for the probabilities may be written

$$p(\alpha) = \int dx_0 dy_0 \langle y_0 | C_{\alpha}^{\dagger} C_{\alpha} | x_0 \rangle \rho(x_0, y_0), \quad (6.2)$$

where

$$\begin{aligned} & \langle y_0 | C_{\alpha}^{\dagger} C_{\alpha} | x_0 \rangle \\ &= \int \prod_{k=1}^n dx_k dy_k \delta(x_n - y_n) \Upsilon(x_k - \bar{x}_k) \Upsilon(y_k - \bar{y}_k) \\ & \quad \times \prod_{k=1}^n J(x_k, y_k, t_k | x_{k-1}, y_{k-1}, t_{k-1}). \end{aligned} \quad (6.3)$$

Here, as in previous sections, $\bar{x}_k = \sigma \alpha_k$. The samplings functions Υ are given by Eq. (3.2). J is the density matrix propagator, which for unitary evolution is given by

$$\begin{aligned} & J(x'', y'', t'' | x', y', t') \\ &= \langle x'', t'' | x', t' \rangle \langle y'', t'' | y', t' \rangle^*. \end{aligned} \quad (6.4)$$

We shall work in the limit that the time separation between each projection is small. The propagators in (6.4) are then given by (4.35) and (4.36). This is exact for linear systems. The case in which J is a nonunitary reduced density matrix propagator is also of interest in the context of decoherence models (see Refs. [5,41], for example). However, such propagators reduce to the unitary expression (6.4) in the short time limit; hence, our results are applicable to that case also.

For simplicity, we study first the free particle case, for which one has

$$\Delta(t'', t') = \left[\frac{m}{2\pi i \hbar (t'' - t')} \right]^{1/2} \quad (6.5)$$

and

$$S(x'', t'' | x', t') = \frac{m(x'' - x')^2}{2(t'' - t')} . \quad (6.6)$$

Also, let all of the projections have the same width σ , and let the time separation between all slits be t (except for t_1 and t_0 ; see below).

We wish to estimate the largest eigenvalue of the operator $C_{\alpha}^{\dagger} C_{\alpha}$. The eigenvalue equation is

$$\int dx_0 \langle y_0 | C_{\alpha}^{\dagger} C_{\alpha} | x_0 \rangle u(x_0) = \lambda u(y_0) . \quad (6.7)$$

The expression (6.3) occurring in (6.7) has the form of a discrete version of a sum over histories. It may be regarded as a sum over pairs of paths, starting at x_0 and y_0 , passing through gates of width σ at times t_1, \dots, t_n , and meeting in the final gate at point x_n , which is integrated over the width σ . We may approximately evaluate (6.3), and hence solve the eigenvalue equation, by looking for the paths which dominate the integral in the regimes of interest.

We follow a heuristic argument previously used by Mensky in a related context [40]. There are two competing effects that will determine which paths dominate. On the one hand, if the slit widths in the projections are very small, this will force the paths to follow the set of alternatives \bar{x}_k specified by the projections. On the other hand, if the action of each path (i.e., the sum of the phases of the propagators) is very large, $S \gg \hbar$, then by the stationary phase approximation, we expect the dominant paths to be those extremizing the action, i.e., classical paths.

Consider first the case in which the slit widths are very small. In this case the paths are forced to follow the sampling positions \bar{x}_k . The action S of each path is of order $m\sigma^2/t$. We therefore take “ σ small” to mean that $S \ll \hbar$. This implies that the exponential part of the propagators in (6.3) is negligible, and only the prefactors contribute. We may therefore approximately evaluate the integral (6.3), with the result

$$\begin{aligned} \langle y_0 | C_{\alpha}^{\dagger} C_{\alpha} | x_0 \rangle &\approx \sigma(\sigma^2)^{n-2} \left[\frac{m}{2\pi\hbar t} \right]^{n-1} \\ &\times \int dx_1 dy_1 \Upsilon(x_1 - \bar{x}_1) \Upsilon(y_1 - \bar{x}_1) \\ &\times J(x_1, y_1, t_1 | x_0, y_0, t_0) . \end{aligned} \quad (6.8)$$

The origin of each part of this expression is as follows: The factor $(m/2\pi\hbar t)^{n-1}$ comes from the $(n-1)$ propagators J ; the factor $(\sigma^2)^{n-2}$ comes from the integrations over x and y at times t_2 to t_{n-1} , noting that J is approximately constant and recalling Eq. (3.4); the factor of σ comes from the final integration over x_n . The remaining integrations over J in (6.8) arise due to the fact that the density matrix in (6.2) is at the initial time t_0 , and not at the time t_1 at which the first projection is made. This is merely a notational inconvenience—the very last part of the chain operators C_{α} is an evolution operator from t_0 to t_1 . It is readily removed by letting $t_1 \rightarrow t_0$; thus, J becomes a product of δ functions and (6.8) becomes

$$\begin{aligned} \langle y_0 | C_{\alpha}^{\dagger} C_{\alpha} | x_0 \rangle \\ \approx \sigma^{-1} \left[\frac{m\sigma^2}{2\pi\hbar t} \right]^{n-1} \Upsilon(x_0 - \bar{x}_1) \Upsilon(y_0 - \bar{x}_1) . \end{aligned} \quad (6.9)$$

Inserting this in the eigenvalue equation (6.7), one thus finds that the spectrum is degenerate, with

$$\lambda \approx \left[\frac{m\sigma^2}{2\pi\hbar t} \right]^{n-1} . \quad (6.10)$$

The eigenfunctions are functions constant in an interval of size σ and zero elsewhere.

Next, let the slit widths be very large. The action of each section of path is then allowed to be large, and it is the stationary phase effect that will dominate. The dominant contribution to the sum over histories will therefore come from the immediate vicinity of the classical paths. When the sampling positions are chosen to line up according to the classical path, it is as if the projections are not there, since most of the integral comes from this regime anyway. It follows that

$$\langle y_0 | C_{\alpha}^{\dagger} C_{\alpha} | x_0 \rangle \approx \delta(x_0 - y_0) , \quad (6.11)$$

and we thus find that $\lambda_{\max} \approx 1$.

Combining these two cases, we thus obtain the following for the lower bound on the information of a sequence of position samplings:

$$\begin{aligned} I(X_1, X_2, \dots, X_n) \\ \geq I_{\min} \approx \begin{cases} 0 & \text{if } m\sigma^2 \gg \hbar t , \\ (n-1) \ln \left[\frac{2\pi\hbar t}{m\sigma^2} \right] & \text{if } m\sigma^2 \ll \hbar t . \end{cases} \end{aligned} \quad (6.12)$$

This lower bound is what one might intuitively expect. First of all, large σ is essentially the classical regime, in which we do not expect to suffer limitations on our ability to describe a history; hence, there is no restriction on the information. Second, the case of small σ is essentially (4.34) generalized to an arbitrary number of samplings. We might expect it because when σ is small the projectors are almost fine grained. They “pinch off” the probability (6.2)—it becomes approximately equal to a product of probabilities for two-time histories of the type discussed in Secs. IV and V. Indeed, the bound in (6.12) is just a sum of bounds of the type (4.34). We will see this in more detail below.

Generalizations of (6.12) may be obtained. The above analysis is readily generalized to the case in which the slits widths σ_j and the time separations $(t_{j+1}-t_j)$ are different, and the short time projector is given by the more general expression (4.35). It is then straightforward to show that the lower bound in (6.12) is, in the small σ_j regime,

$$I_{\min} \approx - \sum_{j=1}^{n-1} \ln(\sigma_{j+1}\sigma_j |\Delta(t_{j+1}, t_j)|^2) \quad (6.13)$$

(and again $I_{\min} \approx 0$ in the large σ_j regime). Equation (6.13) is the leading order behavior of I_{\min} for small σ_j and for small time separations. For linear systems it is valid for arbitrary time separations. How are we to understand this expression?

For the phase space samplings considered earlier, the significance of the lower bounds (4.6) and (5.5) is intuitively clear: The argument of the logarithm is the ratio of the fundamental phase space volume $2\pi\hbar$ to the sampling volume $\sigma_x\sigma_k$.

The lower bound (6.13) has a rather different form; yet an analogous interpretation suggests itself. The propagator prefactor $|\Delta(t_{j+1}, t_j)|^2$ has the dimension of $(\text{length})^{-2}$ and is commonly regarded as the ‘‘density of paths.’’ Introduce the quantity

$$V_H = \prod_{j=1}^{n-1} |\Delta(t_{j+1}, t_j)|^{-2} \quad (6.14)$$

for $n=2, 3, \dots$. For the case of position samplings, it has the dimension $(\text{length})^{2n-2}$. It might therefore reasonably be regarded as the fundamental ‘‘history space

volume.’’ Equation (6.13) may then be written in the suggestive form

$$I_{\min} \approx \ln \left[\frac{V_H}{\sigma_1\sigma_2^2 \cdots \sigma_{n-1}^2\sigma_n} \right]. \quad (6.15)$$

Equation (6.15) now has exactly the same structure as the information-theoretic bounds (4.6) and (5.5) on the phase space samplings considered earlier: The argument of the logarithm in (6.15) is the ratio of the fundamental history space volume to the sampling volume.

It is natural to ask how the results of this section might be further generalized to histories characterized by samplings of variables other than position. It is actually not difficult to see that the above results generalize to histories characterized by samplings of any continuous quantity, such as momentum, angular momentum, etc. Let $P_{\bar{\alpha}}$ be an imprecise sampling of some continuous quantity α :

$$P_{\bar{\alpha}} = \int_{\sigma} d\alpha |\alpha\rangle \langle \alpha|. \quad (6.16)$$

The projections partition the variable α into bins of size σ labeled by $\bar{\alpha}$. We may take the projections to be onto different variables at each moment of time. In the limit of small widths, it is not difficult to see that the analysis for the position samplings case described above readily goes over to the case of arbitrary continuous variables α_k . Essentially what happens is that in the small σ_j limit the matrix elements of the operator $C_{\bar{\alpha}}^{\dagger} C_{\bar{\alpha}}$ become products of propagators and slit widths, in analogy with Eq. (6.8). More precisely,

$$\langle \alpha'_0 | C_{\bar{\alpha}}^{\dagger} C_{\bar{\alpha}} | \alpha_0 \rangle \approx \int_{\sigma_1} d\alpha_1 \int_{\sigma_1} d\alpha'_1 \prod_{j=1}^{n-1} \sigma_{j+1} |\langle \alpha_{j-1}, t_{j+1} | \alpha_j, t_j \rangle|^2 \sigma_j \frac{1}{\sigma_1} \langle \alpha_1, t_1 | \alpha_0, t_0 \rangle \langle \alpha'_1, t_1 | \alpha'_0, t_0 \rangle^* . \quad (6.17)$$

We therefore again deduce the lower bound on the uncertainty (6.15) for this much more general class of histories. The factors $\Delta(t_{j+1}, t_j)$ in (6.14) are now identified with the short time limit of the propagators $\langle \alpha_{j+1}, t_{j+1} | \alpha_j, t_j \rangle$ (maximized over the alternatives α_{j+1}, α_j , in the event that the propagator depends on them in the short time limit). We may thus write

$$I(A_1, A_2, \dots, A_n) \geq I_{\min} \approx \ln \left[\frac{V_H}{\sigma_1\sigma_2^2 \cdots \sigma_{n-1}^2\sigma_n} \right] \quad (6.18)$$

in the small σ_j regime. Here A_1, A_2, \dots, A_n denotes a string of alternatives which can be any continuous variables and may be different variables at different times.

Let us test this more general result with a simple case. Consider a history characterized by a position projection at time t_1 and a momentum projection at time t_2 . Thus $\sigma_1 = \sigma_x$ and $\sigma_2 = \sigma_k$. The short time propagator is

$$\langle p, t_2 | x, t_1 \rangle \approx \frac{1}{(2\pi\hbar)^{1/2}} \exp \left[-\frac{ip^2(t_2-t_1)}{2m} - ipx \right]. \quad (6.19)$$

The history space volume is therefore $V_H = |\Delta(t_2, t_1)|^{-2} = 2\pi\hbar$. The history space volume element is not just analogous to the factor of $2\pi\hbar$ for phase space samplings: It is equal to it in this case. Moreover, the general result (6.18) coincides exactly with the expected result (5.5) for phase space samplings.

Equation (6.18) is the main result of this paper: a concise and very general expression of the uncertainty principle, expressed in the language of quantum-mechanical histories, not referring in any way to phase space, but reducing to the phase space form in the appropriate circumstances.

The expression of the uncertainty principle (6.18) refers to a fundamental history space volume V_H . It is obtained in (6.14) from the short time behavior of the propagator and is thus uniquely determined given the unitary evolution operator $e^{-iHt/\hbar}$. That this operator should appear in the statement of the uncertainty principle for histories should come as no surprise. Unlike phase space statements, the description of a history depends on both the projection operators at each moment of time *and* the unitary evolution between them.

Of course, we have not defined the ‘‘history space’’ of which V_H is the volume element. We shall not pursue

this question here, except to note that it appears to be related to the Cartesian product space $s_1 \times s_2 \times \cdots \times s_n$, where s_j is the spectrum of the observable projected at time t_j . This has been discussed by Omnès [42]. It is perhaps interesting to note that the existence and relevance of such a space is indicated by the form of the uncertainty relation (6.18).

VII. DISCUSSION

In this paper we addressed a simple question: How is the uncertainty principle encoded in the probabilities for histories [Eq. (1.3)]? A simple but very general answer is offered: It arises as the lower bound on the Shannon information [Eq. (6.18)].

We have stressed the generality of the lower bound (6.18) within the framework of standard quantum mechanics (or, at least, its modest generalization to histories). Yet the information-theoretic approach employed here has a potentially greater degree of generality. Information as a measure of uncertainty depends solely on the probabilities for histories. This is in contrast with the usual variance form of the uncertainty principle [Eq. (1.6)] which depends on the wave function of the system at a fixed moment of time. The generality of the information-theoretic form suggests that it might survive to broader forms of quantum mechanics, such as the gen-

eralized quantum mechanics suggested by Hartle [2], which attempts to get away from the Hilbert space formulation. For even if a formulation of quantum mechanics does not deal with wave functions, it must deal with probabilities: Information-theoretic measures may therefore exist where Hilbert-space-dependent measures do not.

To be more precise, we conjecture that the uncertainty principle will most generally arise as a lower bound on the information of the form (1.8), even in generalized formulations of quantum mechanics in which a statement in terms of variances is not available. A stronger conjecture is that the general form of the lower bound (6.18) will also survive such generalizations. These are, however, difficult issues to address in the absence of a concrete generalization of quantum mechanics. They will be taken up elsewhere.

ACKNOWLEDGMENTS

I am very grateful to numerous colleagues for useful conversations, including Arlen Anderson, Carl Caves, Murray Gell-Mann, Jim Hartle, Salman Habib, Bei-Lok Hu, Chris Isham, Raymond Laflamme, Seth Lloyd, Warner Miller, Roland Omnès, Juan Pablo Paz, and Wojciech Zurek. This work was supported by the Royal Society.

-
- [1] M. Gell-Mann and J. B. Hartle, in *Complexity, Entropy and the Physics of Information*, edited by W. Zurek, SFI Studies in the Sciences of Complexity Vol. VIII (Addison-Wesley, Reading, MA, 1990); Phys. Rev. D **47**, 3345 (1993).
 - [2] J. B. Hartle, in *Quantum Cosmology and Baby Universes*: Proceedings of the Jerusalem Winter School as Theoretical Physics, Jerusalem, Israel, 1990, edited by S. Coleman, J. B. Hartle, T. Piran, and S. Weinberg (World Scientific, Singapore, 1991); in *Gravitation and Quantizations*, Proceedings of the Les Houches Summer School, Les Houches, France, 1992, edited by B. Julia and J. Zinn-Justin, Les Houches Summer School Proceedings Vol. LVII (North-Holland, Amsterdam, 1993).
 - [3] R. Griffiths, J. Stat. Phys. **36**, 219 (1984).
 - [4] Y. Aharonov, P. Bergmann, and J. Lebowitz, Phys. Rev. **134**, B1410 (1964).
 - [5] H. F. Dowker and J. J. Halliwell, following paper, Phys. Rev. D **46**, 1580 (1992).
 - [6] See, for example, E. P. Wigner, in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University Press, Princeton, NJ, 1983).
 - [7] Y. Aharonov, P. Bergmann, and J. Lebowitz, Phys. Rev. **134**, B1410 (1964).
 - [8] A. Anderson and J. J. Halliwell, following paper, Phys. Rev. D **48**, 2753 (1993).
 - [9] J. J. Halliwell (unpublished).
 - [10] C. E. Shannon and W. W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, IL, 1949).
 - [11] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
 - [12] H. Everett, in *The Many-Worlds Interpretation of Quantum Mechanics*, edited by B. S. DeWitt and N. Graham (Princeton University Press, Princeton, NJ, 1973).
 - [13] W. Beckner, Ann. Math. **102**, 159 (1975).
 - [14] I. Bialynicki-Birula and J. Mycielski, Commun. Math. Phys. **44**, 129 (1975).
 - [15] I. I. Hirschmann, Am. J. Math. **79**, 152 (1957).
 - [16] J. R. Klauder and E. C. G. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968); J. R. Klauder and B. S. Skagerstam, *Coherent States* (World Scientific, Singapore, 1985).
 - [17] A. Wehrl, Rep. Math. Phys. **16**, 353 (1979).
 - [18] E. H. Lieb, Commun. Math. Phys. **62**, 35 (1978).
 - [19] M. Grabowski, Rep. Math. Phys. **20**, 153 (1984).
 - [20] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).
 - [21] M. H. Partovi, Phys. Rev. Lett. **50**, 1883 (1983).
 - [22] H. Maassen and J. B. M. Uffink, Phys. Rev. Lett. **60**, 1103 (1988).
 - [23] S. Abe and N. Suzuki, Phys. Rev. A **41**, 4608 (1990).
 - [24] F. E. Schroeck, J. Math. Phys. **30**, 2078 (1989).
 - [25] I. Bialynicki-Birula, Phys. Lett. **103A**, 253 (1984).
 - [26] P. Busch and P. J. Lahti, J. Phys. A **20**, 899 (1987).
 - [27] M. Grabowski, Phys. Lett. A **124**, 19 (1987).
 - [28] B. Mamojka, Int. J. Theor. Phys. **11**, 73 (1974).
 - [29] R. Blankenbecler and M. H. Partovi, Phys. Rev. Lett. **54**, 373 (1985).
 - [30] A. K. Rajagopal and S. Teitler, Phys. Lett. A **115**, 313 (1986).
 - [31] J. Sánchez, Phys. Lett. A **173**, 270 (1993).
 - [32] A. Dembo, T. M. Cover, and J. A. Thomas, IEEE Trans. Inf. Theory **37**, 1501 (1991).
 - [33] See, for example, R. Omnès, J. Stat. Phys. **57**, 357 (1989),

- for a review of microlocal analysis as it affects expressions of the form (1.3). For earlier work, see L. Hörmander, *The Analysis of Linear Partial Differential Equations* (Springer, Berlin, 1985).
- [34] A. Anderson, *Phys. Rev. D* **42**, 585 (1990).
- [35] N. Balazs and B. K. Jennings, *Phys. Rep.* **104**, 347 (1984); M. Hillery, R. F. O'Connell, M. O. Scully, and E. P. Wigner, *ibid.* **106**, 121 (1984).
- [36] J. J. Halliwell, *Phys. Rev. D* **46**, 1610 (1992).
- [37] Lower bounds on the information obtained in this way for arbitrary pairs of noncommuting projections have previously been discussed by M. D. Srinivas, *Pramana* **24**, 673 (1985), but the explicit form of the bounds was not computed.
- [38] For an extensive discussion of quantum mechanics for measurements distributed in time, see C. Caves, *Phys. Rev. D* **33**, 1643 (1985).
- [39] J. J. Halliwell, "Quantum-Mechanical Histories and the Uncertainty Principle. II. Fluctuations about Classical Predictability," Imperial College report 1993 (unpublished).
- [40] M. Mensky, *Phys. Lett. A* **155**, 229 (1991).
- [41] A. O. Caldeira and A. J. Leggett, *Physica* **121A**, 587 (1983).
- [42] R. Omnès, *Ann. Phys. (N.Y.)* **201**, 354 (1990).