PHYSICAL REVIEW A

GENERAL PHYSICS

THIRD SERIES, VOLUME 40, NUMBER 2 JULY 15, 1989

Minimum uncertainty products from the principle of maximum entropy

A. K. Rajagopal and S. Teitler
Naval Research Laboratory, Washington, D.C. 20375-5000
(Received 1 February 1989)

The maximum-entropy method is here generalized to obtain many possible extrema of the uncertainty product corresponding to the generalized minimum uncertainty products recently discussed by Lahiri and Menon (LM) [Phys. Rev. A **38**, 5412 (1988)]. Unlike the LM work, the present work applies to mixed states and leads to a new annealing algorithm for obtaining the extrema of the entropy functional.

I. INTRODUCTION

Recently Lahiri and Menon[1] (LM) developed a variational method for finding the class of wave packets that minimizes the quantum-mechanical uncertainty product of position and its conjugate momentum. This work goes beyond the early work of a similar nature by Jackiw[2] in two ways: (1) by obtaining other extrema and (2) by application to cases where a particle may move in some potential V. These latter cases allow different possible domains for the eigenvalue of the position operator to be considered. The variational functional used by LM does not seem to have a direct physical interpretation so that the physical meaning of the additional extrema is not entirely clear. Also, this method involves wave packets so that its suggested application to quantum cosmology may be problematical. A more general formulation incorporating statistical-mechanical content by means of a density matrix seems more appropriate for such an application.

One can formulate a density matrix corresponding to the Heisenberg uncertainty relation by means of an appropriate choice of constraints in the principle of maximum entropy.[3-5] In Sec. II we carry out such a formulation in an extended version that includes cases considered by LM when the coordinate domain is not $[-\infty,\infty]$. This method leads naturally to the minimum uncertainty product as the least maximum-entropy state when the density matrix collapses to a pure state. We point out that this is equivalent to the standard annealing algorithm for reaching the ground state of a system by allowing a parameter that is either the temperature or its analog to be cooled to zero. The parameter corresponding to the reciprocal temperature (or its analog) is a Lagrange multiplier in the maximum-entropy method. To obtain the other extremal solutions, we construct a hierarchy of renormalized density matrices and their associated Lagrange multipliers which play the role of generalized reciprocal temperaturelike quantities. The annealing algorithms for the hierarchy then picks up the other extrema which are the same as in LM. This procedure is also given in Sec. II. In Sec. III we exhibit the solutions for the special cases considered by LM and point out a technical error in their analysis which, however, does not affect their conclusions. In Sec. IV we discuss a more general context for the annealing algorithm which appears to be a promising method for obtaining the states above the ground state for any system. In Sec. V a summary with concluding remarks is given.

II. PRINCIPLE OF MAXIMUM ENTROPY AND ITS EXTENSIONS FOR THE LM PROBLEM

We use dimensionless units such that the Planck constant $\hbar = h/2\pi = 1$ and consider the usual Hermitian position $\hat{x}$ and its conjugate momentum $\hat{p}$ to be dimensionless operators which obey the usual canonical commutation rules. We consider operators relative to their expectations:

$$\hat{P} = \hat{p} - \langle \hat{p} \rangle, \quad \hat{X} = \hat{x} - \langle \hat{x} \rangle , \tag{2.1}$$

where $\langle \hat{A} \rangle \equiv \mathrm{Tr}(\hat{\rho}\hat{A})$, where $\hat{\rho}$ is the density matrix. As discussed in more detail elsewhere[3-5] we seek to maximize the von Neumann entropy

$$S = -\mathrm{Tr}(\hat{\rho}\ln\hat{\rho}) \tag{2.2}$$

subject to the constraints that the quantities $\langle \hat{x} \rangle, \langle \hat{p} \rangle, \langle \hat{X}^2 \rangle, \langle \hat{P}^2 \rangle$, and

$$\langle \hat{X}\hat{P} + \hat{P}\hat{X} \rangle \equiv 2\langle (\hat{X}\hat{P})_s \rangle \tag{2.3}$$

are given. After some algebra involving the evaluation of all the relevant Lagrange multipliers in terms of the constraints, and the introduction of the reciprocal of a temperaturelike parameter κ, one obtains

$$\hat{\rho} = Z^{-1}(\kappa)\exp(-\kappa\hat{K}) . \tag{2.4}$$

Here

$$\hat{K}=\langle\hat{P}^2\rangle\hat{X}^2+\langle\hat{X}^2\rangle\hat{P}^2-2\langle(\hat{X}\hat{P})_s\rangle(\hat{X}\hat{P})_s \tag{2.5}$$

and $Z(\kappa)$ is the partition function $=\mathrm{Tr}\exp(-\kappa\hat{K})$. There is a self-consistency condition in operation here in that the basis states for this maximum entropy are eigenstates of $\hat{K}$. In other words the maximization of the entropy is *not* carried out with respect to a fixed basis set but rather involves a basis set determined by the constraints (and their boundary conditions).

The parameter κ is determined from the condition

$$\langle\hat{K}\rangle=-\frac{\partial}{\partial\kappa}\ln Z(\kappa)=\tfrac{1}{2}\Omega^2 , \tag{2.6}$$

where

$$\Omega=2[\langle\hat{X}^2\rangle\langle\hat{P}^2\rangle-\langle(\hat{X}\hat{P})_s\rangle^2]^{1/2} . \tag{2.7}$$

The last equality in Eq. (2.6) follows from Eqs. (2.3) and (2.5). Ω must be determined self-consistently because averages that determine the density matrix are also determined by the density matrix itself. This self-consistency is a general feature of the principle of maximum entropy noted above. Clearly, by direct use of Schwartz inequality, one has generally the result

$$\Omega\geq 1 . \tag{2.8}$$

It may be pointed out here that we actually obtain the operator

$$\hat{H}=\kappa\{\hat{K}-[\langle\hat{P}^2\rangle\langle\hat{x}\rangle^2+\langle\hat{x}^2\rangle\langle\hat{p}\rangle^2-2\langle(\hat{X}\hat{P})_s\rangle\langle\hat{x}\rangle\langle\hat{p}\rangle]\} . \tag{2.9}$$

However, because of the normalization condition $\mathrm{Tr}\hat{\rho}=1$, the extra constant contribution in square brackets in Eq. (2.9) does not make an explicit appearance. The values of $\langle\hat{x}\rangle$, $\langle\hat{p}\rangle$, etc., may be calculated self-consistently once $\hat{\rho}$ in Eq. (2.4) is set up; this is similar to the self-consistency property of the calculation in the work of LM.

It may be observed that the operator $\hat{K}$ in Eq. (2.5) is diagonal in terms of the harmonic-oscillator states when applied to the infinite domain of $\hat{x}$ operator. To include situations as in Ref. 1 when there is an infinite potential barrier for negative x, we have to relax the infinite domain condition on the operator $\hat{x}$. When this is done, the $\hat{K}$ operator admits of solutions in terms of the Weber functions $D_m(x)$.[6] Only in the infinite domain do these become the familiar harmonic-oscillator solutions where m becomes a positive integer. LM *assumed* harmonic-oscillator solutions to hold in general even when they considered the case where there is an infinite potential for negative x. They *assumed* further that in this case their Lagrange multiplier $\lambda_1=\langle\hat{x}\rangle$. Such an assumption actually invalidates the variational method. This lack of validity may be shown directly by a self-consistent calculation of $[(\Delta\hat{X})(\Delta\hat{P})]_{\mathrm{sc}}$ which yields the result $(\frac{9}{4}-6/\pi)^{1/2}\approx 0.583$. On the other hand, LM obtained $(\frac{3}{2}-2/\pi)\approx 0.87$. It should be emphasized that this is only a minor technical error in the LM framework and their other conclusions remain intact.

Thus in general we have the solution in the form

$$\hat{K}|m\rangle=\Omega\varepsilon_m|m\rangle, \quad m=0,1,2,\ldots \tag{2.10}$$

with $\langle m|n\rangle=\delta_{m,n}$ and $\{|m\rangle\}$ a complete and nondegenerate set, discrete in one dimension. The actual values of $\Omega\varepsilon_m$ depend on the domain of $\hat{x}$ as we shall show in Sec. III. Then

$$\hat{\rho}=\sum_{m=0}^{\infty} w_m(\kappa)|m\rangle\langle m| , \tag{2.11}$$

with

$$w_m(\kappa)=\exp(-\kappa\Omega\varepsilon_m)/Z(\kappa) , \tag{2.12}$$

$$Z(\kappa)=\sum_{m=0}^{\infty}\exp(-\kappa\Omega\varepsilon_m) . \tag{2.13}$$

κ is determined from the condition Eq. (2.6)

$$\Omega/2=\sum_{m=0}^{\infty}\varepsilon_m\exp(-\kappa\Omega\varepsilon_m)/Z(\kappa) . \tag{2.14}$$

The corresponding value of the entropy is

$$S(\kappa)=-\sum_{m=0}^{\infty} w_m(\kappa)\ln[w_m(\kappa)] . \tag{2.15}$$

By taking the limit $\kappa\to\infty$, i.e., by application of an annealing algorithm it is easy to verify that

$$w_m(\kappa)\to\delta_{m,0} . \tag{2.16a}$$

Consequently

$$\hat{\rho}\to|0\rangle\langle 0| , \tag{2.16b}$$

with

$$(\Omega/2)\to\varepsilon_0 \tag{2.16c}$$

and

$$S(\kappa)\to 0 . \tag{2.16d}$$

Thus the *least maximum entropy* is attained in this limit with the value of Ω given by Eq. (2.16c).

To obtain other states $|N\rangle\langle N|$ with the corresponding eigenvalue ε_N by a similar annealing algorithm we construct a new density matrix

$$\hat{\rho}_N(\kappa_N)=\sum_{m=N}^{\infty} w_m^{(N)}(\kappa_N)|m\rangle\langle m| , \tag{2.17}$$

with

$$w_m^{(N)}(\kappa_N)=\exp(-\kappa_N\Omega_N\varepsilon_m)/Z_N(\kappa_N) \tag{2.18}$$

and

$$Z_N(\kappa_N)=\sum_{m=N}^{\infty}\exp(-\kappa_N\Omega_N\varepsilon_m) . \tag{2.19}$$

The new density matrix is obtained by *removal* of the first N states in $\hat{\rho}$ of Eq. (2.11) and a subsequent renormalization leading to the new reciprocal "temperature" κ_N, determined by

$$\langle\hat{K}\rangle=-\frac{\partial}{\partial\kappa_N}\ln Z_N(\kappa_N)=\Omega_N^2/2 . \tag{2.20}$$

Recall that Ω_N is now defined in terms of $\hat{\rho}_N$. Thus

$$\tfrac{1}{2}\Omega_N = \sum_{m=N}^{\infty} \varepsilon_m \exp(-\kappa_N \Omega_N \varepsilon_m)/Z(\kappa_N) \ . \tag{2.21}$$

The corresponding entropy S_N is

$$S_N(\kappa_N) = -\sum_{m=N}^{\infty} w_m^{(N)}(\kappa_N) \ln w_m^{(N)}(\kappa_N) \ . \tag{2.22}$$

We now observe that by setting $\kappa_N \to \infty$ we obtain

$$\tfrac{1}{2}\Omega_N \to \varepsilon_N \ , \tag{2.23}$$

with

$$\hat{\rho}_N \to |N\rangle\langle N|, \quad S_N \to 0 \ . \tag{2.24}$$

Thus we have here a new annealing algorithm which enables us to obtain all the self-consistent eigenvalues of the $\hat{K}$ operator. In Sec. III we briefly give the actual results for the two special cases considered by LM. These examples should make clear the detailed workings of the method.

III. TWO EXAMPLES

A. Infinite domain

Here we have the familiar harmonic-oscillator states

$$\varepsilon_m = (m + \tfrac{1}{2}), \quad m = 0,1,2,\ldots \tag{3.1}$$

and all the sums appearing in Sect. II can be carried out explicitly. They are given here for completeness,

$$Z_N(\kappa_N) = \exp[-\kappa_N \Omega_N (N+\tfrac{1}{2})]/[1-\exp(-\kappa_N \Omega_N)] \ , \tag{3.2}$$

$$\kappa_N = \frac{1}{\Omega_N} \ln\left[\frac{\Omega_N - (2N-1)}{\Omega_N - (2N+1)}\right], \quad \Omega_N \geq 2N+1 \ . \tag{3.3}$$

Hence

$$Z_N(\kappa_N) = \left[\frac{\Omega_N - 2N + 1}{2}\right]\left[\frac{\Omega_N - 2N - 1}{\Omega_N - 2N + 1}\right]^{N+1/2}, \tag{3.4}$$

$$S^{(N)} = \left[\frac{\Omega_N - 2N + 1}{2}\right]\ln\left[\frac{\Omega_N - 2N + 1}{2}\right] - \left[\frac{\Omega_N - 2N - 1}{2}\right]\ln\left[\frac{\Omega_N - 2N - 1}{2}\right], \tag{3.5}$$

$$\hat{\rho}_N = \sum_{m=N}^{\infty} w_m^{(N)} |m\rangle\langle m| \tag{3.6}$$

with

$$w_m^{(N)} = \left[\frac{2}{\Omega_N - 2N - 1}\right]\left[\frac{\Omega_N - 2N - 1}{\Omega_N - 2N + 1}\right]^{m-N} . \tag{3.7}$$

Note that Ω_N contains the appropriate averages of the operator combinations with respect to the corresponding density matrix $\hat{\rho}_N$ and hence is determined self-consistently. For $N=0$ we recover our previous result with the least maximum entropy then involving the entire density matrix and $\Omega_0 = 1$, the minimum uncertainty. For $N = 1,2,\ldots,$ we obtain $\Omega_N = (2N+1)$, the other subsidiary minima associated with the uncertainty products as in Ref. 1 but now for corresponding density matrix $\hat{\rho}_N \to |N\rangle\langle N|$.

B. Infinite potential barrier for $x < 0$

Here the states are not the harmonic oscillator states but rather the Weber solutions. The ε_m's are not simple as in Sec. III A but are obtained numerically as zeros of appropriate Weber functions that satisfy the condition of zero state vector for $x \leq 0$. Recently, Marin and Cruz[7] have given numerical solutions for a more general problem of the eigenvalues of the harmonic oscillator confined between two infinite potential walls. Here again one obtains the Weber functions as solutions and the eigenvalues ε_m are determined numerically. The indicated sums in Sec. II can only be performed numerically in these cases.

IV. GENERAL SCHEME

The method given above may be generalized to allow the determination of the states of a density matrix outside the context of the uncertainty relation. We give here an outline for such a scheme. Suppose we have a system described by a Hermitiam operator bounded from below, $\hat{H}$, whose eigenstates are given by (for simplicity we here consider discrete spectra)

$$\hat{H}|n,\alpha_n\rangle = \varepsilon_n |n,\alpha_n\rangle \ . \tag{4.1}$$

Here α_n is degeneracy index of the state $|n,\alpha_n\rangle$ with D_n the actual degeneracy. The set $\{|n,\alpha_n\rangle\}$ is a complete, orthonormal set. We then construct the hierarchy of density matrices

$$\rho_N(\kappa_N) = \sum_{n=N}\sum_{\alpha_n} |n,\alpha_n\rangle w_n(\kappa_N)\langle n,\alpha_n| \ , \tag{4.2}$$

with

$$w_n(\kappa_N) = \exp(-\kappa_N \varepsilon_n)/Z(\kappa_N) \quad (\alpha_n = 1 \ldots D_n) \tag{4.3}$$

and

$$Z(\kappa_N) = \sum_{n=N} D_n \exp(-\kappa_N \varepsilon_n) \ . \tag{4.4}$$

The parameter κ_N is determined by the constraint that a fixed average of $\hat{H}$ with respect to $\hat{\rho}$ is given,

$$\langle \hat{H} \rangle = -\frac{\partial}{\partial \kappa_N} \ln[Z_N(\kappa_N)] = \sum_{n=N} \varepsilon_n D_n w_n(\kappa_N) \ . \tag{4.5}$$

Note the occurrence of degeneracy of $w_n(\kappa_N)$ in Eq. (4.3). The entropy associated with ρ_N is

$$S_N(\kappa_N) = -\sum_{n=N} D_n w_n(\kappa_N) \ln w_n(\kappa_N) \ . \tag{4.6}$$

In the limit $\kappa_N \to \infty$, we observe that

$$w_n(\kappa_N) \rightarrow \delta_{n,N}/D_N \ , \tag{4.7}$$

$$\rho_N(\kappa_N) \rightarrow \frac{1}{D_N} \sum_{\alpha_n} |N,\alpha_n\rangle\langle N,\alpha_n| \ , \tag{4.8}$$

$$S_N(\kappa_N) \rightarrow \ln(D_N) \ , \tag{4.9}$$

and

$$\langle \hat{H} \rangle_N \rightarrow \varepsilon_N \ . \tag{4.10}$$

If the ground eigenstate of $\hat{H}$ for ε_N is nondegenerate then $S \rightarrow 0$ and the density matrix $\rho \rightarrow |N\rangle\langle N|$. The idea then is to delete the first "N" states and construct a renormalized density matrix with appropriate reciprocal "temperature" parameter through which the annealing algorithm is set up to yield the various states of the operator $\hat{H}$. The uncertainty relation considered in detail above had the feature that the operator corresponding to $\hat{H}$ contained quantities which depended on the $\hat{\rho}$. A somewhat similar situation is encountered in mean field theories and these will be explored elsewhere. For the present we note the simple scheme outlined here exhibits the feature that the microscopic states are obtained by the annealing procedure $\kappa_N \rightarrow \infty$ where the macroscopic average energy $\langle \hat{H} \rangle_N$ approaches the microscopic eigenvalue ε_N and the density matrix $\hat{\rho}_N$ approaches the associated (possibly degenerate) set of states Eq. (4.8).

V. SUMMARY AND CONCLUDING REMARKS

In statistical mechanics a weighting procedure is employed for all the microscopic states of the system in obtaining macroscopic physical quantities of interest. When the system is cooled to zero temperature only the weight associated with the lowest eigenstate(s) is nonzero and the macroscopic physical quantity becomes the lowest eigenstate average. This principle or its analog has become a useful technique called simulated annealing in a variety of other problems of optimization. The method suggested here introduces further steps in the simulated annealing procedure. One arrives at the averages appropriate to the next state and so forth on to the other desired state averages as well. We have illustrated this in the context of uncertainty products where the stringent requirements of self-consistency, indicate that the method can be employed elsewhere (e.g., many-body problems). The simple example given in Sec. IV involves only the usual self-consistency requirement inherent in the application of the principle of maximum entropy subject to the constraint that the average of an operator bounded from below is given. However, it is illustrative of the principles involved in the suggested scheme. We plan to explore this method in other contexts in future communications.

ACKNOWLEDGMENTS

One of us (A.K.R.) is supported in part by Contract No. N0001489WX24034.

[1]J. Lahiri and V. J. Menon, Phys. Rev. A **38**, 5412 (1988).
[2]R. Jackiw, J. Math. Phys. **9**, 339 (1968).
[3]A. K. Rajagopal and S. Teitler, in *Maximum-Entropy and Bayesian Methods in Science and Engineering*, edited by G. J. Erickson and C. R. Smith (Reidel, New York, 1988), Vol. 2, p. 121.
[4]E. H. Wichmann, J. Math. Phys. **4**, 886 (1963).
[5]W. J. Titus, Am. J. Phys. **47**, 357 (1979).
[6]P. M. Morse and H. Feshbach, *Methods in Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. II, pp. 1565 and 1566.
[7]J. L. Marin and S. A. Cruz, Am. J. Phys. **56**, 1134 (1988). We thank Professor A. R. P. Rau for bringing this paper to our attention.