Extended mean-field treatments and information theory

R. Rossignoli and A. Plastino
Department of Physics, National University of La Plata, Casilla de Correo 67, (1900) La Plata, Argentina
(Received 18 January 1990)

An extended mean-field approximation for statistical operators is derived within the context of information theory. The ensuing scheme is applied to (many) boson mean-field approaches within a completely general statistical framework. Numerical examples are given within the context of the finite-temperature anharmonic oscillator problem. Fluctuations are also discussed.

I. INTRODUCTION

Information theory[1,2] (IT) provides a very general and convenient framework on which to base statistical mechanics.[3–5] It allows for the straightforward extension of statistical descriptions to finite systems, off-equilibrium situations and nonconventional ensembles.[6–12]

The aim of the present effort is to discuss approximate self-consistent mean-field descriptions within a general IT statistical context. Approaches such as thermal Hartree-Fock[13] or Hartree-Fock-Bogoliubov[14,15] and ensuing cranked[13] (i.e., rotating frame) extensions[14–16] in fermion systems, as well as similar statistical Hartree-Bose schemes,[17,18] arise as a consequence of the need for tractable descriptions of interacting many-body (albeit possibly finite) systems, and are able to provide a vivid physical picture, including the prediction of phase transitions and metastable solutions.[19] In the case of finite systems, these mean-field transitions constitute in general a signature of significant changes in the structure of the system. These, in turn, arise as the consequence of the variation of certain control parameters, and may manifest themselves more clearly in mean field approaches than in exact descriptions.[20,21]

Within the information-theoretic context, as shall be discussed below, these self-consistent schemes can be implemented by including very specific ingredients in the construction of the density operator. Together with the maximum entropy criterion, the observer chooses the operators appearing in the exponent of $\hat{\rho}$, in terms of which the available data will be processed.

A quite general self-consistent approximation for statistical operators is derived in Sec. II within the information-theory context. The scheme includes, as a particular case, the standard statistical mean-field descriptions, which are generalized in Sec. III to accommodate a completely general statistical context. An illustrative example is given in Sec. IV, where mean-field solutions are constructed directly from a set of expectation values (and not from a set of Lagrange multipliers as usual). This allows for a different picture of the behavior of mean-field quantities, appropriately suited to the IT philosophy. The ensuing effective temperatures and Lagrange multipliers are accordingly calculated, and the inference of fluctuations is also examined. Numerical results are given in Sec. V for the case of an anharmonic oscillator, and compared with exact predictions. Finally, some conclusions are drawn in Sec. VI.

II. GENERALIZED SELF-CONSISTENT APPROXIMATION

A. Introductory remarks

Let us consider a quantum system, about which the only available information consists of the expectation values O_i of n linearly independent observables $\hat{O}_i$. According to IT, the least biased normalized statistical operator describing the system is[3–12]

$$\hat{\rho}=\exp\left[-\lambda_0-\sum_{i=1}^{n}\lambda_i\hat{O}_i\right] . \tag{2.1}$$

The operator (2.1) maximizes the entropy (we set Boltzmann's constant $k_B=1$)

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

subject to the linear constraints

$$\langle\hat{O}_i\rangle=\mathrm{Tr}\hat{\rho}\hat{O}_i=O_i , \tag{2.3}$$

where the trace is taken over the set of accessible states. The parameters λ_i are Lagrange multipliers which are to be adjusted in order to comply with (2.3). The normalization parameter λ_0 is obviously given by

$$\lambda_0=\ln\left[\mathrm{Tr}\exp\left[-\sum_{i=1}^{n}\lambda_i\hat{O}_i\right]\right] , \tag{2.4}$$

and can also be interpreted as the Lagrange multiplier associated with the identity operator $\hat{O}_0\equiv I$.

We remark that, within the information-theoretic context, the observables $\hat{O}_i$ are completely arbitrary. They constitute, in general, a non-Abelian set, and the operator (2.1) is not necessarily stationary, being thus suitable for the description of arbitrary systems in off-equilibrium situations.[3–12]

The ensuing maximum entropy $S=\lambda_0+\sum_{i=1}^{n}\lambda_i\hat{O}_i$, as a function of the n relevant mean values O_i, verifies

$$\frac{\partial S}{\partial O_i}=\lambda_i , \tag{2.5a}$$

whereas its Legendre transform $S'=S-\sum_{i=1}^{n}\lambda_i O_i=\lambda_0$, considered as a function of the n λ_i's, fulfills

$$\frac{\partial S'}{\partial \lambda_i}=-O_i \ . \tag{2.5b}$$

B. Self-consistent approach

Let us state now the basis of our approach. In the spirit of IT, one builds up $\hat{\rho}$ with the given operators $\hat{O}_i$. However, the corresponding statistical operator may constitute a formidable object indeed, especially if some of the $\hat{O}_i$ are many-body operators, since a diagonalization of the exponent over the complete set of accessible states is required. The whole IT machinery may thus become useless, due to the intractability of the concomitant set of equations. Exceptions occur if the observables are of a simple character, and can be written, for instance, as one-body operators in some representation, or are simultaneously diagonal in a given known basis. In these cases the relationship between Lagrange parameters and expectation values is of a straightforward nature, and the inference of arbitrary mean values poses no difficulties. We propose thus an approximate density of the form

$$\hat{\rho}_{\rm app}=\exp\left[-\lambda_0-\sum_{j=1}^{m}\lambda_j\hat{R}_j\right] , \tag{2.6}$$

where $\{\hat{R}_j, j=1,\dots,m, m\geq n\}$ is a set of suitable operators, chosen by the observer in order to render (2.6) "tractable." The parameters λ_j will be determined by maximizing the entropy S associated with (2.6) subject to the n constraints $\mathrm{Tr}\hat{\rho}_{\rm app}\hat{O}_i=O_i$. Introducing n additional multipliers β_i, and making the quantity

$$S'=-\mathrm{Tr}\hat{\rho}_{\rm app}\left[\ln(\hat{\rho}_{\rm app})+\sum_{i=1}^{n}\beta_i\hat{O}_i\right]$$

stationary with respect to the yet unknown expectation values R_j, we obtain, using (2.5a) [applied to (2.6)],

$$\begin{aligned}\lambda_j&=\sum_{i=1}^{n}\beta_i\frac{\partial O_i}{\partial R_j}\\ &=\sum_{i=1}^{n}\sum_{l=1}^{m}\beta_i A_{il}B_{lj}^{-1}, \quad j=1,\dots,m ,\end{aligned} \tag{2.7}$$

where

$$A_{il}=\frac{\partial O_i}{\partial \lambda_l}=-\langle \hat{R}_l^{*}\hat{O}_i\rangle=-\langle \hat{O}_i^{*}\hat{R}_l\rangle \tag{2.8}$$

and

$$\begin{aligned}B_{jl}&=\frac{\partial R_j}{\partial \lambda_l}=-\frac{\partial^2\lambda_0}{\partial\lambda_l\partial\lambda_j}\\ &=-\langle \hat{R}_l^{*}\hat{R}_j\rangle \ .\end{aligned} \tag{2.9}$$

We denote by $\hat{R}_j$ the so-called Kubo transform[22]

$$\hat{R}_j^{*}=\int_0^1(\hat{\rho}_{\rm app})^{-u}\hat{R}_j(\hat{\rho}_{\rm app})^{u}du-\mathrm{Tr}(\hat{\rho}_{\rm app}\hat{R}_j) , \tag{2.10}$$

such that $\partial\hat{\rho}/\partial\lambda_j=-\hat{\rho}\hat{R}_j^{*}$, $\langle\hat{R}_j^{*}\rangle=0$ [λ_0 has been considered as a function of the remaining λ's through (2.4) and (2.5b)]. Using (2.7), the approximate density operator can be written as

$$\hat{\rho}_{\rm app}=\exp\left[-\lambda_0-\sum_{i=1}^{n}\beta_i\hat{o}_i\right] , \tag{2.11}$$

with

$$\hat{o}_i=\sum_{j=1}^{m}\frac{\partial O_i}{\partial R_j}\hat{R}_j \ . \tag{2.12}$$

Expression (2.11) leads obviously to a nonlinear system, since the approximate statistical operator depends upon the very mean values R_j that it determines (self-consistency). The exception occurs if all operators $\hat{O}_i$ are linearly related to the $\hat{R}_j$'s, in which case (2.12) implies $\sum_j\lambda_j\hat{R}_j=\sum_i\beta_i\hat{O}_i$. Otherwise, one faces the nonlinear system (2.7), where the right-hand side depends on the λ_j's through (2.8) and (2.9). If the expectation values appearing in (2.7)–(2.9) can be easily expressed in terms of the mean values R_j's, it is possible to deal instead with the equivalent problem

$$\mathrm{Tr}[\hat{\rho}_{\rm app}(R_j)\hat{R}_j]=R_j, \quad j=1,\dots,m \ . \tag{2.13}$$

C. Some elementary considerations regarding the self-consistent description: Fluctuations

The present, general self-consistent approach yields obviously a lower bound to the exact S and S' for fixed values of the O_i's and β_i's, respectively (if traces are evaluated by summing over the same set of accessible states), since we are constraining the trial densities to adopt the form (2.6). The most general statistical mean-field treatment can be straightforwardly obtained from (2.11), if the operators R_j are chosen to be general quadratic functions of creation and annihilation operators, as will be seen in Sec. III.

Since the Eqs. (2.7) [or (2.13)] are nonlinear, more than one solution may exist for fixed parameters β_i, and not all of them will correspond to maxima of S'. Certainly, minima and saddle points may also occur. The exact solution (2.1) is, however, unique and always yields a maximum.[4]

On the other hand, for fixed mean values O_i, the β_i's must be determined from the constraints (2.3), and will not coincide, in general, with the exact parameters appearing in (2.1). A general solution will not always exist in this case, since the range of mean values O_i, spanned by (2.6), may be smaller than the exact range. At the same time, the nonlinearity may give rise to various simultaneous solutions for the β_i's, the best of which is, in principle, that which yields the highest entropy.

System (2.7) (or 2.13) can be solved by iteration, starting with a set λ_j^0, with which the initial matrices (2.8) and (2.9) are to be calculated, i.e.,

$$\lambda_j^{s+1}=\sum_{i=1}^{n}\sum_{l=1}^{m}\beta_i[A_{il}^{s}(B^{s})_{lj}^{-1}] , \tag{2.14}$$

where the superscript s denotes the corresponding iterative step. The iteration process is to be continued until self-consistency is reached, i.e., convergence in the λ_j's or in the mean values R_j's.

The approximate S and S', constructed with the solution (2.11), still fulfill the relationships (2.5), i.e.,

$$\frac{\partial S}{\partial O_i}=\beta_i \ , \tag{2.15a}$$

$$\frac{\partial S'}{\partial \beta_i}=-O_i \ , \tag{2.15b}$$

due to the stationary condition (2.7), which may be used to relate expectation values and Lagrange parameters. However, λ_0 and S' cease to be identical:

$$S'-\lambda_0=\sum_{i=1}^{n}\beta_i\langle\hat{o}_i-\hat{O}_i\rangle \ .$$

In the exact picture, it is possible to identify thermodynamic derivatives with fluctuations, or, within the general non-Abelian context [see (2.10)], with[23]

$$\frac{\partial O_i}{\partial \beta_k}=-\frac{\partial^2 S'}{\partial\beta_k\partial\beta_i}=-\langle\hat{O}\,{}_k^*\hat{O}_i\rangle \ . \tag{2.16}$$

However, in the present self-consistent picture (2.16) is no longer valid due to the β dependence of the operators (2.12). We have instead [cf. (2.11) and (2.12)]

$$\frac{\partial R_j}{\partial \beta_k}=C_{jk}+\sum_{l=1}^{m}D_{jl}\frac{\partial R_l}{\beta_k} \ , \tag{2.17}$$

where [see (2.9)]

$$C_{jk}=-\langle\hat{o}\,{}_k^*\hat{R}_j\rangle=-\sum_{l=1}^{m}\frac{\partial O_k}{\partial R_l}B_{lj} \tag{2.18}$$

$$\begin{aligned}D_{jl}&=-\sum_{i=1}^{n}\beta_i\left\langle\left[\frac{\partial\hat{o}_i}{\partial R_l}\right]^*\hat{R}_j\right\rangle\\&=-\sum_{i=1}^{n}\sum_{l'=1}^{m}\beta_i\frac{\partial^2 O_i}{\partial R_l\partial R_{l'}}B_{l'j} \ .\end{aligned} \tag{2.19}$$

Therefore,

$$\frac{\partial R_j}{\partial \beta_k}=\sum_{l=1}^{m}(I-D)_{jl}^{-1}C_{lk} \ , \tag{2.20}$$

and finally,

$$\frac{\partial O_i}{\partial \beta_k}=\sum_{j,l=1}^{m}\frac{\partial O_i}{\partial R_j}(I-D)_{jl}^{-1}C_{lk} \ , \tag{2.21}$$

so that (2.16) is recovered when $D=0$. Actually, the thermodynamic derivative may provide a better estimate of the "true" fluctuation than a direct evaluation (see Secs. IV and V).

III. GENERAL STATISTICAL MEAN-FIELD APPROACH FOR BOSE SYSTEMS

We shall apply now the previous formalism to the case where the operators $\hat{R}_j$ are general one-body operators. Let us first discuss the statistical one-body description from the IT viewpoint. We shall examine the case of a boson system. The corresponding fermion expressions can be straightforwardly obtained from the boson case. The most general one-body statistical operator can be written as

$$\begin{aligned}\hat{\rho}=\exp\Bigg[&-\lambda_0-\sum_{i,j}[\Lambda_{ij}b_i^\dagger b_j+\tfrac{1}{2}(\Gamma_{ij}b_ib_j+\Gamma_{ij}^*b_j^\dagger b_i^\dagger)]\\&-\sum_i(\eta_ib_i^\dagger+\eta_i^*b_i)\Bigg] \ ,\end{aligned} \tag{3.1}$$

where $b_i^\dagger$ (b_i) creates (annihilates) a boson in the single-particle (SP) state labeled by the index $i=1,\dots,L$ (assumed discrete). Obviously, the Hermiticity of $\hat{\rho}$ implies $\Lambda_{ij}=\Lambda_{ji}^*$, and we can set $\Gamma_{ij}=\Gamma_{ji}$. It is convenient in this context to define $Z^\dagger=(b_1^\dagger,\dots,b_L^\dagger,b_1,\dots,b_L)$, $F^\dagger=(\eta_1^*,\dots,\eta_L^*,\eta_1,\dots,\eta_L)$ (with Z and F the adjoint column vectors), and the Hermitic matrix of multipliers

$$M=\begin{bmatrix}\Lambda & \Gamma^*\\ \Gamma & \Lambda^*\end{bmatrix} \ . \tag{3.2}$$

With the above definitions, the boson commutation properties $[b_i,b_j^\dagger]=\delta_{ij}$ can be written as $ZZ^\dagger-[(Z^\dagger)^{\mathrm{tr}}Z^{\mathrm{tr}}]^{\mathrm{tr}}=\Pi$, with

$$\Pi=\begin{bmatrix}I & 0\\ 0 & -I\end{bmatrix} \ , \tag{3.3}$$

and the operator (3.1) can be succinctly cast as

$$\begin{aligned}\hat{\rho}&=\exp[-\lambda_0'-\tfrac{1}{2}(Z^\dagger MZ+F^\dagger Z+Z^\dagger F)]\\&=\exp[-\lambda_0^c-\tfrac{1}{2}(Z^c)^\dagger MZ^c] \ ,\end{aligned} \tag{3.4}$$

where $\lambda_0'=\lambda_0-\frac{1}{2}\mathrm{tr}(\Lambda)$, $\lambda_0^c=\lambda_0'-\frac{1}{2}F^\dagger M^{-1}F$, and

$$Z^c=Z+M^{-1}F \ . \tag{3.5}$$

Z^c denotes the set of "centered" boson operators b_i^c, with vanishing mean values. The complete set of one-body expectation values can be conveniently accommodated in the column vector $\langle Z\rangle$ and in the generalized SP density matrix, which can be defined in this context as

$$D=\langle ZZ^\dagger\rangle-\Pi=D^c+\langle Z\rangle\langle Z^\dagger\rangle \ , \tag{3.6}$$

where D^c is the "covariance" matrix

$$D^c=\langle Z^c(Z^c)^\dagger\rangle-\Pi=\begin{bmatrix}A & B\\ B^* & I+A^*\end{bmatrix} \ , \tag{3.7}$$

with $A_{ij}=\langle(b_j^c)^\dagger b_i^c\rangle$, $B_{ij}=\langle b_j^c b_i^c\rangle$. In this way [see (2.5b)],

$$\tfrac{1}{2}D_{ji}=-\frac{\partial\lambda_0'}{\partial M_{ij}} \ , \tag{3.8}$$

and $\frac{1}{2}\langle Z\rangle_i=-\partial\lambda_0'/\partial F_i^\dagger$, $\frac{1}{2}D_{ji}^c=-\partial\lambda_0^c/\partial M_{ij}$. Both D and A are Hermitic, while B is symmetric. The operator (3.1) can be put into diagonal form by means of a general Bogoliubov transformation for bosons,[24]

$$Z'=WZ^c=W(Z+M^{-1}F) \ , \tag{3.9}$$

with

$$W=\begin{bmatrix} X & Y \\ Y^* & X^* \end{bmatrix} , \tag{3.10}$$

a unitary matrix with respect to the metric $\Pi(W\Pi W^\dagger=\Pi)$, in order to preserve the boson commutation relations. In the primed SP basis defined by (3.9), we have, obviously, $M=W^\dagger M'W$, $D'=WD^cW^\dagger$, which can be conveniently rewritten as

$$(M'\Pi)=U^{-1}(M\Pi)U,\quad (\Pi D')=U^{-1}(\Pi D^c)U , \tag{3.11}$$

where $U=W^\dagger$. By choosing M' diagonal [which entails, according to (3.11), the diagonalization of the non-Hermitic matrix $M\Pi$], (3.1) can be finally written as

$$\hat{\rho}=\exp\{-\lambda_0^d-\sum \lambda_i b_i'^\dagger b_i'\} , \tag{3.12}$$

with $\lambda_0^d=\lambda_0^c+\frac{1}{2}\sum_i \lambda_i$. Therefore, the corresponding density matrix $D'=\langle Z'Z'^\dagger\rangle-\Pi$ is also diagonal, so that

$$\langle b_i'^\dagger b_j'^\dagger\rangle=\langle b_j' b_i'\rangle^*=0,\quad \langle b_i'^\dagger b_j'\rangle=f_i\delta_{ij} .$$

In a standard grand canonical (GC) ensemble, we obtain the well-known formulas

$$\lambda_0^d=-\sum_i \ln[1-\exp(-\lambda_i)]$$

and

$$f_i=[\exp(\lambda_i)-1]^{-1} ,$$

which can be written in matrix notation as $\Pi D'=[\exp\{M'\Pi\}-I]^{-1}$. Therefore, by means of (3.11) and (3.6), we are led, in the original basis, to the final result

$$\Pi D^c=[\exp(M\Pi)-I]^{-1} , \tag{3.13a}$$

$$\langle Z\rangle=-M^{-1}F , \tag{3.13b}$$

which gives the relationship between Lagrange multipliers and expectation values in the GC ensemble, enabling one to determine $\hat{\rho}$ from a complete knowledge of D^c and $\langle Z\rangle$, and vice versa. Our IT formalism allows one also to conventiently deal with situations of *incomplete* one-body information, in which case (3.13) provides a set of appropriate equations, namely,

$$0=M_{ij}=\{\ln[(\Pi D)^{-1}+I]\}_{ij} \tag{3.14a}$$

$$0=F_i=[M\langle Z\rangle]_i , \tag{3.14b}$$

which determines those elements D_{ij}^c, $\langle Z\rangle_i$ previously unknown, according to the maximum entropy principle.

The mean value of a general one-body boson operator $\hat{O}=Z^\dagger OZ+G^\dagger Z$ can be expressed in terms of (3.13) as $\langle\hat{O}\rangle=\mathrm{tr}[OD]+G^\dagger\langle Z\rangle$, where tr denotes the trace in the $2L$-dimensional quasiparticle space. In particular, the entropy acquires the appearance

$$S=\lambda_0^c+\tfrac{1}{2}\,\mathrm{tr}(MD^c) , \tag{3.15}$$

which in a standard GC ensemble coincides with the well-known formula

$$S=\sum_i [(f_i+1)\ln(f_i+1)-f_i\ln(f_i)] .$$

The entropy is evidently independent of the multipliers η_i, depending solely on the covariance matrix D^c, determined completely by M.

For $\mathrm{Tr}(\hat{\rho})$ to be finite, it is necessary that $\lambda_i>0\forall i$. We should remark, however, that the diagonalization of the exponent of $\hat{\rho}$ by means of a transformation of the type (3.9) is not always possible. Physically meaningful expectation values (which fulfill, for instance, the uncertainty principle) correspond to a normalizable statistical operator (see Appendix for a specific example). In such a case, the matrices $M\Pi$ and ΠD^c can be diagonalized, and have real eigenvalues (λ_i, f_i, respectively), that appear in pairs of opposite sign. This entails that the hermitic matrices M and D^c possess only positive nonvanishing eigenvalues. The vector F and the expectation values $\langle b_i\rangle$ can adopt, obviously, arbitrary values.

A very important situation is that in which the available information is expressed in terms of generalized coordinates and momenta

$$\hat{P}_i=-i(b_i-b_i^+)/2^{1/2} , \tag{3.16a}$$

$$\hat{Q}_i=(b_i+b_i^\dagger)/2^{1/2} , \tag{3.16b}$$

with $[\hat{Q}_i,\hat{P}_j]=i\delta_{ij}$. The whole formalism can be recast in terms of the operators (3.16), in which case, the exponent of (3.1) is a general (hermitic) quadratic function of $\hat{P}_i$ and $\hat{Q}_i$, and (3.9) represents a linear canonical transformation, with (3.5) a translation (in $\hat{P}$ and $\hat{Q}$).

On the other hand, if $\Gamma=F=0$ in (3.1), W reduces, obviously, to a standard (boson-number conserving) unitary transformation X ($Y=0$), and $\hat{\rho}$ can always be put into the diagonal form (3.12). (3.13) reduces in this case to

$$A=[\exp(\Lambda)-I]^{-1} , \tag{3.17}$$

and $\langle Z\rangle=0$. However, in this situation canonical ensembles (with a fixed number N of bosons) can be considered, in which case (3.17) should be replaced by an appropriate numerical relationship between the eigenvalues of D and Λ. No restrictions upon the sign of λ_i would arise in this case.

In the case of a fermion system, the multipliers η_i vanish, and the Bogoliubov matrix W is Hermitic with respect to the standard metric ($\Pi\rightarrow I$). The corresponding fermion matrices M and D [defined as in (3.8)] are formally equivalent to $M\Pi$ and $D\Pi$, but are now hermitic, and (3.13a) should be replaced, in a GC ensemble, by $D=[I+\exp(M)]^{-1}$.

Let us turn our attention now to the generalized mean-field approach, in which the exact statistical operator is approximated by a density of the form (3.1). We assume that the available information consists of the expectation values of n arbitrary boson operators $\hat{O}_i$. In a GC ensemble, it is possible to employ the general statistical version of Wick's theorem,[9] which enables one to express expectation values of arbitrary n-body boson operators [with respect to (3.1)] as products of SP mean values, and hence, to use (2.13). According to Sec. II, the effective operators (3.12) will now be SP operators of the form

$$\hat{o}_i=\tfrac{1}{2}(Z^{\dagger}M^iZ+Z^{\dagger}F^i+F^{i\dagger}Z)\ , \qquad (3.18)$$

where $\frac{1}{2}M^i_{kj}=\partial O_i/\partial D_{jk}$, $\frac{1}{2}F^i_j=\partial O_i/\partial\langle Z^{\dagger}\rangle_j$, or, in a more explicit fashion,

$$\Lambda^i_{kj}=\frac{\partial\langle\hat{O}_i\rangle}{\partial\langle b^{\dagger}_k b_j\rangle},\quad \tfrac{1}{2}\Gamma^i_{kj}=\frac{\partial\langle\hat{O}_i\rangle}{\partial\langle b^{\dagger}_k b^{\dagger}_j\rangle}\ , \qquad (3.19)$$

and $\eta^*_i=\partial\langle O_i\rangle/\partial\langle b_i\rangle$, where $\langle\hat{O}_i\rangle$ is considered a function of the elements of (3.8) and $\langle Z\rangle$. Both M^i and F^i will depend on D^c and $\langle Z\rangle$. The relationship between Lagrange multiplers and mean values in a GC ensemble is thus given by the nonlinear system

$$\Pi D^c=\{\exp[M(D^c,\langle Z\rangle)\Pi]-I\}^{-1} \qquad (3.20a)$$

$$\langle Z\rangle=-M^{-1}F(D^c,\langle Z\rangle)\ , \qquad (3.20b)$$

where $M=\sum_{i=1}^n\beta_iM^i$, $F=\sum_{i=1}^n\beta_iF^i$, which condenses the general statistical Hartree-Bogoliubov Boson equations. Notice that if $\langle\hat{O}_i\rangle$ is considered a function of D^c (instead of D) and $\langle Z\rangle$, $\langle Z^{\dagger}\rangle$, (3.20b) is equivalent to

$$\partial\left[\sum_{i=1}^n\beta_iO_i(D^c,\langle Z\rangle,\langle Z^{\dagger}\rangle)\right]\Big/\partial\langle Z^{\dagger}\rangle=0\ , \qquad (3.21)$$

which asserts that $F=0$ in the basis where $\langle Z'\rangle=0$.

System (3.20) can be solved by iteration, starting, for instance, with an initial matrix D^0 and an initial vector $\langle Z\rangle^0$. This entails the diagonalization of $M\Pi$ in each step, i.e., (see 3.11)

$$\begin{bmatrix}\Lambda^s-\Gamma^{s*}\\ \Gamma^s-\Lambda^{s*}\end{bmatrix}\begin{bmatrix}X^{\dagger}\\ Y^{\dagger}\end{bmatrix}=\begin{bmatrix}X^{\dagger} & \Lambda'^s\\ Y^{\dagger} & \Lambda'^s\end{bmatrix}\ , \qquad (3.22)$$

where $\Lambda'^s_{ij}=\lambda^s_i\delta_{ij}$, so that [writing (3.20) explicitly],

$$\begin{aligned}A^{s+1}&=X^{\dagger}A'X+Y^{\rm tr}(I+A')Y^*\ ,\\ B^{s+1}&=-X^{\dagger}A'Y-Y^{\rm tr}(I+A')X^*\ ,\end{aligned} \qquad (3.23)$$

$$\begin{aligned}\langle b\rangle^{s+1}=&-(X^{\dagger}C'X+Y^{\rm tr}C'Y^*)\eta^{s*}\\ &+(X^{\dagger}C'Y+Y^{\rm tr}C'X^*)\eta^s\ ,\end{aligned} \qquad (3.24)$$

where

$$A'_{ij}=[\exp(\lambda^s_i)-1]^{-1}\delta_{ij},\quad C'_{ij}=(\lambda^s_i)^{-1}\delta_{ij}\ .$$

M^{s+1} and F^{s+1} are then built according to (3.19) using A^{s+1},B^{s+1} and $\langle Z\rangle^{s+1}$. In canonical ensemble treatments, (3.20) is no longer valid, and it may be convenient to work directly with (2.7) and (2.14).

Finally we would like to remark that the operator (3.1) preserves its form in time if the temporal evolution is described by a general SP Hamiltonian $\hat{H}=\frac{1}{2}(Z^{\dagger}HZ+G^{\dagger}Z+Z^{\dagger}G)$, basically because the semialgebra formed by $\hat{H}$ and the set of general one-body operators is closed under commutation. The ensuing equations of motion can be conveniently cast in terms of D^c and $\langle Z\rangle$, and are easily shown to be, by means of Ehrenfest theorem (we set $\hbar=1$),

$$\begin{aligned}i\frac{d(\Pi D^c)}{dt}&=[H\Pi,\Pi D^c]\ ,\\ i\frac{d(\Pi\langle Z\rangle)}{dt}&=H\Pi(\Pi\langle Z\rangle)+G\ .\end{aligned} \qquad (3.25)$$

The evolution of D^c is thus decoupled from that of $\langle Z\rangle$. In a similar fashion, we obtain analogous equations for the Lagrange multipliers,

$$\begin{aligned}i\frac{d(M\Pi)}{dt}&=[H\Pi,M\Pi]\\ i\frac{dF}{dt}&=H\Pi F-(M\Pi)G\ .\end{aligned} \qquad (3.26)$$

If $\hat{H}$ is an effective mean-field Hamiltonian, Eqs. (3.25) [or (3.26)] constitute the general (statistical) time-dependent Hartree-Bogoliubov boson equations. In this case, H and G depend on D^c and $\langle Z\rangle$. Nevertheless, the eigenvalues of ΠD_c and $M\Pi$ (and hence the entropy) are preserved by (3.25) and (3.26). The temporal evolution can thus be completely described by an optimal (time-dependent) Bogoliubov transformation (3.9).

IV. APPLICATION

We shall consider, as a simple and illustrative example, the case of a one-dimensional, one-particle system, described by a Hamiltonian (we set $m=1$) $\hat{H}=\frac{1}{2}\hat{P}^2+V(\hat{Q})$, with $\hat{P}$ and $\hat{Q}$ defined according to (3.16). The general case where the information deals with expectation values of arbitrary functions of $\hat{P}$ and $\hat{Q}$ is dealt with in the Appendix. Let us consider here a situation in which the available information deals with the expectation values of $\hat{H}$ and of a set of functions of the coordinate $F_i(\hat{Q})$. The corresponding (exact) statistical operator can be written as

$$\hat{\rho}=\exp(-\lambda_0-\beta\hat{H}')\ , \qquad (4.1)$$

with $\hat{H}'=\hat{H}+\sum_i(\beta_i/\beta)F_i(\hat{Q})$. This $\hat{\rho}$ constitutes a quite difficult object to deal with, even in this simple situation. One requires, first of all, the knowledge of the eigenvalues of $\hat{H}'$, and then, the determination of the particular values of the Lagrange parameters which adjust the available data. Hence, it is reasonable to consider a first-order description based on an approximate one-body density cperator of the form

$$\hat{\rho}=\exp(-\lambda_0-\beta\hat{h}')=\exp(-\lambda'_0-\lambda' b'^{\dagger}b')\ , \qquad (4.2)$$

where $\hat{h}'$ is an effective SP operator constructed as in (2.12), and b' a boson "quasiparticle" operator related to the "unperturbed" ones by [see (3.9)]

$$b'=xb+yb^{\dagger}+z\ , \qquad (4.3)$$

with $|x|^2-|y|^2=1$. Since the only relevant one-body mean values are $\langle\hat{P}^2\rangle$, $\langle\hat{Q}^2\rangle$, and $\langle\hat{Q}\rangle$, $\hat{h}'$ can be written as [cf. Eq. (2.12)]

$$\hat{h}'=\hat{P}^2/2+\sum_{j=1}^{2}\frac{\partial\langle V'(\hat{Q})\rangle}{\partial\langle\hat{Q}^j\rangle}\hat{Q}^j\ , \qquad (4.4)$$

where

$$V'(\hat{Q})=V(\hat{Q})+\sum_i(\beta_i/\beta)F_i(\hat{Q})\ .$$

Moreover, in this situation, x, y, and z are real, so that the Bogoliubov transformation (4.3) becomes equivalent to a scale factor plus a translation in $\hat{P}$ and $\hat{Q}$,[18] i.e., $\hat{Q}=t\hat{Q}'+q$, $\hat{P}=\hat{P}'/t$, with $t=x-y$, $q=-tz\sqrt{2}$. The basic elements for calculating expectation values with respect to (4.2) are [see (A10)–(A13) with $\delta=\xi=0$]

$$\sigma\equiv\langle(\hat{Q}-\langle\hat{Q}\rangle)^2\rangle=t^2(f+\tfrac{1}{2}),\quad \langle\hat{Q}\rangle=q\ , \tag{4.5a}$$

$$\langle\hat{P}^2\rangle=(f+\tfrac{1}{2})/t^2=(f+\tfrac{1}{2})^2/\sigma\ , \tag{4.5b}$$

where $f=\langle b'^\dagger b'\rangle$ is the average boson quasiparticle number. By recourse to Wick's theorem, it can be easily shown that

$$\langle(\hat{Q}-q)^{2n}\rangle=\sigma^n 2n!/(2^n n!)\ , \tag{4.6}$$

which corresponds, obviously, to a Gaussian density, i.e.,

$$\langle V(\hat{Q})\rangle=\int_{-\infty}^{+\infty}V(x)\exp[-(x-q)^2/2\sigma]/(2\pi\sigma)^{1/2}dx \equiv V(\sigma,q)\ . \tag{4.7}$$

A similar expression holds obviously for the expectation value of any function of $\hat{P}$, with σ replaced by $\langle\hat{P}^2\rangle$ (and $q=0$).

The mean-field equations in this situation, for fixed β and β_i's, can be found in the Appendix (see also Ref. 18). We shall focus here our attention upon the form that these equations adopt for (fixed) given mean values. Using (4.5), the mean energy can be written as

$$\langle\hat{H}\rangle=\tfrac{1}{2}(f+\tfrac{1}{2})^2/\sigma+V(\sigma,q)\equiv H\ . \tag{4.8}$$

It is evident from (4.5) that in this situation σ, q, and f suffice to determine $\hat{\rho}$. Therefore, the mean-field approach is able to adjust, in addition to the mean energy, the mean values of just two independent observables F_i (which would determine σ and q). Let us examine thus the case where $F_1=(\hat{Q}-\langle\hat{Q}\rangle)^2$, $F_2=\hat{Q}$. the operator (4.2) will be completely determined by σ, q, and H, with f given by

$$f=\{2\sigma[H-V(\sigma,q)]\}^{1/2}-\tfrac{1}{2}\ . \tag{4.9}$$

The corresponding effective multipliers (2.15a) will be

$$\beta=\frac{\partial S}{\partial H}=\alpha\sigma\ , \tag{4.10}$$

$$\beta_1=\frac{\partial S}{\partial\sigma}=\alpha(H-V-\sigma\partial V/\partial\sigma)\ , \tag{4.11}$$

$$\beta_2=\frac{\partial S}{\partial q}=-\alpha\sigma(\partial V/\partial q)\ , \tag{4.12}$$

where $\alpha=(\partial S/\partial f)/(f+\frac{1}{2})=\ln[(f+1)/f]/(f+\frac{1}{2})$.

In case σ and (or) q are unknown, the maximum entropy criterion implies $\beta_1=0$ and (or) $\beta_2=0$, and hence the equations

$$H-V-\sigma\partial V/\partial\sigma=0\ , \tag{4.13a}$$

$$\partial V/\partial q=0\ , \tag{4.13b}$$

which express the mean-field equations directly in terms of the available data [and not in terms of the unknown Lagrange multipliers, which do not appear explicitly in (4.13)]. Equations (4.13) are obviously equivalent to (A18)–(A19), after the appropriate replacements are made.

A physical self-consistent solution will exist wherever the value of f obtained from (4.9) is real and nonnegative. In case $\langle\hat{H}\rangle$ is the sole information, a solution will obviously exist only if H is greater than the ground-state energy of the zero temperature Hartree-Bose approximation,[17,18] obtained from $\frac{1}{8}-\sigma^2\partial V/\partial\sigma=0$, $\partial V/\partial q=0$.

The thermodynamic derivatives (2.20) can be explicitly calculated in this context using equations (A17)–(A19). For the derivatives with respect to β, we obtain (subscripts denoting partial derivatives)

$$\frac{\partial f}{\partial\beta}=[-a/\sigma+(\beta a/\sigma^2)\sigma_\beta]/[b-(\beta a/\sigma^2)\sigma_f] \tag{4.14}$$

and

$$\partial\sigma/\partial\beta=\sigma_\beta+\sigma_f(\partial f/\partial\beta),\quad \partial q/\partial\beta=q_\beta+q_\sigma(\partial\sigma/\partial\beta)\ ,$$

where

$$\begin{aligned}\sigma_\beta&=-[V_\sigma-\tfrac{1}{2}(a/\sigma)^2-V'_{q\sigma}V_q/V'_{qq}]/(D\beta)\ ,\\ \sigma_f&=(a/\sigma^2)/D\ ,\\ D&=V'_{\sigma\sigma}-(V'_{\sigma q})^2/V'_{qq}+a^2/\sigma^3\ ,\end{aligned} \tag{4.15}$$

with

$$a=(f+\tfrac{1}{2}),\quad b=\beta/\sigma+1/[f(f+1)]\ .$$

Using (4.8) and (4.14), $\partial\langle\hat{H}\rangle/\partial\beta$ can be calculated. It is to be remarked that β and the remaining Lagrange parameters entering (4.14) and (4.15) are inferred from the available data by means of (4.10)–(4.12). In case $\langle\hat{H}\rangle$ is the only information ($\beta_i=0$), we obtain

$$\frac{\partial\langle\hat{H}\rangle}{\partial\beta}=-(a/\sigma)^2/[b-\beta(a/\sigma^2)^2/D]\ , \tag{4.16}$$

which for a quadratic potential ($V_{\sigma\sigma}=V_{q\sigma}=0$) reduces to the exact expression $-f(f+1)(a/\sigma)^2$. Equation (4.16) provides a much more accurate estimate of the exact energy fluctuation $\langle\hat{H}^2\rangle-H^2$ than a direct evaluation by means of Wick's theorem (see Sec. V).

All the above formulas can be straightforwardly generalized to a d-dimensional case. In the special case where $\hat{H}=\frac{1}{2}\hat{P}^2+V(\hat{R})$ (with $\hat{R}^2=\sum_{i=1}^d\hat{Q}_i^2$, $\hat{P}^2=\sum_{i=1}^d\hat{P}_i^2$), and the functions F_i depend solely on $\hat{R}$ [O(d)-invariant case], we have instead of (4.2),

$$\hat{\rho}=\exp\left[-\lambda_0-\lambda'\sum_{i=1}^d b_i'^\dagger b_i'\right]\ , \tag{4.17}$$

where all b_i' are equally related to the unperturbed operators by (4.3). In this situation we have obviously $\langle\hat{Q}_i\rangle=0$, and (4.5) must be replaced by

$$\langle\hat{R}^2\rangle=dt^2(f+\tfrac{1}{2})\equiv d\sigma,\quad \langle\hat{P}^2\rangle=d(f+\tfrac{1}{2})/t^2\ .$$

Finally, we obtain instead of (4.7)

$$\langle V(\hat{R})\rangle=C\int_0^\infty V(r)\{\exp[-r^2/2\sigma]/(2\pi\sigma)^{d/2}\}r^{d-1}dr\ , \tag{4.18}$$

where $C=2\pi^{d/2}/\Gamma(d/2)$ is the area of the d-dimensional unit sphere. Hence, things behave as in the one-dimensional case, with an effective even potential

$$V_{\rm eff}(\hat{r})=CV(\hat{r})\hat{r}^{\,d-1}/[2d(2\pi\sigma)^{(d-1)/2}] \ . \qquad (4.19)$$

In this way,

$$\langle \hat{H}\rangle/d=\tfrac{1}{2}(f+\tfrac{1}{2})^2/\sigma+V_{\rm eff}(\sigma) \ ,$$

and expressions (4.9)–(4.11) can be utilized.

V. RESULTS

As a numerical example, we have chosen the case of a general quartic anharmonic oscillator

$$V(\hat{Q})=A\hat{Q}^{2}+B\hat{Q}^{2}+C\hat{Q}^{4} \ , \qquad (5.1)$$

in which case, using (4.7),

$$V(\sigma,q)=A(q^2+\sigma)+B(q^3+3q\sigma)+C(q^4+6q^2\sigma+3\sigma^2) \ . \qquad (5.2)$$

This type of potential has attracted a great deal of work during the past years (see, for instance, Refs. 25–27, 17, and 18) due to its relevance for studying molecular vibrations and nonlinear quantum field theories. Exact statistical averages for this Hamiltonian can be calculated by means of a diagonalization in a truncated optimized Hartree basis.[25,26,18]

Figures 1–5 depict results corresponding to $B=0$, $A=\frac{1}{2}$, and two different values of C (obviously, by means of a proper scaling, one can bring one of the coupling constants to a fixed desired value; this choice corresponds to an unperturbed oscillator energy $\hbar\omega=1$). We have examined two different situations, in which the mean-field solution (4.2) is constructed with information concerning $\langle\hat{H}\rangle$ (case 1), and $\{\langle\hat{H}\rangle,\langle\hat{Q}^2\rangle\}$ (case 2). These expectation values were obtained from the exact averages corresponding to the system heated at a temperature T, in which case the exact multipliers are $\beta=1/T$ (we set Boltzmann constant $k_B=1$), $\beta_1=0$. In this situation, q vanishes ($\beta_2=0$). Quantities are plotted in terms of the exact temperature.

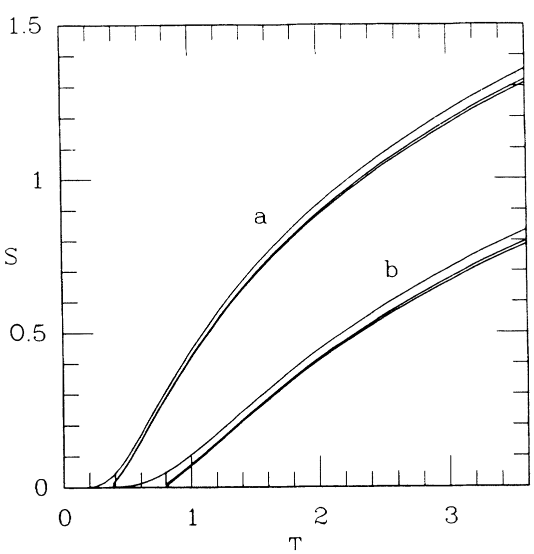

FIG. 1. Entropy (in units of Boltzmann constant k_B) vs temperature (in units of $\hbar\omega/k_B$ in all figures, with ω the frequency of the unperturbed harmonic oscillator) for the anharmonic potential (5.1), for $A=\frac{1}{2}$, $B=0$, and $C=1$ (curves a), $C=10$ (curves b). In each group, the upper curve corresponds to exact results, the intermediate one to the mean-field approach (1), constructed with information about $\langle\hat{H}\rangle$, and the lower to the constrained mean-field approach (2), constructed with $\langle\hat{H}\rangle$ and $\langle\hat{Q}^2\rangle$ (see Sec. V).

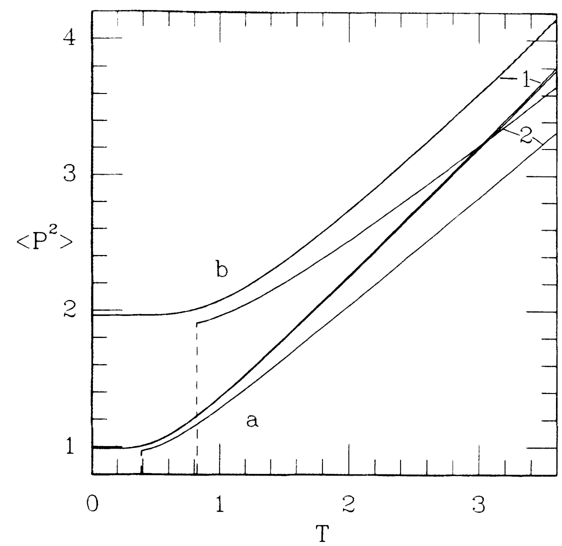

FIG. 2. Predicted value of $\langle\hat{P}^2\rangle$ (in units of $m\hbar\omega$). Details are similar to those of Fig. 1. Curves 1 depict exact results and those given by the mean-field approach (1), almost undistinguishable in the scale of the figure, whereas curves 2 correspond to the mean-field approach (2). Dashed lines indicate threshold points of mean-field solutions.

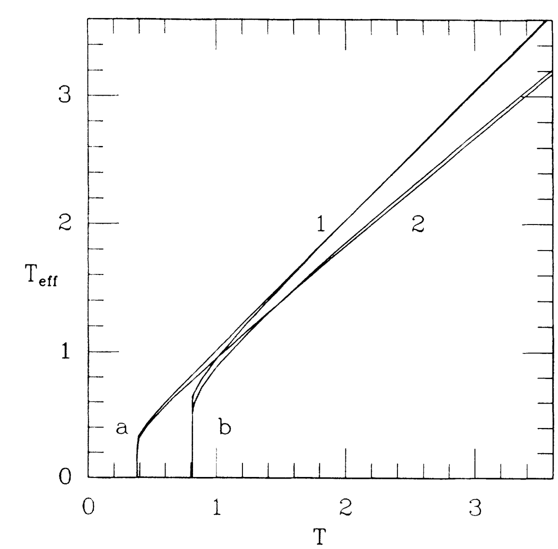

FIG. 3. Effective temperatures. Curves 1 and 2 depict the inverse of the multiplier (4.10) according to the respective mean-field solution.

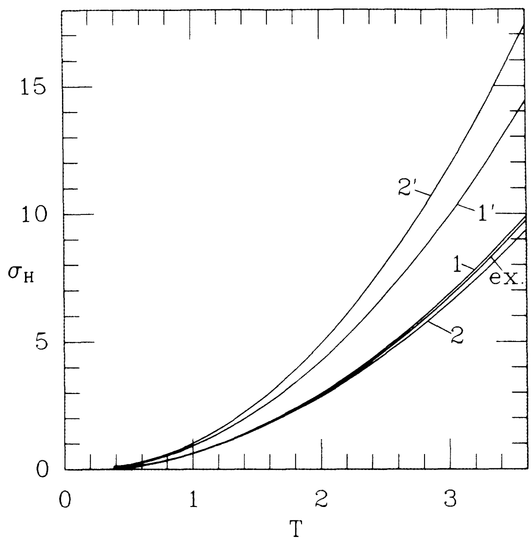

FIG. 4. The prediction of the energy fluctuation $\sigma_H = \langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2$ [in units of $(\hbar\omega)^2$] for the case $A=\frac{1}{2}$, $B=0$, $C=1$. Curve 1 depicts results given by the mean-field approach (1), employing expression (4.16), which lies only slightly above the exact prediction (denoted by ex.). Curve 2 depicts the corresponding results of approach (2) [employing (4.14) and (4.15)]. Curves 1' and 2' show the respective direct evaluation using Wick's theorem.

Results indicate that in case 1, the approach (4.2) is quite reliable, independent of the value of C, yielding very accurate inferred values of S, $T_{\text{eff}} = 1/\beta$ [cf. (4.10)] and one-body observables (such as $\hat{P}^2$), for temperatures above the threshold value. Moreover, the prediction of fluctuations, such as $\sigma_H = \langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2$, is also in excellent agreement with the exact results, provided the corresponding thermodynamic derivative (cf. 4.16) is utilized. However, as can be seen in Fig 4, the results yielded by these derivatives differ appreciably from those given by a direct evaluation of the fluctuation (by means of Wick's theorem), which yields a finite value even at the threshold point, where the inferred temperature vanishes.

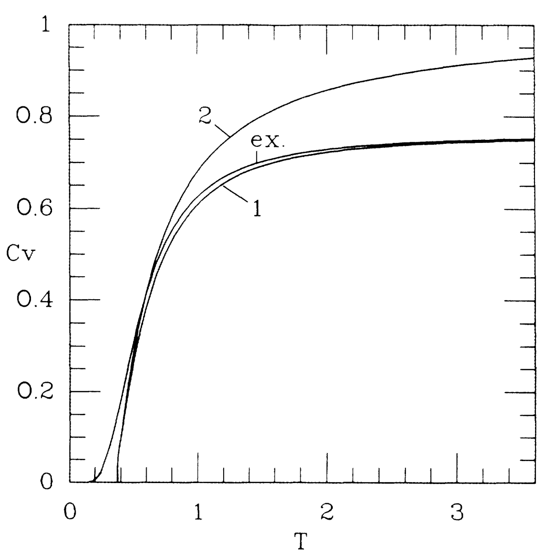

FIG. 5. Predicted value of the specific heat $C_v = \partial \langle H \rangle / \partial T$ (in units of k_B) for the same case of Fig. 4, according to exact (ex.) and mean-field results (1) and (2) [constructed using (4.14)–(4.16) and the inferred effective temperature].

The specific heat $C_v = -\beta^2 \partial \langle \hat{H} \rangle / \partial \beta$, shown in Fig. 5, is constructed with the inferred value of β (4.10), and also lies quite close to the exact value. As $T \to \infty$, the mean-field approach yields the exact leading order of first-order averages (not fluctuations) and thermodynamic derivatives, providing the correct limiting value of a quartic oscillator $C_v = \frac{3}{4}$ (see Ref. 17 for results in the conventional formulation). For sufficiently high temperatures, the effective inferred temperature is slightly greater than the exact one. At the threshold point f vanishes, while its derivative with respect to H remains finite, implying thus an infinite initial slope of the inferred entropy and temperature.

On the other hand, in case 2, where $\langle \hat{Q}^2 \rangle$ is clamped at that value corresponding to the exact average, the approach (4.2) obviously yields a lower value of the entropy (due to a higher amount of available information), and in general poorer estimates of the remaining quantities, due to a reduced number of degrees of freedom. The constrained description also becomes feasible at a slightly higher value of H. The additional multiplier β_1 (4.11) does not vanish, and is negative, since the unconstrained solution underestimates the value of $\langle \hat{Q}^2 \rangle$. β_1 diverges at the threshold point, although the "chemical potential" β_1/β remains finite. The finite value of β_1 implies a departure from the exact behavior in most inferred quantities, even as T increases. We should remark that just a single physical self-consistent solution exists in both cases 1 and 2, for given mean values of $\langle \hat{H} \rangle$ (and $\langle \hat{Q}^2 \rangle$ in 2), and that no "phase transitions" are encountered.

The addition of a small cubic term to the potential does not alter the previous conclusions. The picture varies dramatically however, in case the coefficient B in (5.1) increases, or A becomes negative (bistable case). For instance, Fig. 6 depicts results for the inferred entropy for $C=1$, $B=-0.1$, and $A=-\frac{3}{2}A_c$, in which case the potential possesses two wells of different depth [$A_c = (243C^2/2)^{1/3}$ is the critical value such that for $A < -A_c$, the zero-temperature Hartree approximation exhibits a displaced symmetry-breaking solution[18] for $B=0$, centered in either of the wells].

On this quasibistable situation, the agreement between mean field and exact predictions is only of a qualitative character. Up to five different types of unconstrained mean-field solutions (information about $\langle \hat{H} \rangle$ only) may exist for the same value of $\langle \hat{H} \rangle$, starting at different threshold energies. They correspond to three maxima of S and two intermediate saddle points, if S is viewed as a function of q [with σ and f expressed in terms of q by means of (4.9) and (4.13b)]. Only the solutions corresponding to a local maximum are exhibited in Fig. 6.

The solution labeled as b corresponds to a density localized at the deepest well of the potential (q close to the value which minimizes V), and is the first to appear.

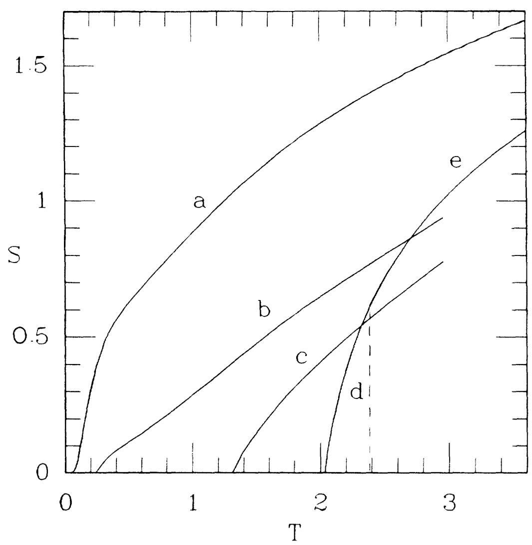

FIG. 6. Entropies for the bistable cubic case, with $A=-\frac{3}{2}A_c$, $B=-0.1$, $C=1$. Curve a denotes the exact entropy, whereas b, c, and e correspond to different mean-field solutions (constructed with $\langle \hat{H} \rangle$ alone). Curve d depicts the mean-field solution built up with $\langle \hat{H} \rangle$ and $\langle \hat{Q} \rangle$. The dashed line indicates the onset of solution e, which starts at a nonvanishing entropy.

Solution c corresponds to a density localized at the shallow well, whereas solution e corresponds to a density almost centered at the origin (q small). This solution starts with a nonvanishing value of S. The solution d depicts a constrained situation, in which the value of $\langle \hat{Q} \rangle$ is supplied in addition to that of $\langle \hat{H} \rangle$. This solution practically coincides with e above the threshold value, yielding thus a very small value of β_2 (4.12), except at the threshold point. For higher values of T, solutions b and c disappear, and only d and e remain. A fully constrained mean-field solution, in which H, σ, and q are to be adjusted, does not exist for the shown range.

In the unconstrained (inferred) picture, the solution yielding the highest value of the entropy, changes from b to e at the critical value $H_c \cong -0.553$, which corresponds to the (exact) temperature $T_c \cong 2.7$. Therefore, a "phase transition" arises as $\langle \hat{H} \rangle$ increases, in which the system, according to the unconstrained mean-field description, would "jump" from a situation appropriately described by a solution localized in the deepest well, to that corresponding to a near symmetric solution. This abrupt transition is correlated with the sudden (albeit smooth) decrease in the value of $\langle \hat{Q} \rangle$, as T increases, exhibited by the exact density at low temperatures.[18]

Within the IT context, one can assert that a mean-field treatment of the available information amplifies a nearly critical behavior. Notice however that the inclusion of additional information may induce one to choose a solution different from that yielding the highest entropy. For instance, the addition of information concerning $\langle \hat{Q} \rangle$ leads one to choose solution e even before T_c is reached, since this solution practically coincides with the constrained solution d.

VI. CONCLUSIONS

The problem of (many) interacting particles exhibits severe difficulties if one wishes to deal with exact descriptions. This is true in classical mechanics and to a much larger degree in quantum mechanics. The Information-theory approach, although widely regarded as an extremely convenient and illuminating one, cannot escape the restrictions any one faces in looking for exact solutions. The physicist is to be satisfied with approximate treatments, which either contain just a few of the essential ingredients of the "real" problem or, at the very least, contain within themselves a criterion for their validity. In this respect, IT can often lead to rewarding results.

Mean-field methods constitute, in many instances, the standard basic approach to the quantum many-body problem, and we have here reformulated them from a particular IT viewpoint. In this view, they arise as a consequence of building up the density operator with a special set of observables chosen by us, and not forced upon as by the nature of the available information, as is the case of the (by now) classical formalism of Jaynes. The maximum entropy principle leads then, in a natural fashion, to a general self-consistent approximation for statistical operators, which contains, as a very particular case, the (previously) known statistical mean-field treatments. The approach provides a generalization of thermal many-body mean-field theories, inserting them within a completely general statistical context, and allowing for the possibility of dealing with arbitrary ensembles and trail density operators.

The IT framework allows also for a new interpretation of these mean-field descriptions, viewing them as an approximate processing of the available data in terms of a particular choice of relevant operators. In this sense, we have examined the feasibility of the construction of mean-field descriptions directly from the knowledge of a given set of expectation values, obtaining in this way different lower bounds for the exact entropy, and a set of effective temperatures and Lagrange multipliers. Statistical mean-field descriptions of Bose systems have been examined within this general context.

Finally, different possibilities for the approximate inference of fluctuations from the available information have been studied, and general formulas for obtaining thermodynamic derivatives within the general self-consistent scheme have been derived. We have shown, by recourse to a numerical example, that statistic mean-field approaches are able to predict, in general, quite accurate values of fluctuations, provided a thermodynamic expression is utilized instead of a direct evaluation.

ACKNOWLEDGMENTS

R.R. acknowledges support by the Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC), as Investibador. A.P. was supported by the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) de Argentina.

APPENDIX

For the sake of completeness, we shall consider explicitly the statistical operator corresponding to the general one-dimensional case, where there is just a single accessible boson state,

$$\hat{\rho}=\exp\{-\lambda_0-\lambda b^\dagger b-\tfrac{1}{2}[\gamma b^2+\gamma^*(b^\dagger)^2]-\eta b^\dagger-\eta^* b\} \tag{A1a}$$

$$=\exp(-\lambda_0'-\lambda' b'^\dagger b') , \tag{A1b}$$

where

$$\lambda'=\lambda[1-(|\gamma|/\lambda)^2]^{1/2} , \tag{A2}$$

and $b'=xb+yb^\dagger+d$, with

$$|x|^2=\tfrac{1}{2}(\lambda/\lambda'+1)=1+|y|^2, \quad 2\lambda' y^* x=\gamma ,$$

and $d=(x\eta-y\eta^*)/\lambda'^2$. The normalizable case corresponds thus to $|\gamma|<|\lambda|$ and $\lambda>0$. The matrices D^c and M [(3.2) and (3.7)] are now two dimensional and the relation (3.13a) between them can be written explicitly as

$$\Pi D^c=(M\Pi)(f+\tfrac{1}{2})/\lambda'-\tfrac{1}{2}I ,$$

i.e.,

$$\langle b^\dagger b\rangle-|\langle b\rangle|^2=\lambda(f+\tfrac{1}{2})/\lambda'-\tfrac{1}{2} , \tag{A3a}$$

$$\langle b^2\rangle-\langle b\rangle^2=-\gamma^*(f+\tfrac{1}{2})/\lambda' , \tag{A3b}$$

where

$$f=\langle b'^\dagger b'\rangle=[\exp(\lambda')-1]^{-1} , \tag{A3c}$$

and

$$\langle b\rangle=-(\lambda\eta-\gamma^*\eta^*)/\lambda'^2 . \tag{A4}$$

In order to obtain the Lagrange parameters from the available data, we can employ again (A3) and (A4) but expressing f and $\lambda'=\ln[f/(f+1)]$ in terms of the expectation values, i.e.,

$$(f+\tfrac{1}{2})^2=(\langle b^\dagger b\rangle-|\langle b\rangle|^2+\tfrac{1}{2})^2 - |\langle b^2\rangle-\langle b\rangle^2|^2)\equiv\Delta . \tag{A5}$$

Consequently, since $f\geq 0$, the statistical operator can be constructed only if $\Delta\geq\frac{1}{4}$, which represents the uncertainty relation in terms of the mean values of $b^\dagger$, b. The ensuing maximum entropy $S=(f+1)\ln(f+1)-f\ln(f)$ depends thus only upon Δ, vanishing for $\Delta=\frac{1}{4}$.

When written in terms of $\hat{P}$ and $\hat{Q}$, (A1) reads

$$\hat{\rho}=\exp(-\lambda_0'-\alpha\hat{Q}^2-\beta\hat{P}^2-\delta\hat{L}-\zeta\hat{Q}-\xi\hat{P}) \tag{A6}$$

$$=\exp[-\lambda_0'-\tfrac{1}{2}\lambda'(\hat{Q}'^2+\hat{P}'^2)] , \tag{A7}$$

where $\hat{L}=\hat{P}\hat{Q}+\hat{Q}\hat{P}$, $\lambda'=2(\alpha\beta-\delta^2)^{1/2}$, and [choosing phases such that $\mathrm{Im}(y-x)=0$]

$$\hat{Q}'=(\hat{Q}-q)[\lambda'/(2\beta)]^{1/2} , \tag{A8}$$

$$\hat{P}'=[2\beta/\lambda']^{1/2}[(\hat{P}-p)+\delta(\hat{Q}-q)/\beta] , \tag{A9}$$

where

$$q=\langle\hat{Q}\rangle=-2(\beta\zeta-\delta\xi)/\lambda'^2 , \tag{A10a}$$

$$p=\langle\hat{P}\rangle=-2(\alpha\xi-\delta\zeta)/\lambda'^2 . \tag{A10b}$$

The relations (A3) correspond to

$$\sigma_q\equiv\langle\hat{Q}^2\rangle-\langle Q\rangle^2=2\beta(f+\tfrac{1}{2})/\lambda' , \tag{A11}$$

$$\sigma_P\equiv\langle\hat{P}^2\rangle-\langle\hat{P}\rangle^2=2\alpha(f+\tfrac{1}{2})/\lambda' , \tag{A12}$$

$$\sigma_L\equiv\langle\hat{L}\rangle-2\langle\hat{Q}\rangle\langle\hat{P}\rangle=-4\delta(f+\tfrac{1}{2})/\lambda' . \tag{A13}$$

Besides, the quantity Δ (A5) can be written now as

$$\Delta=\sigma_p\sigma_q-\sigma_L^2/4 . \tag{A14}$$

It is clear from (A14) that (A1) can be constructed only if the available information fulfills the uncertainty relation. The minimum uncertainty corresponds to the pure case ($S=0$).

In the particular situation where $\xi=\delta=0$, (A10)–(A13) lead to (4.5), if we set $t^2=(\beta/\alpha)^{1/2}$. The normalizable case corresponds to $\delta^2<\alpha\beta$. For $\lambda'=0$, both f and λ_0 diverge, although some mean values converge. For instance, if δ and α are vanishing quantities of the same order and β remains finite, $\hat{\rho}$ tends to a free particle density, and it can be seen from (A11) and (A12) that σ_q diverges as $1/(2\alpha)$, whereas $\sigma_p=1/(2\beta)$, $\sigma_L=-\delta/(\alpha\beta)$. The system possesses in this case an infinite entropy, corresponding to an infinite fluctuation of the coordinate.

The general mean-field equations for the one-dimensional case (for fixed multipliers β_i) can now be straightforwardly obtained from (A3) and (A4) or (A10)–(A13) by expressing the Lagrange parameters in terms of the one-body mean values (or vice versa) by means of (2.7), which can be easily accomplished using Wick's theorem. We obtain in general a set of five coupled nonlinear real equations. The matrix B (2.9) [and with it the thermodynamic derivatives (2.20) and (2.21)] can be easily calculated in the general case by deriving relations (A3) and (A4), or alternatively (A10)–(A13).

In the particular situation of Sec. IV, the system reduces to a set of two equations in σ_q and q, namely, (A10a) and (A11), with

$$\alpha=\beta' \frac{\partial\langle \hat{V}'\rangle}{\partial\langle \hat{Q}^2\rangle}=\beta' \frac{\partial V'(\sigma_q,q)}{\partial\sigma_q} , \tag{A15}$$

$$\zeta=\beta' \frac{\partial\langle \hat{V}'\rangle}{\partial\langle \hat{Q}\rangle}=\beta' \left[\frac{\partial}{\partial q}-2q\frac{\partial}{\partial\sigma_q}\right]V'(\sigma_q,q) , \tag{A16}$$

$$\lambda'=\beta' \left[2\frac{\partial V'(\sigma_q,q)}{\partial\sigma_q}\right]^{1/2} , \tag{A17}$$

and $\beta=\beta'/2$, with f given by (A3c) and β' the multiplier associated with $\hat{H}'$. The ensuing final equations can be rewritten in terms of f, σ, and q:

$$\frac{\partial V'(\sigma_q,q)}{\partial\sigma_q}-\frac{(f+\frac{1}{2})^2}{2\sigma_q^2}=0 , \tag{A18}$$

$$\frac{\partial V'(\sigma_q,q)}{\partial q}=0 , \tag{A19}$$

which could have also been obtained through a direct minimization of S'.

[1]C. Shannon, Bell Syst. Tech. J. **27**, 379 (1948); **27**, 623 (1948).
[2]L. Brillouin, *Science and Information Theory* (Academic, New York, 1956).
[3]E. T. Jaynes, Phys. Rev. **106**, 620 (1957); **108**, 171 (1957).
[4]A. Katz, *Principles of Statistical Mechanics* (Freeman, San Francisco, 1967); A. Hobson, *Concepts in Statistical Mechanics* (Gordon and Breach, London, 1971).
[5]R. Balian and N. Balazs, Ann. Phys. (N.Y.) **179**, 143 (1987).
[6]Y. Alhassid and R. Levine, Phys. Rev. A **18**, 89 (1978); Phys. Rev. C **20**, 1775 (1979).
[7]N. Tishby and R. Levine, Phys. Rev. A **30**, 1477 (1984).
[8]R. Balian, Y. Alhassid, and H. Reinhardt, Phys. Rep. **131**, 1 (1986).
[9]H. Reinhardt, R. Balian, and Y. Alhassid, Nucl. Phys. A **422**, 349 (1984).
[10]R. Balian and M. Veneroni, Ann. Phys. (N.Y.) **164**, 334 (1985); **174**, 229 (1987).
[11]D. Otero, A. Plastino, A. Proto, and G. Zannoli, Phys. Rev. A **26**, 1209 (1982); Z. Phys. A **316**, 323 (1984); J. Aliaga, D. Otero, A. Plastino, and A. Proto, Phys. Rev. A **35**, 2304 (1987).
[12]R. Rossignoli and A. Plastino, Phys. Rev. C **36**, 1595 (1987); Phys. Rev. C **40**, 1798 (1989).
[13]D. J. Thouless, *The Quantum Mechanics of Many-Body Systems* (Academic, New York, 1961).
[14]A. Goodman, Nucl. Phys. A **352**, 30 (1981); **352**, 45 (1981).
[15]M. Brack and P. Quentin, Nucl. Phys. A **361**, 35 (1981).
[16]K. Tanabe, K. Sugawara Tanabe, and H. J. Mang, Nucl. Phys. A **357**, 20 (1981); **357**, 45 (1981).
[17]J. Nuñez, A. Plastino, and R. Rossignoli, Phys. Rev. D **33**, 1709 (1986).
[18]J. Nuñez, A. Plastino, and R. Rossignoli, J. Phys. A **21**, 729 (1988).
[19]G. Bozzolo, A. Plastino, and J. Ferrante, Z. Phys. **A 333**, 119 (1989); H. G. Miller, R. M. Quick, G. Bozzolo, and J. P. Vary, Phys. Lett. **168B**, 13 (1986); R. Rossignoli and A. Plastino, Phys. Rev. C **32**, 1040 (1985).
[20]R. Gilmore, *Catastrophe Theory for Scientists and Engineers* (Wiley, New York, 1981); D. H. Feng, R. Gilmore, and L. M. Narducci, Phys. Rev. C **19**, 1119 (1979).
[21]H. G. Kümmel, Nucl. Phys. A **317**, 199 (1979); M. Cambiaggio and A. Plastino, *ibid.* **403**, 365 (1983).
[22]R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957); Rep. Prog. Phys. **29**, 255 (1966).
[23]D. Otero, A. Plastino, and A. Proto, Phys. Rev. A **38**, 918 (1988).
[24]P. Ring and P. Schuck, *The Nuclear Many Body Problem* (Springer-Verlag, New York, 1980).
[25]C. Esebbag, J. Nuñez, A. Plastino, and G. Bozzolo, Phys. Rev. D **32**, 522 (1985); J. Phys. A **18**, (1985).
[26]C. S. Hsue and J. Chern, Phys. Rev. D **29**, 643 (1984); C. S. Hsue, Phys. Rev. A **33**, 1392 (1986).
[27]P. K. Patnaik, Phys. Rev. D **33**, 3145 (1986); C. M. Bender and K. A. Milton, Phys. Rev. D **34**, 3149 (1986).