Journal of Statistical Physics, Vol. 36, Nos. 3/4, 1984

On the Validity of the Inverse Conjecture in Classical Density Functional Theory

J. T. Chayes[1,2,3] and L. Chayes[1,2,3]

Received December 6, 1983; revised March 9, 1984

It is shown that the basic assumptions of the classical density functional approach are rigorously correct for H-stable systems in the grand canonical ensemble. Moreover, it is established that the set of all single-particle densities is convex. These results are derived by providing necessary and sufficient conditions for the solution of the classical inverse problem for single-particle densities. Analogous results are obtained for the solution of the higher-order correlation inverse problem, and the ramifications of these results for the validity of two-body decomposition of forces are discussed.

KEY WORDS: Inverse problem; density functional theory; V-representability; two-body decomposition.

1. INTRODUCTION

The classical inverse problem is the question of whether a preprescribed function corresponds to the equilibrium single-particle density in a given system under the influence of an external potential. This problem was thoroughly addressed in Ref. 1 for both the canonical and grand canonical ensembles. Of notable absence was a substantive result for systems with hard-core interactions.

In this paper, we examine the inverse problem for H-stable systems in the grand canonical ensemble. Since hard cores are typically introduced into

[1] Present address: Department of Physics, Harvard University, Cambridge, Massachusetts 02138.

[2] Department of Physics, Princeton University, Princeton, New Jersey 08544.

[3] Research supported in part by the National Science Foundation under grants PHY-8116101 A01 (J.T.C.), PHY-8301493 (L.C.) and PHY-8203669 (J.T.C. and L.C.).

0022-4715/84/0800-0471$03.50/0

822/36/3-4-12*

model systems to guarantee H-stability, the class of problems analyzed here, together with those treated in Ref. 1, includes essentially all cases of physical relevance (in the grand canonical distribution).

The inverse problem arises in the density functional approach to the theory of nonuniform fluids.[2-8] There one is concerned with the relationship between three kinds of functions: probability distributions, one-body densities, and external potentials. The standard approach,[4,5] reviewed in Section 2, entails the study of certain functionals whose extremum values correspond to thermodynamic potentials of the system. Although these functionals are naturally defined on a set of probability distributions, it has been found useful to view them as functionals of single-particle densities. Hence the density functional approach requires that every density ρ, which is the single-particle reduction of a probability distribution F, is the equilibrium single-particle density at some external potential U.

The inverse problem is the question of whether the map $\rho \mapsto U$ exists for some *given* density. For the purposes of density functional theory, one is concerned only with those densities which are obtained as single-particle reductions of probability distributions. In Section 3, we show that the necessary and sufficient condition for existence of the map $\rho \mapsto U$ in H-stable systems is that ρ is the single-particle reduction of an admissible probability distribution. (Our criterion of admissibility is simply that the positive part of one of the corresponding thermodynamic functionals is finite.) Our results imply that the functionals introduced in Section 2 achieve minima over the set of all probability distributions which reduce to a given density. This assertion is the cornerstone of the classical density functional approach.[4,5]

The above results demonstate that the basic assumptions of density functional theory are rigorously correct for H-stable systems. However, in order to proceed with the density functional approach, it is necessary to establish certain properties of the set of all single-particle densities. For H-stable systems in the grand canonical ensemble, it was shown in Ref. 1 that this set is open (in the L^1 strong topology). In Section 3, we prove that the set is convex. Furthermore, we show that the density functionals introduced in Section 2 are convex functionals on this set. As stressed in Ref. 9 (see also Ref. 10), convexity is a crucial property in variational theories of this sort. Moreover, in various applications of classical density functional theory,[2-4,6] it is assumed that there is a one-parameter family which continuously extrapolates between any two admissible densities. Convexity guarantees that such a family may be constructed. In this regard, it should be noted that convexity fails for certain formulations of the (quantum) density functional approach to the inhomogeneous electron gas.[9]

We should like to point out that, although the conditions specified in

Section 3 for the solution of the $\rho \mapsto U$ problem are clearly necessary, they are not verifiable in the sense of the sufficient conditions imposed in Ref. 1. Thus, while our analysis closes the question of the inverse problem for classical density functional theory of H-stable systems, it does not provide a complete characterization of all possible single-particle densities. The authors feel that those aspects of the inverse problem which do not concern density functional theory are nevertheless interesting and important.

The techniques developed in Section 3 easily extend to higher-order generalizations of the inverse problem. For example, one may ask whether a given function of two variables corresponds to the two-point function of a system under the action of some two-body potential. The inverse conjecture for two-point functions has recently been used in extensions of the density functional approach.[11]

In contrast to the one-body problem, the inverse problem for two-point functions in a noninteracting system is not explicitly solvable. Indeed, in this case, the existence of a solution implies that a given two-point function is realizable by the application of a two-body potential. As such, a proof of the inverse conjecture even in this special case constitutes a nontrivial statement on the validity of two-body decomposition.

In Section 4, we examine the m-point inverse conjecture in the truncated grand canonical ensemble. It is shown that the inverse problem possesses a unique solution for any log summable m-point function obtained as an m-particle reduction of a probability distribution.

2. DENSITY FUNCTIONAL THEORY

Although the density functional approach has typically been used for systems defined on $\mathbb{R}^d$, there is no difficulty in extending the formulation to the general setting introduced in Ref. 1. Let $x \in \Lambda$ denote *all* coordinates of a given particle. We need only assume that the corresponding measure space $\langle \Lambda, dx \rangle$, henceforth called the *single-particle space*, is σ-finite. All spaces of physical relevance are σ-finite. Thus the results derived here, as in Ref. 1, may be applied to lattice or continuum systems in any dimension. Moreover, the space, which need not be Euclidean, may include momenta or internal coordinates (e.g., intrinsic spin). We denote by $\int dx$ all relevant summations and integrations.

The N-particle state is defined on the product space $\langle \Lambda^N, d^N x \rangle$. The measure space for the grand canonical ensemble is obtained by taking the direct sum:

$$\langle \Lambda, d\overline{\mathbf{X}} \rangle = \bigoplus_{N=0}^{\infty} \left\langle \Lambda^N, \frac{1}{N!} d^N x \right\rangle; \qquad \langle \Lambda^0, d^0 x \rangle \equiv 1 \tag{2.1}$$

A *system* in the grand canonical ensemble is specified by a family of interactions

$$\mathbf{W} = (W_N : \Lambda^N \to \mathbb{R}^* \mid N = 1, 2,...); \qquad \mathbb{R}^* \equiv \{\mathbb{R} \cup \infty\} \tag{2.2}$$

The function $W_N = W_N(x_1,..., x_N)$, which is assumed to be measurable, represents the interaction of the N-particle state in the *absence* of an external field. For convenience, we absorb the chemical potential, $-N\mu$, into the definition of W_N and take inverse temperature $\beta = 1$. As in Ref. 1, W_N need not be endowed with any particular symmetry (e.g., exchange), nor bear any functional relationship to W_M, $M \neq N$ (e.g., two-body decomposition).

For each N, the measurable set

$$Q_N = \{(x_1,..., x_N) \in \Lambda^N \mid W_N(x_1,..., x_N) = \infty\} \tag{2.3}$$

represents the hard-core excluded region. Henceforth, all equalities shall only be assumed to hold almost everywhere on $\Lambda^N \backslash Q_N$.

The grand canonical average of a family of functions $f = (f_N : \Lambda^N \to \mathbb{R}^+ \mid N = 1, 2,...)$ is defined by

$$\langle f \rangle_{\mathbf{W}} = 1 + \sum_{N=1}^{\infty} \frac{1}{N!} \int f_N \, e^{-W_N} \, d^N x \tag{2.4}$$

Any family f satisfying $\langle f \rangle_{\mathbf{W}} < \infty$ defines a *probability distribution* $F = (F_N \mid N = 0, 1, 2,...)$ on $\langle \Lambda, d\bar{\mathbf{X}} \rangle$ given by

$$F_0 = [\langle f \rangle_{\mathbf{W}}]^{-1}, \qquad F_N = f_N \, e^{-W_N} [\langle f \rangle_{\mathbf{W}}]^{-1} \tag{2.5}$$

Note that the functions F_N are automatically zero on the sets Q_N, and thus respect the existing hard-core structure.

Consider the family of product functions given by $(\prod_{i=1}^N e^{-U(x_i)} \mid N = 1, 2,...) \equiv \pi(e^{-U})$ for some measurable function U: $\Lambda \to \mathbb{R}^*$. Clearly, U has the interpretation of an *external potential.* The average $\langle \pi(e^{-U}) \rangle_{\mathbf{W}}$ is the partition function of the system at external potential U, and shall be denoted by $\Xi(U)$. If $\Xi(U) < \infty$, the family of product functions defines a *Gibbs distribution*,

$$G_U = \left(\frac{1}{\Xi(U)} e^{-W_N} \prod^N e^{-U} \mid N = 0, 1, 2,... \right)$$

A *one-body density* is defined to be the (symmetrized) single-particle reduction of a probability distribution F:

$$\rho_F(y) = \sum_{N=1}^{\infty} \frac{1}{N!} \sum_{j=1}^{N} \int F_N(x_1,..., x_{j-1}, y, x_{j+1},..., x_N) \, dx_1 \cdots \widehat{dx_j} \cdots dx_N \tag{2.6}$$

Here we shall only consider distributions with finite first moments, so that the corresponding one-body densities represent systems with finite expected particle numbers (i.e., $n_F \equiv \int \rho_F(y)\,dy < \infty$). We denote the set of all such distributions by $\mathscr{D}$.

We shall use the notation ρ_U, rather than ρ_{G_U}, for single-particle reductions of Gibbs distributions. These special one-body densities are called *single-particle densities*. The set of external potentials for which a single-particle density exists [i.e., $\Xi(U) < \infty$] and is in L^1 (i.e., $\int \rho_U < \infty$) shall be denoted by $\mathscr{U}_{\mathbf{W}}$.

The starting point of density functional theory is the introduction of functionals of probability distributions, whose extremum values correspond to thermodynamic potentials of the system. The principal functionals are

$$\mathfrak{F}(F) = \int F(W + \log F) \equiv F_0 \log F_0 + \sum_{N=1}^{\infty} \frac{1}{N!} \int_{\Lambda^N \setminus Q_N} F_N(W_N + \log F_N)\, d^N x \tag{2.7}$$

and, for $S \in \mathscr{U}_{\mathbf{W}}$, the grand canonical potential

$$\begin{aligned} \Omega_S(F) &= \int F\left(W - \textstyle\sum S + \log F\right) \\ &= F_0 \log F_0 + \sum_{N=1}^{\infty} \frac{1}{N!} \int_{\Lambda^N \setminus Q_N} F_N \left(W_N + \sum_{i=1}^{N} S(x_i) + \log F_N\right) d^N x \end{aligned} \tag{2.8}$$

It is easy to show that $\Omega_S(\cdot)$ is minimized by the Gibbs distribution G_S, and that its extremum value is $-\log \Xi(S)$.

The functionals Ω_S and $\mathfrak{F}$ are related by

$$\Omega_S(F) = \int \rho_F S + \mathfrak{F}(F) \tag{2.9}$$

provided that the right-hand side of the above is well defined. In particular, for Gibbs distributions

$$\mathfrak{F}(G_U) = -\int \rho_U U - \log \Xi(U) \tag{2.10}$$

and

$$\Omega_S(G_U) = \int \rho_U(S - U) - \log \Xi(U) \tag{2.11}$$

In some cases of physical relevance, certain terms in equations (2.9)–(2.11) are *not* separately well defined. This problem may often be traced to the divergence of $\int \rho_F |\log \rho_F|$. In such circumstances, this difficulty can be circumvented by considering instead "renormalized" functionals. For example,

$$\mathfrak{F}^{\text{ren}}(F) = \int F \left[W + \log F - \log \left(\prod \rho_F \right) \right] \tag{2.12}$$

$$\equiv F_0 \log F_0 + \sum \frac{1}{N!} \int_{\Lambda^N \setminus \Omega_N} F_N \left\{ W_N + \log F_N - \log \left[\prod_{i=1}^{N} \rho_F(x_i) \right] \right\} d^N x$$

This renormalization corresponds, in some sense, to a subtraction of the ideal gas ($\mathbf{W} \equiv 0$) behavior of the functional $\mathfrak{F}$. Such a procedure was invoked in Refs 1 and 5, and is discussed in Ref. 10. Moreover, when viewed as a function of single-particle densities, $\mathfrak{F}^{\text{ren}}$ serves as a generating functional for the direct correlation functions of a nonuniform fluid.[5,10]

The strategy in density functional theory is to view $\mathfrak{F}$ and Ω_S as functionals of densities rather than probability distributions. However, one is immediately faced with the nonuniqueness problem: There are in general many distribution functions, $F \in \mathscr{D}$, which reduce to the same density ρ_F. For $\rho \in L^1$, let us define the set

$$\rho^{-1}(\mathscr{D}) = \{F \in \mathscr{D} \mid \rho_F = \rho\} \tag{2.13}$$

In light of the above degeneracy, it is not possible to expresses $\mathfrak{F}$ and Ω_S directly as functionals of densities. However, it is always possible to define

$$\bar{\mathfrak{F}}(\rho) = \inf\{\mathfrak{F}(F) \mid F \in \rho^{-1}(\mathscr{D})\} \tag{2.14}$$

$$\bar{\Omega}_S(\rho) = \inf\{\Omega_S(F) \mid F \in \rho^{-1}(\mathscr{D})\} \tag{2.15}$$

and similarly for the renormalized functionals. The utility of these functionals is a consequence of the following proposition.

Proposition 2.1. If, for some $U \in \mathscr{U}_{\mathbf{W}}$, $G_U \in \rho^{-1}(\mathscr{D})$ and $\int \rho |U| < \infty$, then (a) U is unique, and (b) $\bar{\mathfrak{F}}(\rho) = \mathfrak{F}(G_U)$, i.e., the infimum in equation (2.14) is a minimum.

Proof. (a) This was proved in Ref. 1 (see also Ref. 5). For completeness, observe that this follows from a straightforward application of either the Jensen or the Cauchy–Schwarz inequality.

(b) First note that, for each N, W_N may be written in the form

$$W_N = -\log(G_U)_N - \sum_{i=1}^{N} U(x_i) - \log \Xi(U) \tag{2.16}$$

Hence for every $F \in \rho^{-1}(\mathscr{D})$

$$\begin{aligned}\mathfrak{F}(F) &= \left\{\sum \frac{1}{N!} \int F_N \left[-\log(G_U)_N - \sum^N U + \log F_N\right]\right\} - \log \Xi(U) \\ &= -\left\{\sum \frac{1}{N!} \int F_N \log[(G_U)_N/F_N]\right\} + \mathfrak{F}(G_U) \qquad (2.17)\end{aligned}$$

However, by Jensen's inequality

$$-\sum \frac{1}{N!} \int F_N \log[(G_U)_N/F_N] \geqslant -\log\left[\sum \frac{1}{N!} \int (G_U)_N\right] = 0 \quad \blacksquare \quad (2.18)$$

Remarks. (1) In the proof of uniqueness, the hypothesis $\int \rho\, |U| < \infty$ may be replaced by the hypothesis $\int \rho\, |U + \log \rho| < \infty$. As will be shown in Section 3 (Proposition 3.2), the latter condition is automatically satisfied in H-stable systems whenever $G_U \in \rho^{-1}(\mathscr{D})$.

(2) Assuming all terms in $\Omega_S(G_U)$ are defined, a proof along the lines of (b) shows that $\bar{\Omega}_S(\rho) = \Omega_S(G_U)$. Similar proofs also hold for the renormalized functionals.

The principal assumption of density functional theory is that the functionals $\bar{\mathfrak{F}}$ and $\bar{\Omega}_S$, which we take to be defined by infima as in Eqs. (2.14) and (2.15), do in fact achieve minima. The above proposition reduces a proof of this conjecture to the inverse problem, namely that there exists a unique $U \in \mathscr{U}_W$ such that $\rho_U = \rho$. In the following section we shall show that, for H-stable systems, the inverse conjecture holds whenever an appropriate density functional can be defined.

3. THE INVERSE PROBLEM FOR H-STABLE SYSTEMS

In this section, we consider the inverse problem for H-stable systems. Our principal result (Theorem 3.4) is that the single-particle reduction of any distribution $F \in \mathscr{D}$, with the positive part of $\mathfrak{F}^{\text{ren}}(F)$ finite, is also a single-particle density. Note that if for some density ρ, the above condition is not satisfied for *any* $F \in \rho^{-1}(\mathscr{D})$, then $\overline{\mathfrak{F}^{\text{ren}}}(\rho)$ is infinite, or even undefined. Thus this theorem establishes the validity of the inverse conjecture for all densities which are relevant in the functional approach.

As a corollary, we prove that the set of single-particle densities is convex. Moreover, we show that the functionals $\bar{\mathfrak{F}}$ and $\bar{\Omega}_S$ are convex on this set.

Definition. An interacting system is said to be *H-stable* if there is a constant $B < \infty$ such that, for each N,

$$W_N \geqslant -BN \text{ a.e. } [d^N x] \tag{3.1}$$

Furthermore, we take $W_1 \in L^\infty(dx)$.

Remark. The above differs from the standard definition of H-stability in that we require W_1 to be bounded above. Note, however that this condition is no restriction in physical systems since $-W_1$ is simply the chemical potential.

Proposition 3.1. For any H-stable interaction, $\mathbf{W}$, the set of external potentials, $\mathscr{U}_{\mathbf{W}}$, coincides precisely with the set

$$\mathscr{U} = \{U: \Lambda \to \mathbb{R}^* \mid e^{-U} \in L^1(dx)\} \tag{3.2}$$

Proof. This is easily established by noting that

$$1 + \|e^{-U}\|_1 \exp[-\|W_1\|_\infty] \leqslant \Xi(U) \leqslant \exp[\|e^{-U}\|_1 B] \tag{3.3}$$

while $\|\rho_U\|_1$ is bounded above by the ratio of these two estimates. ∎

Proposition 3.2. Let $U \in \mathscr{U}$. For any H-stable system, the function $V: \Lambda \to \mathbb{R}$ defined by

$$e^{-U} = \rho_U e^{-V} \tag{3.4}$$

is in L^∞.

Proof. This is exactly proposition 8.14 of Ref. 1. The proof follows from an estimate similar to Eq. (3.3). ∎

Remark. For certain $F \in \mathscr{D}$, the functionals $\mathfrak{F}, \mathfrak{F}^{\text{ren}}$, and Ω_S introduced in Section 2 may be infinite or even undefined. However, the positive parts of the functionals are always unambiguously defined and are given, for example, by

$$[\mathfrak{F}(F)]_+ = \sum_{N=0}^{\infty} \frac{1}{N!} \int F_N (W_N + \log F_N)_+ \, d^N x \tag{3.5}$$

where $(W_N + \log F_N)_+ \equiv \max\{(W_N + \log F_N), 0\}$. For finite volumes, $[\mathfrak{F}(F)]_+ < \infty$ implies $[\mathfrak{F}^{\text{ren}}(F)]_+ < \infty$; however, the reverse implication does *not* hold unless $\int \rho_F |\log \rho_F| < \infty$.

We denote by $\tilde{\mathscr{D}}$ the set of distributions

$$\tilde{\mathscr{D}} = \{F \in \mathscr{D} \mid [\mathfrak{F}^{\text{ren}}(F)]_+ < \infty\} \tag{3.6}$$

It is clear that $\tilde{\mathscr{D}}$ is the natural domain of the functional $\mathfrak{F}^{\text{ren}}$.

Lemma 3.3. Let $F \in \rho^{-1}(\tilde{\mathscr{D}})$. Then for $\varepsilon > 0$ sufficiently small, there exists an $F_\varepsilon \in \tilde{\mathscr{D}}$ such that $\rho_{F_\varepsilon} = (1+\varepsilon)\rho$.

Proof. Let $F = (F_N) \in \rho^{-1}(\tilde{\mathscr{D}})$ and take $t > 1$. Define $f_N(t; x_1,..., x_N) = tF_N(x_1,..., x_N)$ and

$$(F_N)_{\varepsilon(t)} = f_N(t)/\langle f(t)\rangle_{\mathbf{W}} \tag{3.7}$$

Then $F_{\varepsilon(t)}$ satisfies

$$\rho_{F_{\varepsilon(t)}} = [1 + \varepsilon(t)]\rho \tag{3.8}$$

with

$$\varepsilon(t) = F_0(t-1)/[t - F_0(t-1)] \tag{3.9}$$

Thus, by taking t sufficiently large, any $\varepsilon < F_0/[1 - F_0]$ may be achieved. It is easily verified that $F_\varepsilon \in \tilde{\mathscr{D}}$. ∎

Theorem 3.4. In H-stable systems, the necessary and sufficient condition for a nonnegative function $\rho \in L^1$ to be the single-particle density of a Gibbs distribution G_U with $U \in \mathscr{U}$ is that there exist an $F \in \rho^{-1}(\tilde{\mathscr{D}})$.

Proof. Necessity is straightforward. Suppose that $\rho = \rho_U$ for some $U \in \mathscr{U}$. Then, writing U in the form $e^{-U} = \rho e^{-V}$, we have

$$[\mathfrak{F}^{\mathrm{ren}}(G_U)]_+ \leqslant \int \rho\,|V| + \log \Xi(U) \tag{3.10}$$

which is finite since $V \in L^\infty$.

In order to prove sufficiency, we shall use a variational approach similar to that employed in Ref. 1. Consider the functional

$$\mathfrak{G}(V) = \exp\left(-\int \rho V\right)\Big/\Xi[V] \tag{3.11}$$

where $\Xi[V] \equiv \Xi(U = -\log\rho + V)$. The class of functions on which $\mathfrak{G}$ is defined is given by

$$\mathscr{V} = \left\{ V: \Lambda \to \mathbb{R} \;\middle|\; \int \rho\,|V| < \infty, \Xi[V] < \infty \right\} \tag{3.12}$$

Note that $\mathscr{V}$ is not empty; indeed, by H-stability, $\mathscr{V}$ contains all bounded functions.

First we claim that $\mathfrak{G}$ is uniformly bounded above in $\mathscr{V}$. To see this, recall that by hypothesis there exists $F=(F_N)\in\rho^{-1}(\tilde{\mathscr{D}})$ in terms of which we may write

$$\begin{aligned}\Xi[V]&=\sum_{N=0}^{\infty}\frac{1}{N!}\int e^{-W_N}\prod^{N}\rho e^{-V}\,d^Nx\\ &\geqslant\sum_{N=0}^{\infty}\frac{1}{N!}\int\exp\left(-W_N-\log F_N+\log\prod^{N}\rho-\sum^{N}V\right)F_N\,d^Nx\\ &\geqslant\sum_{N=0}^{\infty}\frac{1}{N!}\int\exp\left[-\left(W_N+\log F_N-\log\prod^{N}\rho\right)_+\right]\exp\left(-\sum^{N}V\right)F_N\,d^Nx\end{aligned}\tag{3.13}$$

We note that the measure $\mathrm{dF}=[(1/N!)\,F_N\,d^Nx\mid N=0,1,2,...]$ is a normalized measure on the space $\bigoplus_{N=0}^{\infty}\Lambda^N$. Moreover, since $[\mathfrak{F}^{\mathrm{ren}}(F)]_+<\infty$ and $V\in\mathscr{V}$, the arguments of the two exponents in Eq. (3.13) are in the Fock space $\mathbb{L}^1(dF)\equiv l^1(\bigoplus_{N=0}^{\infty}L^1(\Lambda^N,dF_N))$. Applying Jensen's inequality, we obtain

$$\Xi[V]\geqslant\exp(-[\mathfrak{F}^{\mathrm{ren}}(F)]_+)\exp\left(-\int\rho V\right)\tag{3.14}$$

from which it follows that $\mathfrak{G}$ is uniformly bounded.

Let (V_k) denote a maximizing sequence for $\mathfrak{G}$ in $\mathscr{V}$. Consider the Fock space $\mathbb{L}^2(d\Sigma)$ where the measure $d\Sigma$ is given by

$$d\Sigma=\left(\frac{1}{N!}e^{-W_N}\prod^{N}\rho\,d^Nx\mid N=0,1,2,...\right)\tag{3.15}$$

For each k, the product function

$$\pi(e^{-V_k/2})\equiv\left(\prod_{i=1}^{N}e^{-V_k(x_i)/2}\mid N=1,2,...\right)$$

is in $\mathbb{L}^2(d\Sigma)$ since the square of the $\mathbb{L}^2(d\Sigma)$ norm is simply the partition function, i.e.,

$$|||\pi(e^{-V_k/2})|||_2^2=\Xi[V_k]\tag{3.16}$$

We claim that a subsequence of the functions $\pi(e^{-V_k/2})$ converges weakly in $\mathbb{L}^2(d\Sigma)$ to a function $\pi(e^{-V/2})$. Moreover, the function V so obtained is the unique maximizer of $\mathfrak{G}$, and the potential $U=-\log\rho+V$ satisfies $\rho_U=\rho$.

In order to establish weak convergence, we first show that the norms $|||e^{-V_k/2}|||_2$ are uniformly bounded above. Indeed, suppose this is not the case. Then there is a subsequence, which we also denote by (V_k), such that

$$\Xi[V_k] \to \infty \tag{3.17}$$

However, since any subsequence is also a maximizing sequence, this implies

$$\exp\left(-\int \rho V_k\right) \to \infty \tag{3.18}$$

Now recall (Lemma 3.3) that there is an $F_\varepsilon \in \tilde{\mathscr{D}}$ with $\rho_{F_\varepsilon} = (1+\varepsilon)\rho$. Repeating the arguments used to bound the functional with F replaced F_ε, we obtain

$$\mathfrak{G}(V_k) \leqslant (1+\varepsilon)^n \exp\{[\mathfrak{F}^{\text{ren}}(F_\varepsilon)]_+\} \exp\left(\varepsilon \int \rho V_k\right) \tag{3.19}$$

with $n \equiv \int \rho$. This, however, implies $\mathfrak{G}(V_k) \to 0$, which contradicts the assertion that (V_k) is a maximizing sequence.

The remainder of the proof is essentially the same as that of Proposition 8.3 and Theorem 8.4 in Ref. 1. For completeness, we shall sketch the principal steps. First, since the sequence of functions $\pi(e^{-V_k/2})$ is $\mathbb{L}^2(d\Sigma)$-norm bounded, the Banach–Alaoglu theorem ensures that a subsequence converges weakly. Although it is by no means obvious that the limit function is of the form $\pi(e^{-V/2})$, this follows from the fact that, for each N, the measure $e^{-W_N} \prod^N \rho \, d^N x$ is absolutely continuous with respect to $\prod^N \rho \, d^N x$, while for $N = 1$, the corresponding measures are equivalent (see Ref. 1, Theorem A.6 for more details).

Next, it must be shown that the limiting V so defined is in the set $\mathscr{V}$ and that it maximizes $\mathfrak{G}$ in this set. It is not difficult to show that $V_- \in L^1(\rho \, dx)$. The fact that $V_+ \in L^1(\rho \, dx)$ (which proves $V \in \mathscr{V}$) and that V actually maximizes the functional can be ascertained from the inequality

$$\exp\left(-\int \rho V\right) \geqslant \lim_{k \to \infty} \exp\left(-\int \rho V_k\right) \tag{3.20}$$

a proof of which is given in Ref. 1.

Finally, a standard variational argument shows that the function U defined by $e^{-U} = \rho e^{-V}$ satisfies $\rho = \rho_U$ a.e. $[dx]$. ∎

Remark. The above theorem provides a necessary and sufficient condition for the existence of a solution to the inverse problem in any H-stable system. By contrast, in Ref. 1 we found sufficient (but not necessary) conditions for the existence of a solution in systems which need not be H-stable. In particular, rather than H-stability, we required that every constant potential $V = \text{const}$ be in the set $\mathscr{V}$. The results of the previous paper may in fact be recovered as a special case of those in this paper if the assumption of H-stability in Theorem 3.4 is replaced by the above condition on constant potentials. In that case, the above proof of sufficiency is still legitimate. The additional condition imposed in Ref. 1 is easily seen to correspond to the explicit construction of a particular F satisfying the conditions of Theorem 3.4.

Corollary. In any H-stable system, the set of single-particle densities is convex.

Proof. Let ρ_1 and ρ_2 be single-particle densities corresponding to the potentials U_1 and U_2, respectively. We must show that for $\lambda \in [0, 1]$

$$\rho_\lambda = \lambda\rho_1 + (1-\lambda)\rho_2 \tag{3.21}$$

is also a single-particle density. By Theorem 3.4, it is enough to establish the existence of an $F \in \tilde{\mathscr{D}}$ such that $\rho_{F_\lambda} = \rho_\lambda$. Indeed, it is easy to verify that the convex combination of the Gibbs distributions

$$F_\lambda = \lambda G_{U_1} + (1-\lambda) G_{U_2} \tag{3.22}$$

is a probability distribution with the desired properties. Clearly $\rho_{F_\lambda} = \rho_\lambda$. Moreover,

$$[\mathfrak{F}^{\text{ren}}(F_\lambda)]_+ \leqslant |\log\alpha| + |\log\beta|\,(\lambda n_1 + (1-\lambda)n_2) < \infty \tag{3.23}$$

where $\alpha = \max\{1/\Xi(U_1),\ 1/\Xi(U_2)\}$, $\beta = \max\{\|e^{-V_1}\|_\infty/\lambda,\ \|e^{-V_2}\|_\infty/(1-\lambda)\}$ and $n_i = \int \rho_i$. ∎

Proposition 3.5. The density functional $\bar{\mathfrak{F}}(\rho)$ is convex on the set of single-particle densities.

Remark. As noted earlier, if $\int \rho_U |\log \rho_U|$ is divergent, the functional $\bar{\mathfrak{F}}(\rho_U) = -\int \rho_U U - \log \Xi(U)$ may be ill defined, i.e., of the form "infinity minus infinity." However, even in these cases, the proposition is legitimate provided that convexity of $\bar{\mathfrak{F}}$ is interpreted as negativity of the convex difference:

$$\bar{\mathfrak{F}}(\lambda\rho_1 + (1-\lambda)\rho_2) - \lambda\bar{\mathfrak{F}}(\rho_1) - (1-\lambda)\,\bar{\mathfrak{F}}(\rho_2) \tag{3.24}$$

a quantity which is always well defined (although possibly equal to $-\infty$). In the proof given below, we shall assume for the sake of simplicity that all terms are well defined.

Proof. Let ρ_1 and ρ_2 be single-particle densities corresponding to the potentials U_1 and U_2. Define ρ_λ and F_λ as in the proof of the corollary to Theorem 3.4. It then follows immediately from the manifest convexity of $\mathfrak{F}(F)$ and Proposition 2.1b that

$$\begin{aligned}\bar{\mathfrak{F}}(\rho_\lambda) &= \inf\{\mathfrak{F}(F) \mid F \in \rho_\lambda^{-1}(\tilde{\mathscr{D}})\} \\ &\leqslant \mathfrak{F}(F_\lambda) \\ &\leqslant \lambda \mathfrak{F}(G_{U_1}) + (1-\lambda)\,\mathfrak{F}(G_{U_2}) \\ &= \lambda \bar{\mathfrak{F}}(\rho_1) + (1-\lambda)\,\bar{\mathfrak{F}}(\rho_2) \quad \blacksquare \end{aligned} \tag{3.25}$$

Remark. The functional $\bar{\Omega}_S(\rho)$ differs from $\bar{\mathfrak{F}}(\rho)$ simply by a term linear in ρ, and thus is also convex.

4. HIGHER-ORDER CORRELATION FUNCTIONS

The inverse problem for higher-order correlation functions is the question of whether it is possible to produce any given mth order correlation function by augmenting a given system of interactions, $\mathbf{W} = (W_N)$, with some "external" m-body interaction. Although the physical interpretation of an "external" m-body interaction is somewhat obscure, this form of the inverse problem has interesting applications in many systems[11] (see also discussion below).

Under conditions analogous to those imposed in Section 3, one expects that if a given function of m variables is the m-body reduction of an appropriate probability distribution, it is also the mth order correlation function of the system augmented by a unique m-body interaction. Indeed, the conditions under which such a result is established (Theorem 4.1) amount again to the finiteness of certain thermodynamic potentials.

This result is of some interest with regards to the *modeling* of physical systems. It is often convenient to approximate physical forces by the sum of two-body interactions. Although it is *not* generically the case that "real" interactions are exclusively two-body, it is widely believed that, in some sense, this is a good approximation. Theorem 4.1, specialized to the case $\mathbf{W} = 0$ and $m = 2$, sheds some light on the validity of this approximation.

Suppose that various measurements are performed on a physical system in equilibrium. It can be argued that the "best" two-body approximation to the system—should it exist—is the one which produces a second-order correlation function identical to that observed. Theorem 4.1 says that such a two-body approximation does indeed exist if the measured correlation function is the two-particle reduction of a probability distribution corresponding to certain finite thermodynamic potentials. However, the existence of the required probability distribution is guaranteed—it is simply the actual distribution of the physical system under observation.

Of course these results also apply to the question of m-body decomposition, $m > 2$; however this issue is of less relevance to existing models.

The inverse problem for higher-order correlation functions may be formulated in analogy with previous sections. Consider a system specified by an interaction $\mathbf{W} = (W_N)$. The total interaction in the presence of an m-body "external" potential, $S: \Lambda^m \to \mathbb{R}$, is given by

$$W_N(x_1,..., x_N) + \sum_{p \in \mathscr{P}_{N,m}} S(x_{p_1},..., x_{p_m}) \tag{4.1}$$

where $\mathscr{P}_{N,m}$ denotes the $\binom{N}{m}$ permutations of N objects, taken m at a time. In order to simplify our notation, we shall henceforth assume that the interactions W_N and S are symmetric functions.

The mth order correlation function of the augmented system is given by

$$\alpha_m(S; x_1,..., x_m) \\ = \frac{1}{\Xi(S)} \sum_{N \geqslant m} \frac{1}{N!} \binom{N}{m} \int_{\Lambda^N \backslash Q_N} \exp\left[-\left(W_N + \sum_{\mathscr{P}_{N,m}} S\right)\right] d^{N-m}x \tag{4.2}$$

In the above, $\Xi(S)$ is the partition function of the augmented system, and Q_N again denotes the N-particle hard-core excluded region. The m-body reductions of probability distributions, $\alpha_m(F \mid x_1,..., x_m)$, can be defined by a similar procedure.

In a particular system, $\mathbf{W} = (W_N)$, the m-body inverse problem is the question of whether, for a given nonnegative (symmetric) function $\alpha_m \in L^1(d^m x)$, there exists an m-body interaction S such that $\alpha_m = \alpha_m(S)$.

For the purposes of the higher-order problem, we shall restrict attention to systems which obey certain finiteness conditions. First, we require that the volume $|\Lambda| = \int_\Lambda dx$, is finite. We also work in the truncated ensemble (i.e., $W_N \equiv \infty$ for $N > N'$)—which is typical in finite volumes if hard cores are present. Finally, we assume that in the absence of externally imposed interactions, the partition function is finite. Of course this last condition is automatically satisfied if the system is H-stable; here, however, we need not impose any pointwise bounds on the W_N.

In addition, we shall impose a "mechanical stability" condition on the structure of the hard-core excluded regions. For the m-point inverse problem, we require that if $(y_1,..., y_m) \in Q_m$, then for almost every $(x_{m+1},..., x_N)$,

$$(y_1,..., y_m, x_{m+1},..., x_N) \in Q_N \tag{4.3}$$

A violation of the above condition would imply that the injection of additional particles into the system allows more freedom of movement to those already present—a physical absurdity.

For clarity of exposition, we shall take $W_m \in L^\infty$ on $\Lambda^m \backslash Q_m$ for the solution of the m-point problem. However, it should be noted that this condition can be replaced by the weaker assumption

$$\int_{\Lambda^m \backslash Q_m} (W_m)_+ \, \alpha_m \, d^m x < \infty \tag{4.4}$$

Consideration of systems which violate the above condition would introduce a new, and not particularly interesting complication; namely, that the *external* m-body potential would have to compensate for a pathological m-body interaction, in addition to affecting the interactions for $N > m$.

In our analysis, we shall assume that the given m-point function is log summable:

$$\int a_m \log \alpha_m \, d^m x < \infty \tag{4.5}$$

If this condition is not satisfied, then the entropy per particle is infinite. We note also that the assumption of finite volumes implies that for any measurable function f

$$\int |f| \log |f| > -\infty \tag{4.6}$$

It follows that, in finite volumes, the set of log summable functions is a convex subset of L^1. In order to circumvent the assumption of log summability, a renormalization scheme analogous to that employed in Section 3 would be required. This may prove difficult since for $m > 1$ there is no exact solution for the ideal gas ($\mathbf{W} \equiv 0$) case.

Our analysis is quite similar to that of Section 3. Here we consider the functional

$$\mathfrak{F}_m(F) = \sum_{N \geqslant m} \frac{1}{N!} \int_{\Lambda^N \backslash Q_N} F_N(W_N + \log F_N) \, d^N x \tag{4.7}$$

which is well defined for distributions in the set

$$\tilde{\mathscr{D}}_m = \{F \in \mathscr{D} \mid [\mathfrak{F}_m(F)]_+ < \infty\} \tag{4.8}$$

Our principal result is the following:

Theorem 4.1. Let $\mathbf{W} = (W_N)$ be a system of interactions for a truncated grand canonical ensemble. Furthermore, let $\mathbf{W}$ be such that $W_m \in L^\infty$ on $\Lambda^m \backslash Q_m$, $\Xi(0) < \infty$ and the hard-core excluded sets satisfy the mechanical stability condition (4.3). Then the necessary and sufficient condition for a nonnegative, log summable function $\alpha \in L^1(d^m x)$ to be the mth-order correlation function of the system augmented by an m-body potential S: $\Lambda^m \to \mathbb{R}$ is that there exists an $F \in \alpha^{-1}(\tilde{\mathscr{D}}_m)$. Furthermore, S is unique.

Proof. Necessity is easily established. Suppose that such an S exists. Then, taking $F_N = [1/\Xi(S)]\, e^{-W_N} \prod e^{-S}$, we have

$$[\mathfrak{F}_m(F)]_+ \leqslant \frac{1}{\Xi(S)} \sum_{N \geqslant m} \frac{1}{N!} \int e^{-W_N} \prod e^{-S} \left(-\sum S\right)_+ + \log \Xi(S)$$

$$\leqslant \int \alpha S_-\, d^m x + \log \Xi(S) \tag{4.9}$$

It therefore suffices to show that $\int \alpha S_- < \infty$. Let A be the set on which S is negative and $a = \int_A \alpha > 0$. By Jensen's inequality, we have

$$\infty > \Xi(S) \geqslant a \exp\left[-\|W_m\|_\infty + \frac{1}{a}\left(\int \alpha S_- - \int_A \alpha \log \alpha\right)\right] \tag{4.10}$$

from which it follows that $\int \alpha S_- < \infty$.

In order to prove sufficiency, we again use a variational approach. Consider the functional

$$\mathfrak{I} = \exp\left(-\int \alpha S\right)\Big/\Xi(S) \tag{4.11}$$

defined on the set of functions

$$\mathscr{S} = \left\{S : \Lambda^m \to \mathbb{R} \,\Big|\, \int \alpha\,|S| < \infty, \Xi(S) < \infty\right\} \tag{4.12}$$

The set $\mathscr{S}$ is not empty since $0 \in \mathscr{S}$. Furthermore, since by hypothesis there is an $F \in \tilde{\mathscr{D}}_m$ with $\alpha_m(F) = \alpha$, the functional is uniformly bounded in $\mathscr{S}$ by an estimate analogous to that in Eqs. (3.13) and (3.14).

Let (S_k) be a maximizing sequence for $\mathfrak{I}$ in $\mathscr{S}$. By an argument identical to that in the proof of Theorem 3.4, we may establish weak $\mathbb{L}^2(d\Theta)$ convergence of a subsequence of the functions $\pi_m(e^{-S_k/2})$. Here

$$\pi_m(e^{-S/2}) \equiv \left\{\prod_{p \in \mathscr{P}_{N,m}} \exp[-S(x_{p_1}, \ldots, x_{p_m})/2] \,\Big|\, N \geqslant m\right\} \tag{4.13}$$

and $d\Theta$ denotes the measure

$$d\Theta = \left(\frac{1}{N!} e^{-W_N} d^N x \mid N \geqslant m\right) \tag{4.14}$$

Consider a particular N-particle term in Eq. (4.13). Although the term is a product of functions, it differs from the corresponding products in Section 3 in that the individual factors are not defined on separate spaces. Therefore the results of Ref. 1 (Appendix A) cannot be used to prove that the limit of the sequence $\pi_m(e^{-S_k/2})$ is of the form $\pi_m(e^{-S/2})$. It is, however, possible to establish strong convergence, from which the desired result will be shown to follow.

Strong $\mathbb{L}^2(d\Theta)$ convergence of $\pi_m(e^{-S_k/2})$ is a consequence of the weak convergence and certain convexity properties of the functional $\mathfrak{I}$. A similar result is established in Ref. 1 (Theorems 6.2 and 8.16), to which the reader is referred for details.

From the *strongly* convergent sequence, we may extract a further subsequence which converges pointwise a.e. $[d\Theta]$. This implies, in particular, that $S_k(x_1,..., x_m)$ converges pointwise a.e. to some function $S(x_1,..., x_m)$ on $\Lambda^m \backslash Q_m$. The required form of the limit now follows directly from the condition of mechanical stability.

An argument along the lines of that in the previous section shows that $S \in \mathscr{S}$ and that S actually maximizes $\mathfrak{I}$. That $\alpha_m(S) = \alpha$ follows from the variational principle, a proof of which is facilitated here by the truncation of the ensemble.

The proof of uniqueness requires two steps. First, we must show that if $\alpha_m(S) = \alpha$, then $S \in \mathscr{S}$. This follows from the Jensen inequality and the fact that $\Xi(0) < \infty$. Next, again using Jensen's inequality, it can be shown that if $\alpha_m(S_1) = \alpha_m(S_2)$ with $S_1, S_2 \in \mathscr{S}$, then $S_1 = S_2$ a.e. (see Ref. 1, Theorems 2.4 and 8.5, for details). ∎

Corollary. Let $\mathbf{W} = (W_N)$ be a system of interactions satisfying the conditions of Theorem 4.1 on a finite measure space $\langle \Lambda, dx \rangle$. The set of all log summable mth-order correlation functions which may be produced by augmenting $\mathbf{W}$ be an external m-body potential is a convex set.

Remark. As noted in the Introduction, the hypothesis that there exist an $F \in \alpha^{-1}(\tilde{\mathscr{D}}_m)$, while necessary for the solution of the inverse problem, is by no means trivial to verify. This is particularly true for $m > 1$. In this context, it should be noted that the question of constructive criteria for α_2 to arise from some F has been considered in the literature (see, for example, Yamada[12] for necessary conditions in some spcial cases).

ACKNOWLEDGMENTS

We should like to thank Frank Stillinger for discussing with us the implications of the m-point problem for the question of two-body decomposition of forces. This conversation reopened our interest in the inverse problem. We should also like to thank Elliott Lieb for numerous discussions on various aspects of the inverse problem. Finally, we should like to thank J. D. Weeks for a useful discussion on density functional theory and the inverse problem for higher order correlation functions.

REFERENCES

1. J. T. Chayes, L. Chayes, and E. H. Lieb, *Commun. Math. Phys.* **93**:57 (1984).
2. F. H. Stillinger and F. P. Buff, *J. Chem. Phys.* **37**:1 (1962).
3. J. L. Lebowitz and J. K. Percus, *J. Math. Phys.* **4**:116 (1963).
4. R. Evans, *Adv. Phys.* **28**, 143 (1979).
5. A. Robledo and C. Varea, *J. Stat. Phys.* **26**:513 (1981).
6. J. K. Percus, in *The Liquid State of Matter, Studies in Statistical Mechanics*, Vol. VIII, E. W. Montroll and J. L. Lebowitz, eds. (North-Holland, Amsterdam, 1982), p. 31.
7. J. K. Percus, *J. Stat. Phys.* **15**:505 (1976).
8. J. K. Percus, *J. Stat. Phys.* **28**:67 (1982).
9. E. H. Lieb, in *Physics as Natural Philosophy: Essays in Honor of Laszlo Tisza on His 75th Birthday*, A. Shimony and H. Feshbach, eds. (M.I.T. Press, Cambridge, 1982), p. 111.
10. J. T. Chayes, thesis, Princeton University, 1983.
11. J. D. Weeks, private communication.
12. M. Yamada, *Progr. Theor. Phys.* **25**:579 (1961).