INTERNATIONAL JOURNAL OF QUANTUM CHEMISTRY, VOL XXIV, 243–277 (1983)

Density Functionals for Coulomb Systems

ELLIOTT H. LIEB

Departments of Mathematics and Physics, Princeton University, P.O.B. 708, Princeton, New Jersey 08544, U.S.A.

Abstract

This paper has three aims: (i) To discuss some of the mathematical connections between N-particle wave functions ψ and their single-particle densities $\rho(x)$. (ii) To establish some of the mathematical underpinnings of "universal density functional" theory for the ground state energy as begun by Hohenberg and Kohn. We show that the HK functional is *not* defined for all ρ and we present several ways around this difficulty. Several less obvious problems remain, however. (iii) Since the functional mentioned above is not computable, we review examples of explicit functionals that have the virtue of yielding rigorous bounds to the energy.

Introduction

It is a pleasure to dedicate this article to Laszlo Tisza on the occasion of his seventy-fifth birthday. As a colleague at MIT he was a source of inspiration and encouragement, especially in drawing our attention to the importance of careful and precise thought in mathematical physics. The subject, if not the content, of this article may therefore not be inappropriate in a book dedicated to Professor Tisza (see the Acknowledgment).

The idea of trying to represent the ground state (and perhaps some of the excited states as well) of atomic, molecular, and solid state systems in terms of the diagonal part of the one-body reduced density matrix $\rho(x)$ is an old one. It goes back at least to the work of Thomas [1] and Fermi [2] in 1927. In 1964 the idea was conceptually extended by Hohenberg and Kohn (HK) [3]. Since then many variations on the theme have been introduced. As the present article is not meant to be a review, I shall not attempt to list the papers in the field. Some recent examples of applications are Refs. 4 and 5. Some recent examples of theoretical papers which will play a role here are Refs. 6–12. A bibliography can be found in the recent review article of Bamzai and Deb [13].

This article has three aims:

(i) To discuss and prove some of the mathematical relations between N-particle functions ψ and their corresponding single-particle densities ρ.

(ii) To discuss the mathematical underpinnings of general density functional theory along the lines initiated by HK. In that theory a universal energy functional $F(\rho)$ is introduced. Despite the hopes of HK, $F(\rho)$ is *not* defined for all ρ because it is not true (see Theorem 3.4) that every ρ (even a "nice" ρ) comes from the ground state of some single-particle potential $v(x)$. This problem can be remedied by replacing the HK functional by the Legendre transform of the energy, as is done here. However, the new theory is also not free of difficulties, and these

 CCC 0020-7608/83/090243-35$04.50

can be traced to the fact that the connection between v and ρ is extremely complicated and poorly understood.

(iii) To present briefly another approach to the ground state energy problem by means of functionals that, while not exact, are explicitly computable and yield upper and lower bounds to the energy.

The analysis in this paper gives rise to many interesting open problems. It is my hope that the incompleteness of the results presented here will be partly compensated if others are encouraged to pursue some of the questions raised by them.

It is not my intention to present a brief for HK theory. However, it deserves to be analyzed for at least two reasons: The HK theory is used by many workers and it gives rise to some deep problems in analysis. While it is my opinion that density functionals are a useful way to approach Coulomb systems, there are other approaches besides the HK approach [e.g., see (iii) above]. Apart from the difficulties mentioned above, the HK approach may be too general because *all* potentials have to be considered. Coulomb potentials are special and do lend themselves to a density functional approach; for example, Thomas–Fermi theory is asymptotically exact as $Z \to \infty$ (see Sect. 5E and Ref. 14). In addition to this question of generality there is also the crucial point that the "universal functional" is very complicated and essentially uncomputable. If one is going to make uncontrolled approximations for this functional, then the general theory is not very helpful.

It is a pleasure to thank Barry Simon for some very helpful conversations and the proofs of Theorems 4.4 and 4.8. I also thank Haïm Brezis for the proof of Theorem 1.3.

1. Single-Particle Densities

The first order of business is to describe the single-particle densities of interest. For simplicity we confine our attention to three dimensions whenever dimensionality is important. $z = (x, \sigma)$ will denote a space-spin variable, that is, $x \in \mathbb{R}^3$ and $\sigma \in \{1, \ldots, q\}$. $q = 2$ for electrons, of course, but one might wish to consider $q = 1$, which would mean that a ferromagnetic state is under consideration. We use the notation

$$\int dz = \sum_{\sigma=1}^{q} \int dx. \tag{1.1}$$

Let $\psi = \psi(z_1, \ldots, z_N)$ be an N-particle function (which may be complex valued). To simplify notation we will not indicate N explicitly except where needed. However, the condition of fixed N is crucial and frequently glossed over. The density functionals that will be introduced later are explicitly N dependent in a highly nontrivial way (see Sect. 4A). ψ is assumed to be *normalized*:

$$\int |\psi|^2 = 1, \tag{1.2}$$

(with $\int = \int dz_1, \ldots, dz_N$) and to have *finite kinetic energy*, that is,

$$T(\psi) = \sum_{i=1}^{N} \int |\nabla_i \psi|^2 < \infty. \tag{1.3}$$

Notes. $f \in L^p$ means that f is a function satisfying $\|f\|_p \equiv \{\int |f|^p\}^{1/p} < \infty$. $f \in H^1$ means that f and each component of ∇f are in L^2. Thus, the above $\psi \in H^1$. If f is differentiable everywhere, then ∇f and $T(f) = \int |\nabla f|^2$ are well defined. Otherwise, the correct definition of $T(f)$ for functions in $L^2(\mathbb{R}^m)$ is

$$T(f) = (2\pi)^{-3} \int k^2 |\hat{f}(k)|^2 \, dk, \tag{1.4}$$

where $\hat{f}$ is the Fourier transform of f. Since $f \in L^2$, $\hat{f}$ exists and $\hat{f} \in L^2$. H^1 is a Hilbert space with inner product $(f, g) = \int f^* g + \int \nabla f^* \cdot \nabla g$.

In most of the following it will be assumed that ψ satisfies the Pauli principle, that is, ψ is antisymmetric. However, some of the theorems are easier for symmetric (i.e., bosonic) ψ with $q = 1$, and occasionally this will be mentioned explicitly. In either case, the symmetry implies that

$$T(\psi) = N \int |\nabla_1 \psi|^2. \tag{1.5}$$

We define the *single-particle density* to be [see Eq. (A.1)]

$$\rho(x) = N \sum_{\sigma} \int |\psi((x, \sigma_1), \ldots, (x_N, \sigma_N))|^2 \, dx_2 \cdots dx_N. \tag{1.6}$$

Notice that $\int \rho(x) \, dx = N$, not 1.

Determinants. If $\phi_1(z), \ldots, \phi_N(z)$ are orthonormal functions, we can form the determinantal

$$\psi(z_1, \ldots, z_N) = (N!)^{-1/2} \det \{\phi_i(z_j)\}, \tag{1.7}$$

which is normalized. Then

$$\rho(x) = \sum_{i=1}^{N} \sum_{\sigma=1}^{q} |\phi_i(x, \sigma)|^2. \tag{1.8}$$

$$T(\psi) = \sum_{i=1}^{N} \sum_{\sigma=1}^{q} \int |\nabla \phi_i(x, \sigma)|^2 \, dx. \tag{1.9}$$

Returning to the general case, the finiteness of $T(\psi)$ implies the following [15].

Theorem 1.1. *$\rho(x)^{1/2} \in L^2(\mathbb{R}^3)$ and $\nabla \rho(x)^{1/2} \in L^2(\mathbb{R}^3)$, that is, $\rho(x)^{1/2} \in H^1(\mathbb{R}^3)$. Moreover, $\int (\nabla \rho^{1/2})^2 \leq T(\psi)$.*

Proof. $\rho^{1/2} \in L^2(\mathbb{R}^3)$ because $\int \rho = N$. Now $\nabla \rho(x) = N \int^\dagger (\nabla_1 \psi)^* \psi + N \int^\dagger \psi^* \nabla_1 \psi$, where $\int^\dagger$ means the integral in (1.6). By the Schwarz inequality,

$$[\nabla \rho(x)]^2 \leq 4N\rho(x) \int^\dagger |\nabla_1 \psi|^2.$$

Thus $\int (\nabla \rho^{1/2})^2 \, dx = \frac{1}{4} \int (\nabla \rho)^2 \rho^{-1} \, dx \leq T(\psi)$. ■

We know $\rho^{1/2} \in H^1(\mathbb{R}^3) = \{f | f \in L^2,\ \nabla f \in L^2\}$. (Here we use the standard convention that $\{A|C\}$ means the set of A such that condition C holds.) To discuss the converse of Theorem 1.1 some definitions are useful.

Definition. $\mathscr{I}_N = \{\rho | \rho(x) \geqslant 0,\ \rho^{1/2} \in H^1(\mathbb{R}^3), \int \rho(x)\,dx = N\}$.

Definition. $\mathscr{R}_N = \{\rho | \rho(x) \geqslant 0, \int \rho(x)\,dx = N,\ \rho \in L^3(\mathbb{R}^3)\}$.

$\mathscr{R}_N$ contains $\mathscr{I}_N$ by the Sobolev inequality (see Ref. 16) because if $f \in H^1(\mathbb{R}^3)$, then

$$\int |\nabla f(x)|^2\,dx \geqslant 3(\pi/2)^{4/3}\left[\int |f(x)|^6\,dx\right]^{1/3}. \tag{1.10}$$

Equation (1.10) is true only in three dimensions, but analogous inequalities hold in other dimensions. By Theorem 1.1,

$$T(\psi) \geqslant 3(\pi/2)^{4/3}\|\rho\|_3.$$

$\mathscr{R}_N$ is clearly a *convex set*; that is, if ρ_1 and $\rho_2 \in \mathscr{R}_N$, then $\rho \equiv \lambda\rho_1 + (1-\lambda)\rho_2 \in \mathscr{R}_N$ for all $0 \leqslant \lambda \leqslant 1$. *$\mathscr{I}_N$ is also convex* by the same proof as in Theorem 1.1; that is, by the Schwarz inequality

$$(\nabla\rho)^2 \leqslant 4\rho[\lambda(\nabla\rho_1^{1/2})^2 + (1-\lambda)(\nabla\rho_2^{1/2})^2].$$

In particular, the *functional* $\int [\nabla\rho^{1/2}]^2$ *is convex*. The convexity of $\mathscr{I}_N$ will be important in Sect. 3.

Definition. A function (or functional) f is convex if

$$f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda)f(y)$$

for all $0 \leqslant \lambda \leqslant 1$ and all x and y in the domain of f.

Theorem 1.2. *Suppose $\rho \in \mathscr{I}_N$. Then for either Bose or Fermi statistics there exists a ψ (which is a determinant in the fermion case) such that* (1.6) *holds and, moreover,*

$$T(\psi) \leqslant \int [\nabla\rho^{1/2}(x)]^2\,dx \quad (\textit{bosons}), \tag{1.11}$$

$$T(\psi) \leqslant (4\pi)^2 N^2 \int [\nabla\rho^{1/2}(x)]^2\,dx \quad (\textit{fermions}). \tag{1.12}$$

Proof. For bosons the proof is easy; simply take

$$\psi(x_1, \ldots, x_N) = \prod_{i=1}^N \left(\frac{\rho(x_i)}{N}\right)^{1/2}.$$

For fermions the construction is much more complicated. Some ideas from Ref. 17 will be used in the following. Write $x = (x^1, x^2, x^3)$ and define

$$f(x^1) = \left(\frac{2\pi}{N}\right)\int_{-\infty}^{x^1} ds \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} du\,\rho(s, t, u).$$

Then f is monotone increasing from 0 to 2π. For $k=0,\ldots,N-1$ define

$$\phi^k(x)=[\rho(x)/N]^{1/2}\exp[ikf(x^1)].$$

It is easy to check that the ϕ^k are orthonormal functions in $L^2(\mathbb{R}^3)$. [First do the x^2 and x^3 integrations and then note that the overlap integral is of the form $\int_{-\infty}^{\infty}(df/dx^1)\exp[i(\Delta k)f(x^1)]\,dx^1=\{\exp[i(\Delta k)f(\infty)]-\exp[i(\Delta k)f(-\infty)]\}/i(\Delta k)=0$. Furthermore,

$$N\int|\nabla\phi^k|^2=\int(\nabla\rho^{1/2})^2+\left(\frac{2\pi k}{N}\right)^2\int_{-\infty}^{\infty}g(s)^6\,ds, \tag{1.13}$$

with

$$g(s)^2\equiv\int_{-\infty}^{\infty}\!\!\int dt\,du\,\rho(s,t,u).$$

As in Theorem 1.1, we conclude that

$$g\in H^1(\mathbb{R}^1)\qquad\text{and}\qquad\int\left(\frac{dg}{ds}\right)^2ds\leq\int(\nabla\rho^{1/2})^2\equiv A.$$

Since

$$g(s)^2=2\int_{-\infty}^{s}g(y)\left(\frac{dg(y)}{dy}\right)dy,$$

we conclude by the Schwarz inequality that $g(s)^4\leq 4[\int g^2][\int(dg/dy)^2]$. Thus, the last term in (1.13) is less than $4(2\pi k/N)^2N^2A$. Finally, we take ψ to be a determinant as in (1.7) using the functions $\phi^k(x)\times$(spin up). Equation (1.12) follows by summing on k. ■

Theorem 1.2 is closely related to the results of Gilbert [8] and Harriman [9].

For fermions, the extra factor N^2 in (1.12) is noticeably different from the factor N^0 in Theorem 1.1. Although (1.12) can be improved, it is not easy to do so. In any case, the conclusion is that the map from ψ to $\rho^{1/2}$ given by (1.6) is a map from $H^1(\mathbb{R}^{3N})$ *onto* $H^1(\mathbb{R}^3)$. But the map is clearly not 1 : 1; different ψ's can give the same ρ.

Question 1. Is this map continuous as a map from $H^1(\mathbb{R}^{3N})$ to $H^1(\mathbb{R}^3)$? That is, if ψ is fixed and ψ_j is a sequence (with corresponding ρ and ρ_j) such that $\int|\psi-\psi_j|^2\to 0$ and $\int|\nabla\psi-\nabla\psi_j|^2\to 0$, does it follow that $\int|\rho^{1/2}-\rho_j^{1/2}|^2\to 0$ and $\int|\nabla\rho^{1/2}-\nabla\rho_j^{1/2}|^2\to 0$?

Question 2. Although the map is not invertible (since it is not 1 : 1), we can ask the following: Given a sequence $\rho_j^{1/2}$ that converges to $\rho^{1/2}$ in the above $H^1(\mathbb{R}^3)$ since, and given some ψ satisfying (1.6) for ρ, does there exist a sequence ψ_j [related to ρ_j by (1.6)] that converges to ψ in the above $H^1(\mathbb{R}^3)$ sense? [This is equivalent to the statement that the map $\psi\mapsto\rho^{1/2}$ is "open," that is, the map takes open sets in $H^1(\mathbb{R}^{3N})$ into open sets in $H^1(\mathbb{R}^3)$.]

Intuitively, the answer to both questions should be affirmative. The continuity can indeed be proved, but the proof is not entirely elementary. A proof of Theorem 1.3, due to H. Brezis, is given in the appendix.

Theorem 1.3. *The map* $\psi \mapsto \rho^{1/2}$ *given by* (1.6) *is continuous as a map from* $H^1(\mathbb{R}^{3N})$ *to* $H^1(\mathbb{R}^3)$.

I cannot offer any proof of the openness of the map, however. The fact that these questions do not have simple answers should serve as a warning that the connection between ψ and ρ is not as obvious as one might intuitively think.

2. Single-Particle Density Matrices

If ψ is given as before, we can define the *single-particle density matrix*

$$\gamma(x, x') = N \sum_{\sigma} \int \psi((x, \sigma_1), \ldots, (x_N, \sigma_N)) \times \psi((x', \sigma_1), \ldots, (x_N, \sigma_N))^* \, dx_2 \cdots dx_N. \tag{2.1}$$

This definition is different from the usual one because we sum on σ_1 in (2.1). Usually one defines the quantity $\tilde{\gamma}(x, \sigma; x', \sigma')$, so our $\gamma(x, x') = \sum_\sigma \tilde{\gamma}(x, \sigma; x', \sigma)$. Clearly, $\rho(x) = \gamma(x, x)$.

Theorem 2.1. γ *satisfies*
(i) $\operatorname{Tr} \gamma = \int \gamma(x, x)\, dx = N$.
(ii) *As an operator,* $0 \le \gamma \le qI$, *for fermions; that is,* $0 \le (f, \gamma f) \le q(f, f)$. *For bosons,* $0 \le \gamma \le NI$.

Proof. (i) is "obvious" but not trivial. The point is that if an operator K is given, then its kernel $K(x, y)$ is defined only almost everywhere. In particular, $K(x, x)$ can be anything. Thus, $\operatorname{Tr} K$ need not be $\int K(x, x)\, dx$. However, (i) can be proved from (2.1). This is left as an exercise. To prove (ii), let $M(x, x') = f(x) f(x')^*$ be a one-particle operator with $(f, f) = 1$. Then $A = \sum_{i=1}^N M(x_i, x_i')$ has as its largest eigenvalue on the antisymmetric space the value q. Moreover, A is clearly positive semidefinite. Thus, $0 \le (f, \gamma f) = \operatorname{Tr} \gamma M = (\psi, A\psi) \le q$. ∎

Definition. Let $\gamma(x, y)$ be any kernel. γ is said to be *admissible* if $\operatorname{Tr} \gamma = N$ and $0 \le \gamma \le qI$ (fermions) or $0 \le \gamma \le N$ (bosons). The set of admissible γ is clearly convex; that is, if γ and δ are admissible, then so is $\alpha\gamma + (1-\alpha)\delta$ for $0 \le \alpha \le 1$.

Now we come to a subtle point. If γ is an admissible operator, we can ask two questions:

Question 3. Does an *N-particle density matrix* Γ always exist, where $\Gamma = \Gamma(z_1, \ldots, z_N; z_1', \ldots, z_N')$, so that γ is given by (2.1) with $\psi\psi^*$ replaced by Γ? (Γ is a density matrix if $0 \le \Gamma$ and $\operatorname{Tr} \Gamma = 1$. Γ must also satisfy the appropriate symmetry.)

Question 4. Does a ψ always exist so that (2.1) holds; that is, can Γ be chosen to be a *pure state*, namely, $\Gamma = \psi\rangle\langle\psi$?

The answer to question 4 is No! (for fermions). For bosons, the answer is Yes.

The proof of question 3 (which we now call **Theorem 2.2**) has been known for a long time. An explicit construction is given in Ref. 26. An example in which Γ fails to be of the form $\psi\rangle\langle\psi$, for $N=2$ and $q=1$, is the case in which γ has three nonzero eigenvalues $1, \frac{1}{2}, \frac{1}{2}$. To see this, let the normalized eigenvectors of γ be $f(x)$, $g(x)$, and $h(x)$, respectively; that is,

$$\gamma(x, x') = f(x)f(x')^* + \tfrac{1}{2}g(x)g(x')^* + \tfrac{1}{2}h(x)h(x')^*.$$

Let $A = -\gamma(x_1, x_1') - \gamma(x_2, x_2')$ be an operator on the antisymmetric states. Its lowest eigenvalue is $-1-1/2=-3/2$, which is doubly degenerate. If $\Gamma=\psi\rangle\langle\psi$, then ψ must be a ground state since $\operatorname{Tr}\Gamma A = -\operatorname{Tr}\gamma^2 = -1-1/4-1/4=-3/2$. But every ground state is of the form $\psi = 2^{-1/2}\det(f, p)$, where $p = ag+bh$, $|a|^2+|b|^2=1$. But then $\gamma = f\rangle\langle f + p\rangle\langle p$, and this is never of the form $f\rangle\langle f + \frac{1}{2}g\rangle\langle g + \frac{1}{2}h\rangle\langle h$.

The moral of all this is the following: On the one-particle level we can study density matrices $\gamma(x, x')$ or densities, $\rho(x) = \gamma(x, x)$. The former do not always come from pure states $\psi\rangle\langle\psi$. The latter do, as Theorem 1.2 shows. While γ is more complicated than ρ (it has two variables), it has the distinct advantage that the map $\Gamma \mapsto \gamma$ is *linear*! The map $\psi \mapsto \rho$ is *nonlinear*, and this, as will be seen, is the source of some difficulty.

The relation among ψ, Γ, γ, and ρ can be summarized by the following diagram:

$$\psi \mapsto \Gamma \mapsto \gamma \mapsto \rho, \tag{2.2}$$

by which we mean (i) the map $\psi \mapsto \Gamma = \psi\rangle\langle\psi$, (ii) $\Gamma \mapsto \gamma$ by (2.1) with $\psi\psi^*$ replaced by Γ, (iii) $\gamma \mapsto \gamma(x, x) = \rho(x)$. (ii) and (iii) are linear while (i) is nonlinear.

Notation. We shall use the symbol $\psi \mapsto \rho$ (or any other combination such as $\gamma \mapsto \rho$) to indicate that ψ and ρ are related by the above maps.

Technical remarks. Since γ is self-adjoint and trace class, it can always be written in the form

$$\gamma(x, x') = \sum_{j=1}^{\infty} \lambda_j f_j(x) f_j(x')^*, \tag{2.3}$$

where the f_j are orthonormal and $0 \leq \lambda_j \leq q$ (fermions), $0 \leq \lambda_j \leq N$ (bosons). If $T \in H^1(\mathbb{R}^{3N})$, by which we mean $\operatorname{Tr} -\sum_{i=1}^{N} \Delta_i \Gamma < \infty$, then each $f \in H^1(\mathbb{R}^3)$ (see Theorem 1.1). Although $\gamma(x, x)$ is not *a priori* well defined, as stated before, it is well defined almost everywhere (in $\mathbb{R}^3$) by (2.3) and

$$N = \operatorname{Tr}\gamma = \sum_{j=1}^{\infty} \lambda_j. \tag{2.4}$$

3. General Density Functional Theory

The problem that will concern us in calculating the ground state energy for N electrons interacting with each other via a repulsive Coulomb potential

$|x_i - x_j|^{-1}$ and also interacting with a single-particle potential $v(x)$. If $v = 0$, the Hamiltonian is

$$H_0 = K + \sum_{1 \le i < j \le N} |x_i - x_j|^{-1}, \tag{3.1}$$

where K is the kinetic energy operator

$$K = -\sum_{i=1}^{N} \Delta_i \tag{3.2}$$

in units in which $\hbar^2/2m = 1$. Also of interest is the case where $H_0 = K$ alone (see Sec. 4C). Recall that N is fixed and will not be mentioned unless necessary. Also, to simplify matters we shall confine our attention in the following to fermions. However, many of the following results have obvious analogs for bosons.

The total Hamiltonian is

$$H_v = H_0 + V, \tag{3.3}$$

where

$$V = \sum_{i=1}^{N} v(x_i). \tag{3.4}$$

The ground state energy $E(v)$ is *defined* to be

$$E(v) = \inf \{(\psi, H_v \psi) | \psi \in \mathscr{W}_N\}, \tag{3.5}$$

where

$$\mathscr{W}_N = \{\psi | \|\psi\| = 1, T(\psi) < \infty\}. \tag{3.6}$$

Technical remark. Something should be said about the meaning of $(\psi, H_v\psi)$ and about the class of v's under consideration. We shall always interpret $(\psi, H_v\psi)$ in the sense of a quadratic form; in particular, this means that $(\psi, K\psi) \equiv T(\psi)$. It is not assumed that $\Delta\psi \in L^2$. Since $\psi \in H^1$, it is easy to prove that $(\psi, |x_i - x_j|^{-1}\psi)$ is finite for all $i \neq j$. The part containing v is $\int \rho(x) v(x)\, dx$. As $\rho \in L^1$ and $\rho \in L^3$ (since $\nabla \rho^{1/2} \in L^2$), $\rho \in L^p$ for all $1 \le p \le 3$. The integral is then well defined if $v \in L^{3/2} + L^\infty$. This means that *we consider v's that can be written as $v = v_{3/2} + v_\infty$ with $v_{3/2} \in L^{3/2}$ and with $|v_\infty|$ a bounded function.* This choice precludes v's that go to ∞ as $|x| \to \infty$, such as the harmonic oscillator potential. Unbounded potentials can also be handled by the methods given here, but then we have to place additional restrictions on ρ so that $\int v\rho$ makes sense. We restrict ourselves here to $L^{3/2} + L^\infty$ for simplicity of exposition. The class includes Coulomb potentials because $|x|^{-1} = \theta(x)|x|^{-1} + [1 - \theta(x)]|x|^{-1}$ with $\theta(x) = 1$ if $|x| \le 1$, $\theta(x) = 0$, $|x| > 1$. The two terms on the right are in $L^{3/2}$ and in L^∞, respectively.

$L^{3/2} + L^\infty$ is a Banach space with the norm

$$\|v\| = \inf \{\|g\|_{3/2} + \|h\|_\infty | g + h = v\}. \tag{3.7}$$

Technical remark. $\mathcal{R}_N$ is a subset of the Banach space $X = L^3 \cap L^1$. X^*, the dual of X, is $Y = L^{3/2} + L^\infty$. However, the dual of Y is not X because while L^∞ is the dual of L^1, L^1 is not the dual of L^∞. However, $X \subset Y^*$. The duality will be useful.

There may or may not be a minimizing ψ for (3.5), and if there is one it may not be unique (for bosons it is unique because it is a positive function). Any minimizing ψ (called a *ground state*) would satisfy

$$H_v\psi = E(v)\psi \tag{3.8}$$

in the distributional sense. The proof of this assertion is not difficult. For example, a minimizing ψ will not exist if v is an attractive square well and if N is too large; the extra, unbound electrons will simply "leak away" to infinity. In such a case, $E(v)$ would still have physical significance. It would be the ground state energy for *fewer* than N particles.

There are three simple, but important, properties of $E(v)$:

Theorem 3.1. (i) *$E(v)$ is concave in v: that is,*

$$E(v) \geq \alpha E(v_1) + (1-\alpha)E(v_2), \tag{3.9}$$

for all v_1, v_2, $0 \leq \alpha \leq 1$ and $v = \alpha v_1 + (1-\alpha)v_2$.

(ii) *$E(v)$ is monotone decreasing: that is, if $v_1(x) \leq v_2(x)$ for all x, then $E(v_1) \leq E(v_2)$.*

(iii) *$E(v)$ is continuous in the $L^{3/2} + L^\infty$ norm and is, moreover, locally Lipschitz. In particular, $E(v)$ is finite.*

Proof. (i) If $\psi \in \mathcal{W}_N$, then

$$(\psi, H_v\psi) = \alpha(\psi, H_{v_1}\psi) + (1-\alpha)(\psi, H_{v_2}\psi) \geq \alpha E(v_1) + (1-\alpha)E(v_2).$$

(ii) $(\psi, H_{v_2}\psi) \geq (\psi, H_{v_1}\psi) \geq E(v_1)$.

(iii) Fix v_0 and let $\delta = v - v_0$. We want to show that when $\|\delta\| \leq L/3$, $|E(v) - E(v_0)| < C\|\delta\|$ for some C, independent of v. [Here, L is the constant in (1.10).] Since $E(v)$ is concave, it is sufficient to show that for some fixed D, $E(v) - E(v_0) \geq D$ whenever $\|\delta\| = L/3$; because if $0 \leq \gamma \leq 1$,

$$\gamma[E(v_0+\delta) - E(v_0)] \leq E(v_0 + \gamma\delta) - E(v_0) \leq \gamma[E(v_0) - E(v_0 - \delta)].$$

Let $E(v, \frac{1}{2})$ denote (3.5) with K replaced by $K/2$. Then

$$E(v) \geq E(v_0, \tfrac{1}{2}) + \inf\,(\psi, [\tfrac{1}{2}K + \Sigma\delta(x_i)]\psi).$$

The last term is bounded by $-LN/2$ because $\delta = g + h$ with $\|g\|_{3/2} < L/2$ and $\|h\|_\infty < L/2$. Thus,

$$\int \delta\rho > -(L/2)[\|\rho\|_3 + N].$$

But $(\psi, K\psi)/2 \geq (L/2)\|\rho\|_3$ by (1.10) and Theorem 1.1. Finally, note that $E(v_0, \frac{1}{2}) - E(v_0)$ is a constant, D', independent of v. ■

Now we begin the study of density functional theory in the manner of Hohenberg and Kohn. Their work is based on the following theorem [3]:

Theorem 3.2. *Suppose ψ_1 (respectively, ψ_2) is a ground state for v_1 (respectively v_2) and $v_1 \neq v_2$+constant. Then $\rho_1 \neq \rho_2$.*

Proof. Suppose $\rho_1 = \rho_2 = \rho$. $\psi_1 \neq \psi_2$ because they satisfy different Schrödinger equations, (3.8). [Note. To prove this we must know that $v_1\psi = v_2\psi$ implies that $v_1 = v_2$. This, in turn, requires that $\psi(x)$ does not vanish on a set of positive measure. This technical point is discussed in remark (ii) preceding Theorem 3.5.] Moreover, ψ_2 (respectively, ψ_1) does not satisfy (3.8) for v_1 (respectively, v_2). Therefore,

$$E(v_1) < (\psi_2, H_{v_1}\psi_2) = E(v_2) + \int (v_1 - v_2)\rho.$$

Likewise, $E(v_2) < E(v_1) + \int (v_2 - v_1)\rho$. This is a contradiction. ■

Hohenberg and Kohn assume that every ρ comes from some ψ that is a ground state for some v. For such ρ they *define* the functional

$$F_{\mathrm{HK}}(\rho) = E(v) - \int v\rho, \tag{3.10}$$

and we shall retain this definition for $\rho \in \mathscr{A}_N$, where

$$\mathscr{A}_N = \{\rho \,|\, \rho \text{ comes from a ground state}\}. \tag{3.11}$$

$\mathscr{A}_N \neq \mathscr{I}_N$, as remarked earlier, and it is not convex (see Theorem 3.4)! The definition given by (3.10) requires Theorem 3.2, according to which there is a unique v (up to a constant) associated with ρ. We can also define

$$\mathscr{V}_N = \{v \,|\, H_v \text{ has a ground state}\}. \tag{3.12}$$

It then follows easily that for $v \in \mathscr{V}_N$

$$E(v) = \min\left\{F_{\mathrm{HK}}(\rho) + \int v\rho \,|\, \rho \in \mathscr{A}_N\right\}. \tag{3.13}$$

This is *the* HK *variational principle*, but it is important to note that it holds only for $v \in \mathscr{V}_N$, which is unknown, and that the variation is restricted to the unknown set $\mathscr{A}_N$.

We also do not know what F_{HK} is, and that is a very serious problem. But there are also conceptual problems, which will be addressed here.

If F is to be used in a variational principle, it is clearly desirable that F be a convex functional. In particular, it should be defined everywhere on $\mathscr{I}_N$, or at least on some known convex subset of $\mathscr{I}_N$.

The domain of F_{HK} (i.e., $\mathscr{A}_N$) is not all of $\mathscr{I}_N$ and it is not convex. This last fact is closely connected with the following difficulty: One can define a functional for all ρ in $\mathscr{I}_N$ by*

$$\tilde{F}(\rho)=\inf\{(\psi, H_0\psi)|\psi\mapsto\rho,\ \psi\in\mathscr{W}_N\}. \tag{3.14}$$

It then follows trivially that

$$E(v)=\inf\left\{\tilde{F}(\rho)+\int v\rho\,|\rho\in\mathscr{I}_N\right\}, \tag{3.15}$$

$$\tilde{F}(\rho)=F_{\mathrm{HK}}(\rho), \qquad \text{if} \quad \rho\in\mathscr{A}_N. \tag{3.16}$$

So far, so good. The difficulty is that $\tilde{F}$ is not convex either. However, $\tilde{F}$ has one important property that is proved in the appendix.

Theorem 3.3. *For each ρ in $\mathscr{I}_N$ there is a $\psi\in\mathscr{W}_N$ such that $\tilde{F}(\rho)=(\psi, H_0\psi)$. In other words, the infimum in* (3.14) *is a minimum.*

The following functional F is one choice for "the density functional" that remedies the difficulties mentioned so far:

$$F(\rho)=\sup\left\{E(v)-\int v\rho\,|v\in L^{3/2}+L^{\infty}\right\}. \tag{3.17}$$

We shall explore the properties of F, but it, too will be seen to have subtle difficulties of its own.

Remarks. (i) (3.17) defines $F(\rho)$ for all $\rho\in\mathscr{R}_N$, not just $\mathscr{I}_N$, provided F is interpreted in the extended sense as a function that can have the value $+\infty$. In fact, (2.17) defines F on the much larger set $X=L^3\cap L^1$, without the restrictions $\rho(x)\geq 0$ and $\int\rho=N$. As Theorem 3.5 shows, however, it is only necessary to consider F on the convex subset $\mathscr{I}_N$ of X.

(ii) Recall that F depends explicitly on N through E.

(iii) Since F is the supremum of a family of linear functionals, it is convex.

(iv) Theorem 3.8 shows that $F(\rho)=+\infty$ if $\rho\notin\mathscr{I}_N$. There is an alternative definition of F, namely, F', by which F' is finite on the set

$$\mathscr{I}'_N\equiv\{\rho|\rho(x)\geq 0 \text{ and } \nabla\rho^{1/2}\in L^2\},$$

without requiring $\int\rho=N$. This is

$$F'(\rho)=\left(\int\rho\right)F\left(\rho\Big/\int\rho\right), \qquad \rho\not\equiv 0,$$

$$F'(0)=0. \tag{3.18}$$

It is easy to check that the convexity and lower semicontinuity (a concept to be defined later) of F carry over to F'. This definition has the virtue that F' is finite

* Levy [10] also defined $\tilde{F}(\rho)$ which he called Q, and derived (3.15). He did not prove Theorem 3.3, but assumed the existence of a minimizing ψ. Also, he did not establish the connection between $\tilde{F}$ and the Legendre transform, F (Theorem 3.7). In Ref. 11, Levy proved Theorem 3.4(ii), independently and virtually at the same time as myself, using essentially the same construction. See Ref. 12 for additional remarks about Q.

on a dense subset of the set of nonnegative functions in X. However, this does not change the theory in any important way, so we shall continue to use the definition given by (3.17).

(v) Other characterizations of F, directly in terms of $\tilde{F}$, are given in Theorem 3.7, and in Eqs. (4.5)–(4.7).

There is an obvious relation between F and $\tilde{F}$, namely,

$$F(\rho) \le \tilde{F}(\rho) \qquad \text{for all} \quad \rho \in \mathscr{I}_N, \tag{3.19}$$

since $E(v) \le \tilde{F}(\rho) + \int v\rho$ for all $\rho \in \mathscr{I}_N$. Furthermore, since F is convex and $\tilde{F}$ is not convex (by Theorem 3.4), there are ρ's in $\mathscr{I}_N$ for which $F(\rho) < \tilde{F}(\rho)$.

First we prove that not all ρ's come from ground states. The essential ingredient is the existence of v with a degenerate ground state. (Such v's, incidentally, preclude the existence of a map $v \mapsto \rho$.)

Theorem 3.4. *Let $N > q$ = number of spin states. Then*

(i) *$F(\rho)$) is not convex*

(ii) *There exists a $\rho \in \mathscr{I}_N$ that does not come from a ground state ψ. Moreover this ρ is a convex combination of ρ's that do come from a ground state.*

Proof. Let v be a spherically symmetric potential having a ground state and with the property that its ground state has orbital angular momentum $L \ge 1$. We assume the degeneracy is no greater than necessary, namely $M = 2L + 1$. The orthonormal ground states are $\psi_1, \ldots, \psi_M$ and $\psi_i \mapsto \rho_i$. Under simultaneous rotation of all N coordinates, they transform as a basis for the M-dimensional irreducible representation of $O(3)$. The following fact is easy to prove: (a) If $\bar{\rho} = M^{-1}\Sigma\rho_i$, then $\bar{\rho}(x)$ is spherically symmetric: that is, $\bar{\rho}$ depends only on $r = |x|$.

A second fact that will be needed is (b): if ϕ is any ground state (and hence a linear combination of the ψ_i) and $\phi \mapsto \rho$, then ρ is not spherically symmetric. This fact must follow from some group-theoretic agreement, but I have not found one. However, it is not hard to see that (b) is equivalent to (c): There exists a perturbation of v, $v(x) \mapsto v(x) + \lambda w(x)$, with w bounded and of compact support, so that to first order in λ the M-fold degeneracy is broken. Such pairs v and w certainly exist, so we can henceforth assume that (b) holds. [A proof that a v satisfying (b) exists is the following. First, take the case that $H_0 = K$, that is, independent particles. The ground states are determinants. Choose v so that the ground state has $L \ge 1$, in which case (b) obviously holds. Next, consider $H = K + \lambda\Sigma|x_i - x_j|^{-1} + V$. Angular momentum is still conserved and for sufficiently small λ the ground state will have the same L and, by continuity of the ground states, (b) will continue to hold for small λ. We are interested in $\lambda = 1$ but, under the scaling

$$x \to x/\lambda, \qquad v(x) \to \lambda^{-2} v(x/\lambda) = v'(x),$$

the v, λ problem is converted into the v', $\lambda = 1$ problem. Thus, v' has the desired properties. I thank B. Simon for this remark.]

Clearly $\tilde{F}(\rho_i)=\text{constant}=D=E(v)-\int v\rho_i$. We claim $\tilde{F}(\bar{\rho})>D$, thereby proving lack of convexity. Obviously, $\tilde{F}(\bar{\rho})\geq D$, for otherwise we could use $\bar{\rho}$ instead of ρ_i in (3.15). (Note: $\int v\bar{\rho}=\int v\rho_i=\text{constant}=C$.) Suppose $\tilde{F}(\bar{\rho})=D$. Then $\bar{\rho}$ comes from some ϕ that must be a ground state for v. But (a) and (b) show this to be impossible. Thus $\tilde{F}(\bar{\rho})>D$. Moreover, $\bar{\rho}$ cannot come from any ground state ψ for any other v'. If it did, then

$$E(v')=\tilde{F}(\bar{\rho})+\int v'\bar{\rho}>M^{-1}\sum\left(\tilde{F}(\rho_i)+\int v'\rho_i\right).$$

This implies that for some $1\leq i\leq M$, $\tilde{F}(\rho_i)+\int v'\rho_i<E(v')$, which is a contradiction. ■

Remarks. (i) The foregoing proof holds just as well if $(\psi, H_0\psi)$ is replaced by $T(\psi)$ in the definition of $\tilde{F}$. [This functional will be denoted by $\tilde{T}(\rho)$ and the analog of (3.17) by $T(\rho)$.] In other words, the interelectron Coulomb repulsion plays no role in Theorem 3.4 (see Sec. 4C).

(ii) There are other ρ's that do not come from some v, namely, those $\rho\in\mathcal{I}_N$ that vanish on a nonempty open set. If $v\in L^{3/2}+L^\infty$ and ψ is its ground state, then ψ cannot vanish in an open set by the unique continuation theorem [18]. (Strictly speaking, this theorem is only known to hold for $v\in L^3_{\text{loc}}$, but it is believed to hold for $L^{3/2}+L^\infty$.) Presumably, such ρ's can, in many cases, be obtained as limits in which $v\to\infty$ on the open set. Therefore, if the set of allowed v's can be extended properly to include infinite v's, the existence of such ρ's may not have any particular importance. The question is very delicate, however, as Englisch and Englisch [7] showed recently. Even for one particle there are densities which never vanish but which do not come from any v, even if one allows density matrices (see Sect. 4B) instead of pure states. These densities have regions in which they are "small" so that the obvious v (defined by $v=\rho^{-1/2}\Delta\rho^{1/2}$) has the property that $-\Delta+v$ cannot be defined as a semibounded operator.

Theorem 3.5.

$$E(v)=\inf\left\{F(\rho)+\int\rho v\,|\,\rho\in L^3\cap L^1\right\},\tag{3.20}$$

$$E(v)=\inf\left\{F(\rho)+\int\rho v\,|\,\rho\in\mathcal{I}_N\right\}.\tag{3.21}$$

Remark. The right sides of (3.20) and (3.21) are automatically concave functionals, which is a property we already proved for E.

Proof. Let $M^-(v)$ [respectively, $M^+(v)$] be the infimum in (3.20) [respectively, (3.21)]. Obviously, $M^-(v)\leq M^+(v)$. First, pick v_0. Clearly $F(\rho)\geq E(v_0)-\int v_0\rho\equiv F_1(\rho)$. Therefore,

$$M^-(v)\geq\inf\left\{F_1(\rho)+\int\rho v\,|\,\rho\in L^3\cap L^1\right\}.$$

and hence $M^-(v_0) \geq E(v_0)$. Second, by (3.19), $F(\rho) \leq \tilde{F}(\rho)$, so that $M^+(v) \leq \inf\{\tilde{F}(\rho) + \int \rho v \mid \rho \in \mathscr{I}_N\} = E(v)$. ■

Let us pause briefly to review the situation. Three functionals have been defined:

$$\begin{aligned} &F_{\mathrm{HK}} && \text{defined on} \quad \mathscr{A}_N \subset \mathscr{I}_N; \\ &\tilde{F} && \text{defined on} \quad \mathscr{I}_N \subset L^3 \cap L^1; \\ &F && \text{defined on} \quad X = L^3 \cap L^1. \end{aligned}$$

Of these, only F is convex and only $\tilde{F}$ and F satisfy the variational principle for *all* v.

The next step is to find out something of the nature of F. It is at this point that the analysis becomes complicated and where difficulties and incompleteness arise. The basic reason is that the connection between v and ρ is anything but simple. We have $X = L^3 \cap L^1$ and its dual $X^* = L^{3/2} + L^\infty$. Although X is not the dual of X^*, it is a subset of X^{**}, the dual of X^*.

Definitions. (i) A sequence $\rho_n \in X$ is said to converge to $\rho \in X (\rho_n \to \rho)$ if and only if $\|\rho_n - \rho\|_3 \to 0$ and $\|\rho_n - \rho\|_1 \to 0$. This is also called *norm convergence*. ρ_n *converges weakly* to $\rho (\rho_n \rightharpoonup \rho)$ if and only if $\int v(\rho_n - \rho) \to 0$ for all $v \in Y = X^*$. Clearly, strong convergence implies weak convergence.

(ii) A functional f on X is *continuous* (or norm continuous) if and only if $\rho_n \to \rho$ implies $f(\rho_n) \to f(\rho)$. *Weak continuity* requires the concept of nets to define but, if f is weakly continuous, then whenever $\rho_n \rightharpoonup \rho$, $f(\rho_n) \to f(\rho)$. Weak continuity implies norm continuity.

(iii) A real functional f on X is *lower semicontinuous* (l.s.c.) if and only if $\rho_n \to \rho$ implies $f(\rho) \leq \liminf f(\rho_n)$. *Weak lower semicontinuity* requires nets to define, but if f is weakly l.s.c. then $\rho_n \rightharpoonup \rho$ implies $f(\rho) \leq \liminf f(\rho_n)$. (Weak) lower semicontinuity is equivalent to the following: $\{\rho \mid f(\rho) \leq \lambda\}$ is (weakly) closed for all real λ.

Remarks. (i) Weak lower semicontinuity always implies lower semicontinuity, but not conversely. It is a theorem of Mazur [19], however, that if f is *convex* and norm l.s.c., then it is automatically weakly l.s.c.

(ii) The function $\rho(x) \equiv 0$ is *not* in the $L^3 \cap L^1$ weak closure of $\mathscr{I}_N$.

The reader may be puzzled by all these definitions, especially lower semicontinuity, because finite convex functions on $\mathbb{R}^n$ are always continuous. Unfortunately, this is not true in infinite-dimensional spaces such as the space X we are considering. Even l.s.c. cannot be taken for granted.

Theorem 3.6. $F(\rho)$ *is weakly (and hence also norm) lower semicontinuous.*

Proof.

$$K_\lambda \equiv \{\rho \mid F(\rho) \leq \lambda\} = \left\{\rho \,\middle|\, E(v) - \int v\rho \leq \lambda \text{ for all } v \in Y\right\}.$$

Now if $\rho_n \to \rho$ in norm and $\rho_n \in K_\lambda$ then, for each $v \in Y$.

$$E(v) - \int v\rho = \lim \left(E(v) - \int v\rho_n \right) \le \lambda.$$

Therefore K_λ is norm closed, so that F is norm l.s.c. Weak l.s.c. is a consequence of Mazur's theorem. ∎

Next we define the *convex envelope* (CE).

Definition. Let f be a real functional defined on a subset A of X. $f(\rho)$ is allowed to be $+\infty$, but not for all $\rho \in A$. CE f is defined on all of X as follows: CE $f(\rho) = \sup \{g(\rho) | g$ is weakly l.s.c., g is convex on X, and $g(\rho') \le f(\rho')$ for all $\rho' \in A\}$.

It is easy to check that CE f is convex and weakly l.s.c. and CE $f(\rho) \le f(\rho)$ for all $\rho \in A$. However, CE $f(\rho)$ may be $+\infty$ for some ρ.

The function of interest is CE $\tilde{F}$ with $A = \mathscr{I}_N$. Note that A is convex and that $\tilde{F}$ (and hence CE $\tilde{F}$) is finite on A by Theorem 1.2. Since CE $\tilde{F} \le \tilde{F}$ on A, it is obvious from (3.19) and Theorem 3.6 that $F \le$ CE $\tilde{F}$ on X. On the other hand, suppose we use CE $\tilde{F}$ instead of $\tilde{F}$ in (3.15). This gives a new function, which we call E'. Clearly $E' \le E$. Then, if E' is used in (3.17), we get a new function F', and $F' \le F$. However, an infinite-dimensional generalization of Fenchel's theorem [29] (which uses the Hahn–Banach theorem) states that if the original function (in our case, CE $\tilde{F}$) is convex and weakly l.s.c. on X, then its double Legendre transform (in our case F') is equal to the original function. Thus, $F' =$ CE $\tilde{F}$ and we have

Theorem 3.7. *$F(\rho) =$ CE $\tilde{F}(\rho)$ for all $\rho \in L^3 \cap L^1$.*

The reader may wonder what Theorem 3.7 is good for; the following is an example of the usefulness of the foregoing functional analysis (see Theorem 4.3).

Theorem 3.8. *For all $\rho \in L^3 \cap L^1$ let*

$$\begin{aligned} G(\rho) &\equiv \int (\nabla \rho(x)^{1/2})^2 \, dx \qquad && \textit{if } \rho \in \mathscr{I}_N \\ &= +\infty && \textit{otherwise.} \end{aligned}$$

Then $F(\rho) \ge G(\rho)$, for all $\rho \in L^3 \cap L^1$.

Proof. G is obviously convex on X [see the remark after (1.10)]. We claim that G is norm l.s.c. (Note: The norm in question is $L^3 \cap L^1$, not the H^1 norm on $\rho^{1/2}$.) If so, we are done because G is then weakly l.s.c. and, by Theorem 1.1, $G \le \tilde{F}$ on $\mathscr{I}_N$: but then $G \le$ CE $\tilde{F} = F$.

To prove norm l.s.c., let ρ_n be any sequence in X with $\rho_n \to \rho$; that is, $\|\rho_n - \rho\|_1 \to 0$ and $\|\rho_n - \rho\|_3 \to 0$. We can assume that $G = \lim G(\rho_n)$ exists and is finite, and we have to show that $G \ge G(\rho)$. We can also assume $\rho(x) \ge 0$ a.e. because if $\rho < 0$ on a set S of positive measure, then, for sufficiently large n,

$\rho_n < 0$ on some set of positive measure; hence $\rho_n \notin \mathscr{I}_N$ and $G(\rho_n) = \infty$. For a similar reason we can assume $\int \rho = N$. Since $G(\rho_n) < \infty$, $\rho_n \in \mathscr{I}_N$. Thus, if we define $g_n = \rho_n^{1/2}$ and $g = \rho^{1/2}$, we have: (a) g_n is bounded in H^1; (b) $g_n^2 \to g^2$ in L^3 and L^1. By the Banach–Alaoglu theorem there is an $f \in H^1$ such that $g_n \rightharpoonup f$ and $\nabla g_n \rightharpoonup \nabla f$ weakly in L^2. Clearly $f(x) \geq 0$. It is not hard to prove that if $g_n \rightharpoonup f$ in L^2 and $g_n^2 \to g^2$ in L^1, then $g = f$. Hence $\nabla g = \nabla f$, and thus $\nabla g_n \rightharpoonup \nabla g$. But since $\int (\nabla g)^2$ is H^1-norm continuous, it is H^1 weakly l.s.c., so that $\lim G(\rho_n) \geq G(\rho)$. ■

Theorem 3.8 is certainly not obvious. Among other things it says that if $\rho \notin \mathscr{I}_N$ (and such ρ's can be quite smooth and innocent looking), then there exists a sequence of potentials $v_n \in L^{3/2} + L^\infty$ such that $E(v_n) - \int v_n \rho \to \infty$. The reader is asked to reflect on this fact. Another interesting fact is that F is convex and finite on $\mathscr{I}_N$, but infinite off $\mathscr{I}_N$. However, the complement of $\mathscr{I}_N$ (in X) is dense (in the X norm) in $\mathscr{I}_N$ and $\mathscr{I}_N$ is dense in the cone of nonnegative functions in X.

The following upper bound complements Theorem 3.8.

Theorem 3.9. *If $\rho \in \mathscr{I}_N$, then*

$$F(\rho) \leq \tilde{F}(\rho) \leq (4\pi)^2 N^2 G(\rho) + \tfrac{1}{2} \iint \rho(x)\rho(y)|x-y|^{-1}\, dx\, dy. \tag{3.22}$$

Proof. Use the definition (3.14). By Theorem 1.2 there is a determinantal ψ, with $\psi \mapsto \rho$, such that (1.12) holds. With this ψ we can calculate the Coulomb repulsion $I = (\psi, \Sigma |x_i - x_j|^{-1} \psi)$. I has a direct term, given in (3.22), plus an exchange term. The latter is negative, as is well known, since $|x-y|^{-1}$ is a positive definite kernel. Thus $\tilde{F}(\rho) \leq$ right side of (3.22). Then use (3.19). ■

Remark. By one of Sobolev's inequalities,

$$D \equiv \iint \rho(x)\rho(y)|x-y|^{-1}\, dx\, dy \leq (\text{const.}) \|\rho\|_{6/5}^2.$$

By Hölder's inequality $\|\rho\|_{6/5}^4 \leq \|\rho\|_1^3 \|\rho\|_3$. Again, by Sobolev's inequality, (1.10), $\|\rho\|_3 \leq (\text{const.}) G(\rho)$. Thus,

$$D \leq (\text{const.}) N^{3/2} G(\rho)^{1/2} \leq (\text{const.})[N + N^2 G(\rho)].$$

To continue the study of F the following concept is needed.

Definition. Let f be a real functional on a subset A of a Banach space B, and let $\rho_0 \in A$. A linear functional l on B is said to be a *tangent functional* (TF) *at* ρ_0 if and only if for all $\rho \in A$

$$f(\rho) \geq f(\rho_0) - l(\rho - \rho_0). \tag{3.23}$$

l may not be unique. If l is continuous, then l is a *continuous tangent functional* at ρ_0.

l is a continuous linear functional on X if and only if it has the form $\int v\rho$ with $v \in X^* = Y$. If f is convex, then at every point ρ_0 at which f is finite, f has at least one TF. This is guaranteed by the Hahn–Banach theorem. However, f may have no *continuous* TF at ρ_0.

The functional of interest is obviously F. In general, $F \leq \tilde{F}$, but the following says something about those ρ for which $F(\rho) = \tilde{F}(\rho)$.

Theorem 3.10. *Let $\rho_0 \in \mathscr{I}_N$. The following are equivalent*:
(1) *$F(\rho_0) = \tilde{F}(\rho_0)$ and F has a continuous TF at ρ_0.*
(2) *$\rho_0 \in \mathscr{A}_N$.*
(3) *$\tilde{F}$ has a continuous TF at ρ_0; that is, $\tilde{F}(\rho) \geq \tilde{F}(\rho_0) - \int v(\rho - \rho_0)$ for $\rho \in \mathscr{I}_N$.*
(4) *(3) and (5) hold with the same v.*
(5) *$E(v) = \tilde{F}(\rho_0) + \int v\rho_0$ for some v.*
(6) *(5) holds and, in addition, $v \in \mathscr{V}_N$ and $v \mapsto \rho_0$.*
(7) *$\tilde{F}$ has a continuous TF at ρ_0 and v is unique up to a constant. Moreover, F has the same continuous TFs at ρ_0, and no others.*

Proof. (1)$\Rightarrow$(3): For $\rho \in \mathscr{I}_N$,

$$\tilde{F}(\rho) \geq F(\rho) \geq F(\rho_0) - \int v(\rho - \rho_0) = \tilde{F}(\rho_0) - \int v(\rho - \rho_0).$$

(3)$\Rightarrow$(4): Let $F_1(\rho) = \tilde{F}(\rho_0) - \int v(\rho - \rho_0) \leq \tilde{F}(\rho)$. Then

$$\tilde{F}(\rho_0) + \int v\rho_0 \geq E(v) \geq \inf\left\{F_1(\rho) + \int \rho v \,|\, \rho \in \mathscr{I}_N\right\} = \tilde{F}(\rho_0) + \int v\rho_0.$$

(4)$\Rightarrow$(5), (7)$\Rightarrow$(3), (6)$\Rightarrow$(5): All trivial.

(5)$\Rightarrow$(1): $\tilde{F}(\rho_0) + \int \rho_0 v = E(v) \leq F(\rho_0) + \int \rho_0 v \Rightarrow F(\rho_0) \geq \tilde{F}(\rho_0) \Rightarrow F(\rho_0) = \tilde{F}(\rho_0)$. Then, for all $\rho \in X$, $F(\rho) + \int \rho v \geq E(v) = F(\rho_0) + \int \rho_0 v$.

(2)$\Rightarrow$(5): By (3.16).

(5)$\Rightarrow$(2), (6): By Theorem 3.3, $\tilde{F}(\rho_0) = (\psi, H_0\psi)$ for some ψ with $\psi \mapsto \rho_0$. Then $E(v) = (\psi, H_0\psi) + \int v\rho_0 \Rightarrow \rho_0 \in \mathscr{A}_N$, $v \in \mathscr{V}_N$, and $v \mapsto \rho_0$. Thus (1)–(6) are equivalent and (7)$\Rightarrow$(3). Now we show that (1)–(6)$\Rightarrow$(7). If v is a continuous TF for F, then v is a continuous TF for $\tilde{F}$ [by the proof of (1)$\Rightarrow$(3)]. If v is a continuous TF for $\tilde{F}$, then $F(\rho) \geq E(v) - \int v\rho$, so v is a continuous TF for F. Suppose $\tilde{F}$ has two continuous TFs v and w with $v - w \neq$ constant. Then $E(v) = \tilde{F}(\rho_0) + \int v\rho_0$ and $E(w) = \tilde{F}(\rho_0) + \int w\rho_0$. Since $\rho_0 \in \mathscr{A}_N$, this is impossible by Theorem 3.2. ■

It should be noted that the only place that the HK Theorem 3.2 entered in the analysis of F was in establishing the uniqueness (modulo constants) in (7).

Now we turn to two important questions whose answers we cannot give but that are obviously important for the theory. We replaced F_{HK} by F because F_{HK} was not defined on all of $\mathscr{I}_N$. Theorem 3.10 states that on $\mathscr{A}_N$, where F_{HK} is defined, $F = \tilde{F} = F_{\text{HK}}$ and F has an essentially unique *continuous* TF.

Question 5. For which points of $\mathscr{I}_N$ does F have a continuous TF? Where there is one, is it unique (modulo adding a constant to v)?

Question 6. If F has a continuous TF at $\rho_0 \in \mathscr{I}_N$ given by some $v \in L^{3/2} + L^\infty$, is this $v \in \mathscr{V}_N$?

Questions 5 and 6 have alternative formulations, given below.

Theorem 3.11. *Let $\rho_0 \in \mathscr{I}_N$ and $v \in L^{3/2} + L^\infty$. v is not necessarily in $\mathscr{V}_N$. Then, for all ρ,*

$$F(\rho) \geq F(\rho_0) - \int v(\rho - \rho_0) \qquad (\textit{continuous } \mathrm{TF}) \tag{3.24}$$

if and only if

$$E(v) = F(\rho_0) + v \int \rho_0 \qquad [\textit{minimum in } (3.21)]. \tag{3.25}$$

Proof. Assume (3.24) and let $M_v(\rho)$ be its right side. Then

$$E(v) \geq \inf \left\{ M_v(\rho) + \int \rho v \right\} = F(\rho_0) + \int v\rho_0 \geq E(v).$$

For the converse,

$$F(\rho) + \int v\rho \geq E(v) = F(\rho_0) + \int v\rho_0. \quad \blacksquare$$

Question 5 is equivalent to the following: For which $\rho_0 \in \mathscr{I}_N$ is there a v such that (3.25) holds? Is this v unique (up to constants)? Question 6 is the following: If (3.25) holds, is $v \in \mathscr{V}_N$?

Some insight into the continuous TFs of F are provided by the Bishop–Phelps theorem. We refer the reader to Ref. 20 for this as well as other interesting facts about convexity. A definition is needed.

Definition. Let F be a real functional on a real Banach space B with dual B^* (the set of continuous linear functionals on B). $b^* \in B^*$ is said to be *F-bounded* if there is a constant C (depending on b^* but not on b) such that $F(b) \geq b^*(b) + C$ for all $b \in B$.

In our case $B = X$ and F is our density functional.

Theorem 3.12. *Every $v \in X^* = L^{3/2} + L^\infty$ is F bounded.*

Proof. By Theorem 3.8, $F(\rho) = \infty$ if $\rho \notin \mathscr{I}_N$, so we only have to consider $\rho \in \mathscr{I}_N$ and prove that $G(\rho) \geq \int v\rho + C$ for some C. The proof of this is identical to the last part of the proof of Theorem 3.1. $\blacksquare$

The Bishop–Phelps theorem is the following.

Theorem 3.13. *Let F be a l.s.c. convex functional on a real Banach space B. (Note: Norm and weak l.s.c. are identical.) F can take the value $+\infty$, but not everywhere. Then*

(i) *The continuous tangent functionals to F (over all of B) are B^*-norm dense in the set of F-bounded functionals in B^**

(ii) *Suppose $b_0 \in B$ and $b_0^* \in B^*$ with $F(b_0) < \infty$. For every $\varepsilon > 0$ there exists $b_\varepsilon \in B$ and $b_\varepsilon^* \in B^*$ such that $\|b_\varepsilon^* - b_0^*\|_{B^*} \leq \varepsilon$ and*

$$\varepsilon \|b_\varepsilon - b_0\|_B \leq F(b_0) + b_0^*(b_0) - \inf\{F(x) + b_0^*(x) | x \in B\}.$$

Moreover, b_ε^* is tangent to F at b_ε, namely $F(b) \geq F(b_\varepsilon) - b_\varepsilon^*(b - b_\varepsilon)$ for all b.

The significance of Theorem 3.13(i) is the following. There are certainly many v's in Y that are not in $\mathcal{V}_N$. (Example: Suppose $v \in L^{3/2}$ and $\|v\|_{3/2} < L$, where L is the constant in (1.10). Then $(\psi, H_v\psi) > 0$ for all ψ, but $E(v) = 0$ because we can always take a sequence ψ_n that "leaks away to infinity.") Let $v \notin \mathcal{V}_N$, whence $\tilde{F}(\rho) + \int v\rho$ does not have a minimum. What Theorem 3.13 says is that there always exists a sequence v_n (not necessarily in $\mathcal{V}_N$) such that

(a) $F(\rho) + \int \rho v_n$ has a minimum at some $\rho_n \in \mathcal{I}_N$ and this minimum is $E(v_n)$.
(b) $F(\rho) \geq F(\rho_n) - \int v_n(\rho - \rho_n)$ for all ρ.
(c) $v_n \to v$ in the $L^{3/2} + L^\infty$ norm.

Point (c) means the following: $v_n = v + g_n + h_n$ with $\|g_n\|_{3/2} \to 0$ and $\|h_n\|_\infty \to 0$. In particular, if $v \in L^{3/2}$ with $\|v\|_{3/2} < L$, then $\|v + g_n\| < L$ for large n. Hence $v + g_n \notin \mathcal{V}_N$. If $v_n \in \mathcal{V}_N$, then it can only be because of the (vanishingly small) L^∞ piece h_n.

One consequence of Theorem 3.13(ii) is the following.

Theorem 3.14. *Let $\rho_0 \in \mathcal{I}_N$. Then there exists a sequence $\rho_n \in \mathcal{I}_N$ such that*

(i) *$\rho_n \to \rho_0$ in $L^3 \cap L^1$ norm.*
(ii) *F has a continuous* TF *at ρ_n.*

Proof. Given $n > 0$, by (3.17) there exists v_n such that $E(v_n) - \int \rho_0 v_n > F(\rho_0) - 1/n$. Hence

$$F(\rho) \geq E(v_n) - \int \rho v_n \geq F(\rho_0) - \int v_n(\rho - \rho_0) - 1/n.$$

Take $\varepsilon = 1$ in Theorem 3.13. There exists $w_n \in Y$ such that w_n is a continuous TF at some ρ_n and

$$\|\rho_n - \rho_0\| \leq F(\rho_0) + \int v_n \rho_0 - Z,$$

with

$$Z = \inf\left\{F(\rho) + \int v_n \rho | \rho \in X\right\}.$$

By the above,

$$Z \geq F(\rho_0) + \int \rho_0 v_n - n^{-1}. \quad \blacksquare$$

4. Additional Remarks about Density Functionals

A. The N-Dependence of F

As was stressed earlier, any functional F that satisfies (3.20) or (3.21) must depend explicitly on the particle number N. This fact is unavoidable and frequently overlooked. Let us denote the N dependence by $F(N, \rho)$. It might be hoped that F is jointly convex in N and ρ in the sense that for $N \geq 2$

$$F(N+1, \rho_1)+F(N-1, \rho_2) \geq 2F(N, \tfrac{1}{2}\rho_1+\tfrac{1}{2}\rho_2). \tag{4.1}$$

This convexity definitely does not hold as a general feature, as will be demonstrated.

The importance of convexity is shown by the following.

Theorem 4.1. *Consider the following two statements about any two functionals, F and E:* (i) *$F(N, \rho)$ is jointly convex in N and ρ in the sense of* (4.1). (ii) *$E(N, v)$ is convex in N for all fixed v; that is, for $N \geqq 2$*

$$E(N+1, v)+E(N-1, v) \geq 2E(N, v). \tag{4.2}$$

(a) *If* (3.17) *holds, then* (ii) *implies* (i).
(b) *If either* (3.20) *or* (3.21) *holds, then* (i) *implies* (ii).

Proof. (a) For each v, $E(N, v)-\int v\rho$ is jointly (N, ρ) convex. By (3.17), $F(N, \rho)$ is the supremum of such convex functions and hence is convex.

(b) Pick $\varepsilon > 0$. For $N+1$ there is a ρ_+ such that

$$A \equiv F(N+1, \rho_+)+\int \rho_+ v \leq E(N+1, v)+\varepsilon.$$

Likewise,

$$B \equiv F(N-1, \rho_-)+\int \rho_- v \leq E(N-1, v)+\varepsilon.$$

For the N problem, define $2\rho = \rho_+ + \rho_-$. Then

$$2\left\{F(N, \rho)+\int \rho v\right\} \leq A+B.$$

Since this holds for all $\varepsilon > 0$, (ii) is proved. ■

Equation (4.2) has a simple physical meaning. The ionization potential increases as the number of electrons is decreased. This is intuitively expected to be true, but if it is true, it must be because of some special property of the Coulomb repulsion. A non-Coulombic counterexample is given below.

The kinetic energy functional $\tilde{T}(N, \rho)$ is not even convex in ρ (Theorem 3.4), but the Legendre transform $T(N, \rho)$ is jointly convex. This is so because $E(N, v)$ is indeed convex in N for independent particles as the explicit expression for E, as the sum of the first N eigenvalues (counted with an extra multiplicity q) shows.

What about the convexity of F when the Coulomb repulsion is included? While it has been conjectured that $E(N, v)$ is convex in N (for *all* v) in the case of Coulomb repulsion, this has never been proved. It has not even been proved that $E(3, v)+E(1, v) \geq 2E(2, v)$.

Lest the reader think that convexity in N is a general feature, we present a counterexample. Replace $|x|^{-1}$ by the hard-core repulsion $\theta(x)=\infty$ if $|x|<1$ and $\theta(x)=0$ otherwise. Pick four distinct points x_0, y_1, y_2, y_3 in $\mathbb{R}^3$ such that $|y_i-y_j|>1$ for all $i\neq j$ but $|x_0-y_i|<1$ for all i. Let $v(x)\doteq-2\lambda<0$ in small balls about the y_i, $v(x)=-3\lambda$ in a small ball about x_0, and $v(x)=0$, otherwise. If the kinetic energy be neglected, then $E(1,v)=-3\lambda$, $E(2,v)=-4\lambda$, and $E(3,v)=-6\lambda$. Convexity does not hold. This can be turned into a proper example by letting λ be sufficiently large so that the kinetic energy can effectively be neglected; it is also possible to replace the hard core by a soft core.

Remark. The foregoing example is not applicable if θ is replaced by $|x|^{-1}$, thereby keeping alive the hope that convexity holds in the Coulomb case. The reason is the following: Given any four points x_0, y_1, y_2, y_3, let

$$|x_0-y_1|=\max_i\{|x_0-y_i|\}.$$

Then

$$|x_0-y_1|^{-1}\leq|y_1-y_2|^{-1}+|y_1-y_3|^{-1}.$$

The proof of this is left as an exercise, as well as the implication that if the kinetic energy is neglected, then convexity holds in the Coulomb case.

Question 7. For the case of Coulomb repulsion, is $F(N,\rho)$ jointly convex in N and ρ?

B. Density Matrices

Another possible modification of the theory of Sect. 3 is to replace densities $\rho(x)$ by single-particle admissible density matrices $\gamma(x,x')$. (See Questions 3 and 4 in Sec. 2. We do not restrict ourselves to γ's that come from pure states $\psi\rangle\langle\psi$.) This set of γ's is convex, and $\hat{F}(\gamma)$, defined analogously to (3.14), is convex [see the proof of Theorem 4.1(b)].

Despite the attractive feature just mentioned, there are three drawbacks to the approach:

(i) The problems about continuous tangent functionals remain and may even be more complex than before.

(ii) The original aim of the theory was to express the energy in terms of $\rho(x)$ and not $\gamma(x,x')$.

(iii) While the set of admissible γ's is well defined, it is not easy to identify. Given some γ, it is easy to verify that $\mathrm{Tr}\,\gamma=N$, but it is difficult to verify that $0\leq\gamma\leq qI$.

Still another possible modification is to retain ρ but to consider all N-particle density matrices Γ instead of merely pure states $\psi\rangle\langle\psi$. In other words, consider $\Gamma\mapsto\rho$ instead of $\psi\mapsto\rho$ and define

$$F_{\mathrm{DM}}(\rho)=\inf\{\mathrm{Tr}\,H_0\Gamma\,|\,\Gamma\mapsto\rho\}\tag{4.3}$$

on $\mathcal{I}_N$ and $F_{\mathrm{DM}}(\rho)=+\infty$ otherwise. Because $\Gamma\mapsto\rho$ is linear, F_{DM} is convex on $\mathcal{I}_N$. (Note: The example in Theorem 3.4 does not yield nonconvexity of F_{DM}.)

Obviously, the analog of (3.15) holds, namely,

$$E(v) = \inf\left\{F_{\mathrm{DM}}(\rho) + \int \rho v \,|\, \rho \in \mathscr{I}_N\right\}. \tag{4.4}$$

Since F_{DM} is convex, (4.4) can be used directly instead of (3.20) or (3.21).

Both F and F_{DM} are convex. The amusing fact is that

$$F(\rho) = F_{\mathrm{DM}}(\rho), \qquad \rho \in \mathscr{I}_N. \tag{4.5}$$

Equation (4.5) is not at all obvious, but it does say that the modification does not change the theory in any way. Equation (4.5) also yields another characterization of F. Equation (4.5) is proved in Theorem 4.3.

First, Γ is admissible if and only if

$$\Gamma(\mathbf{z}, \mathbf{z}') = \sum_{i=1}^{\infty} \lambda_i \psi_i(\mathbf{z}) \psi_i(\mathbf{z}')^*$$

with $0 \le \lambda_i$, $\Sigma\lambda_j = 1$, and the ψ_i are orthonormal. If $\psi_i \mapsto \rho_i$, then

$$\operatorname{Tr} H_0\Gamma = \Sigma\lambda_i(\psi_i, H_0\psi_i).$$

Thus we conclude that for all $\rho \in \mathscr{I}_N$

$$F_{\mathrm{DM}}(\rho) = \inf\left\{\sum_{i=1}^{\infty} \lambda_i \tilde{F}(\rho_i) \,\Big|\, \sum \lambda_i\rho_i = \rho,\, \rho_i \in \mathscr{I}_N,\, \lambda_i \ge 0, \sum \lambda_i = 1\right\}. \tag{4.6}$$

A simpler expression (which has to be proved) is

$$F_{\mathrm{DM}}(\rho) = \inf\left\{\sum \lambda_i \tilde{F}(\rho_i) \,\Big|\, \sum \lambda_i\rho_i = \rho,\, \rho_i \in \mathscr{I}_N,\, \lambda_i \ge 0, \sum \lambda_i = 1\right\}, \tag{4.7}$$

where the sums in (4.7) are restricted to finite sums. In view of (4.5), (4.7) is an alternative characterization of $F(\rho)$ for $\rho \in \mathscr{I}_N$.

Theorem 4.2. *Equation* (4.7) *is true.*

Proof. Pick $\varepsilon > 0$. Using (4.6), let $\{\lambda_i, \rho_i\}$ be an infinite sequence satisfying $\Sigma\lambda_i\rho_i = \rho, \rho_i \in \mathscr{I}_N$, and $F_{\mathrm{DM}}(\rho) \ge \Sigma\lambda_i\tilde{F}(\rho_i) - \varepsilon$. Since $\Sigma\lambda_i = 1$ and $\Sigma\lambda_i\tilde{F}(\rho_i) < \infty$, there exists K such that $A \equiv \sum_{i=K}^{\infty}\lambda_i \le \varepsilon$ and $B \equiv \sum_{i=K}^{\infty}\lambda_i F(\rho_i) \le \varepsilon$. Assume $A > 0$ for otherwise we are done. By Theorem 1.1 and the convexity of $G(\rho) = \int (\nabla\rho^{1/2})^2$,

$$\varepsilon \ge \sum_K^{\infty} \lambda_i \tilde{F}(\rho_i) \ge \sum_K^{\infty} \lambda_i G(\rho_i) \ge AG(\rho^K)$$

with $\rho^K = \sum_K^{\infty} \lambda_i\rho_i/A \in \mathscr{I}_N$. By Theorem 3.9 and the remark following it,

$$\tilde{F}(\rho^K) \le C[N^2 G(\rho^K) + N].$$

Therefore the finite sequence $\{\lambda_i, \rho_i\}_{i=1}^K$ with $\{\lambda_K, \rho_K\} = \{A, \rho^K\}$ satisfies $\Sigma\lambda_i\tilde{F}(\rho_i) \le F_{\mathrm{DM}}(\rho) + \varepsilon CN(N+1) + \varepsilon$. ■

Theorem 4.3. *Equation* (4.5) *is true.*

Proof. The easy part is that for $\rho \in \mathscr{I}_N$, $F_{\mathrm{DM}}(\rho) \geq F(\rho)$. By (4.4). $E(v) \leq F_{\mathrm{DM}}(\rho) + \int \rho v$ for all v. Hence, by (3.17), $F_{\mathrm{DM}}(\rho) \geq F(\rho)$. The hard part is contained in Corollary 4.5, which will be assumed for now. Then: (i) $F_{\mathrm{DM}}(\rho) \leq \tilde{F}(\rho)$ by (4.6); (ii) F_{DM} is convex and l.s.c. Hence $F_{\mathrm{DM}}(\rho) \leq \mathrm{CE}\, \tilde{F}(\rho) = F(\rho)$ by Theorem 3.7. ■

Theorem 4.4. *Suppose* $\{\rho_n\}$ *and* $\rho \in \mathscr{I}_N$ *and* $\rho_n \rightharpoonup \rho$ *weakly in* L^1. *Then there exists a density matrix* Γ, *with* $\Gamma \mapsto \rho$, *such that* $\operatorname{Tr} H_0\Gamma \leq \liminf F_{\mathrm{DM}}(\rho_n)$.

The proof of Theorem 4.4, due to Barry Simon, is given in the appendix.

Corollary 4.5. (i) F_{DM} *is* (*norm and weakly*) *l.s.c.*

(ii) *If* $\rho \in \mathscr{I}_N$, *there exists a density matrix* Γ *with* $\Gamma \mapsto \rho$ *such that* $\operatorname{Tr} H_0\Gamma = F_{\mathrm{DM}}(\rho)$ (*see Theorem* 3.3).

Proof. (i) If $\rho_n \to \rho$, $F_{\mathrm{DM}}(\rho) \leq \operatorname{Tr} H_0\Gamma \leq \liminf F_{\mathrm{DM}}(\rho_n)$. Norm l.s.c. implies weak l.s.c.

(ii) Take $\rho_n = \rho$ in Theorem 4.4. ■

C. The Kinetic Energy Functional

Kohn and Sham (KS) [30] define a kinetic energy functional $T_{\mathrm{KS}}(\rho)$. There are several other possible kinetic energy functionals and we shall explore their interrelations, as well as the fact that T_{KS} does *not* have a property assumed by KS. KS define the *exchange and correlation functional* $E_{\mathrm{xc}}(\rho)$ by

$$F_{\mathrm{HK}}(\rho) \equiv \tfrac{1}{2} \iint \rho(x)\rho(y)|x-y|^{-1}\, dx\, dy + T_{\mathrm{KS}}(\rho) + E_{\mathrm{xc}}(\rho). \qquad (4.8)$$

F_{HK} and T_{KS} are defined on different subsets of $\mathscr{I}_N$, so E_{xc} is defined only on a third unknown subset of $\mathscr{I}_N$. This difficulty can be remedied by using $\tilde{F}$ and $\tilde{T}$ in (4.8), but there is another point that should be stressed: There is no reason to believe that E_{xc} is convex on $\mathscr{I}_N$.

First, let us give some definitions. These use K instead of H_0 but otherwise are self-explanatory (with the aid of the equation numbers on the left):

$$\begin{array}{llll} (3.5): & E'(v) & \text{on} & L^{3/2}+L^{\infty}; \\ (3.10): & T_{\mathrm{KS}}(\rho) & \text{on} & \mathscr{A}'_N \quad (3.11); \\ (3.14): & \tilde{T}(\rho) & \text{on} & \mathscr{I}_N; \\ (3.17): & T(\rho) & \text{on} & L^3 \cap L^1. \end{array} \qquad (4.9)$$

($T(\psi) = (\psi, K\psi)$ was defined in (1.3) but it is quite different from $T(\rho)$ above. It is hoped that this notational lapse will not be confusing.) All the previous theorems [except for 3.9, wherein the last term in (3.22) should be omitted] carry over to these quantities. The primes on $E'(v)$ and $\mathscr{A}'_N$ indicate that these are different from before. Since Theorem 3.4 still holds, $\mathscr{A}'_N$ is not $\mathscr{I}_N$. It is left as an exercise to show that $\mathscr{A}_N \neq \mathscr{A}'_N$.

Question 8. What is $\mathscr{A}_N \cap \mathscr{A}'_N$?

There is one more kinetic energy functional that can be defined on $\mathcal{I}_N$, namely,

$$T_{\text{det}}(\rho) = \inf\{(\psi, K\psi) | \psi \mapsto \rho, \psi \in \mathcal{W}_N, \psi \text{ is a determinant}\}. \tag{4.10}$$

Clearly, $T_{\text{det}}(\rho) \geq \tilde{T}(\rho)$. The question to be addressed is whether $T_{\text{det}} = \tilde{T}$. The answer is No!, *not even on all of* $\mathcal{A}'_N$. KS assumed implicitly that $T_{\text{KS}}(\rho) = T_{\text{det}}(\rho)$ for $\rho \in \mathcal{A}'_N$; any such ρ minimizes $K + V$, but it is not true that such a $\rho(x)$ can always be written as $\sum_{i=1}^{N} |\psi_i(x)|^2$ with the ψ_i being orthonormal functions on $\mathbb{R}^3$. (Spin is a complication that is ignored at this point for simplicity.) In other words, not every ground state of $K + V$ is a determinant when degeneracy is present. I thank B. Simon for drawing my attention to this subtlety and for the construction in Theorem 4.8, which is reminiscent of the construction in Theorem 3.4.

Of course, $T_{\text{KS}} = \tilde{T}$ on $\mathcal{A}'_N$ by definition. Also $\tilde{T} = T$ on $\mathcal{A}'_N$ by Theorem 3.10. The following shows that there are cases in which $\tilde{T} = T_{\text{det}}$.

Theorem 4.6. *Suppose $\rho \in \mathcal{A}'_N$, so that $K + V$ has a ground state. If this ground state is nondegenerate, then $T_{\text{det}}(\rho) = \tilde{T}(\rho)$.*

Proof. The ψ that minimizes $(\psi, [K + V]\psi)$ is, of course, a determinant. ■

The following analog of Theorem 3.3 will be needed for Theorem 4.8.

Theorem 4.7. *Let $\rho \in \mathcal{I}_N$. Then there exists a determinant that minimizes $(\psi, K\psi)$ under the condition that $\psi \mapsto \rho$, $\psi \in \mathcal{W}_N$, and ψ is a determinant. Thus,* (4.10) *is actually a minimum.*

Proof. Let D_j be a sequence of determinants with

$$D_j \mapsto \rho \qquad \text{and} \qquad \lim\,(D_j, KD_j) = T_{\text{det}}(\rho).$$

The proof of Theorem 3.3 shows that ψ exists such that (i) $\psi \mapsto \rho$; (ii) $(\psi, K\psi) = T_{\text{det}}(\rho)$; (iii) $D_j \to \psi$ strongly in L^2. It suffices to show that ψ is a determinant. Let f_j^i, $i = 1, \ldots, N$, be the orthonormal single-particle functions of D_j. By the Banach–Alaoglu theorem, N functions $f^1, \ldots, f^N$ exist so that (after passing to a subsequence) $f_j^i \to f^i$ weakly. The f^i are not necessarily orthonormal. The function

$$P_j(z_1, \ldots, z_N) = \prod f_j^i(z_i)$$

then converges weakly to $P = \prod f^i$. This so because any $\psi \in L^2(\mathbb{R}^{3N})$ can be approximated in norm by sums of product functions. Therefore,

$$D_j \rightharpoonup (N!)^{1/2} \det[f^i(z_j)] \equiv D \qquad \text{weakly.}$$

But $D_j \to \psi$, so $D = \psi$. ■

Theorem 4.8. *Let $N = 7$ and $q = 1$. Then there is a $\rho \in \mathcal{A}'_N$ such that $T_{\text{det}}(\rho) > \tilde{T}(\rho)$.*

Proof. Take $v(x) = |x|^{-1}$, the hydrogen potential. The eigenvalues of $-\Delta + v$ are $-1/4$ (onefold), $-1/16$ (fourfold), $-1/36$ (ninefold). All other eigenvalues are greater than $-1/36$. The ground state for $N = 7$ and $q = 1$ is $\binom{9}{2} = 36$-fold degenerate, and a basis for this eigenspace consists of the determinants $(7!)^{-1/2}\det(1S, 2S, 2P_1, 2P_2, 2P_3, f, g)$ where f and g are any orthonormal functions in the nine-dimensional space M spanned by $S, P_1, P_2, P_3, D_1, \ldots, D_5$ (an orthonormal set for the $3S$, $3P$, and $3D$ waves). Let $d(f, g)$ denote the above normalized determinant and let

$$3^{1/2}\psi = d(S, D_1) + d(D_2, D_3) + d(D_4, D_5).$$

Then $\psi \mapsto \rho$ with $\rho = \rho_a + \rho_b$ and

$$\rho_a(x) = |1S(x)|^2 + |2S(x)|^2 + \sum_{i=1}^{3} |2P_i(x)|^2,$$

$$3\rho_b(x) = |S(x)|^2 + \sum_{i=1}^{5} |D_i(x)|^2.$$

Clearly $\rho \in \mathscr{A}'_N$ since ψ is a ground state.

If $T_{\text{det}}(\rho) = \tilde{T}(\rho)$, then there exists a determinant ϕ with $\phi \mapsto \rho$ and such that ϕ *must be* a ground state. Therefore $\phi = d(f, g)$ for some orthonormal $f, g \in M$. Thus,

$$|f(x)|^2 + |g(x)|^2 = \rho_b(x). \tag{4.11}$$

I claim that this is impossible. Write $f = A + D$ and $g = B + d$, with A and B being linear combinations of S and the P_i while D and d are linear combinations of the D_i. Now the S, P, and D waves behave as $|x|^0$, $|x|^1$, $|x|^2$, respectively, near the origin. By examining the behavior of (4.11) near the origin we conclude that

$$|D(x)|^2 + |d(x)|^2 = \tfrac{1}{3}\sum_{1}^{5} |D_i(x)|^2.$$

Since all the D_i waves have the same radial wave functions, this is really an equality about spherical harmonics Y_{2m}. The right side of the last equality is spherically symmetric, so the problem is to find two linear combinations F and G of the Y_{2m} such that

$$|F(\Omega)|^2 + |G(\Omega)|^2 = \text{constant} > 0.$$

This is impossible, and the proof is left as an exercise. (It is easily carried out if the following five basis functions are used: xyr^{-2}, xzr^{-2}, yzr^{-2}, $3x^2r^{-2} - 1$, $3y^2r^{-2} - 1$, with $r^2 = x^2 + y^2 + z^2$.) ■

Remarks. (i) $N = 7$ is not special; it was chosen for convenience in the proof.

(ii) An alternative way of viewing Theorem 4.8 is following. Suppose $K + V$ has a degenerate ground state, so that the ground eigenspace G is more than one-dimensional. $\psi \in G$ is a linear combination of determinants. Consider a perturbation w of v, namely, $v \to v + \lambda w$. In first-order perturbation theory, $V + \lambda W$ picks out a subspace g of G as the new ground eigenspace. If g is one dimensional, then g consists of one determinant since the ground eigenspace of $V + \lambda W$ always contains determinants (see Theorem 4.6). Now we ask, if $\psi_0 \in G$

and $\psi_0 \mapsto \rho_0$, can w be chosen so that g is one dimensional and $g = \{\psi_0\}$? Alternatively, can w be chosen so that $\min\{\int w\rho | \psi \mapsto \rho \text{ and } \psi \in G\}$ occurs uniquely for $\rho = \rho_0$? If so, ψ_0 is a determinant. Theorem 4.8 says that there can be a ρ_0 such that no w can pick it out uniquely.

Even though $T_{\det}(\rho) > \tilde{T}(\rho)$ for some ρ, $T_{\det}$ still satisfies the variational principle for $E(v)$.

Theorem 4.9. *For all* $v \in L^{3/2} + L^\infty$

$$E'(v) = \inf\left\{T_{\det}(\rho) + \int \rho v | \rho \in \mathcal{I}_N\right\}. \tag{4.12}$$

Proof. Equation (4.12) is equivalent to the following:

$$\begin{aligned} E'(v) &\equiv \inf\{(\psi, [K+V]\psi) | \psi \in \mathcal{W}_N\} \\ &= \inf\{(\psi, [K+V]\psi) | \psi \in \mathcal{W}_N, \psi \text{ is a determinant}\} \\ &\equiv \tilde{E}(v). \end{aligned}$$

Clearly $E'(v) \le \tilde{E}(v)$. Consider the operator $-\Delta + v(x)$. We *define* its "eigenvalues" $e_1 \le e_2, \ldots$ (here, spin degeneracy is included) by the min-max principle:

$$e_{n+1} = \sup\{\varepsilon_n(\phi_1, \ldots, \phi_n)\},$$

where

$$\begin{aligned} \varepsilon_n(\phi_1, \ldots, \phi_n) = \inf\{(\phi, [-\Delta + v]\phi) | \phi \in H^1, \|\phi\|_2 = 1 \\ \text{and } \phi \text{ is orthogonal to } \phi_1, \ldots, \phi_n\}. \end{aligned}$$

From this definition, it follows by a standard argument that

$$E_N(v) \equiv \sum_{i=1}^N e_i = \inf\left\{\sum_{i=1}^N (\phi_i, [-\Delta + v]\phi_i) | \phi_1, \ldots, \phi_N \text{ are orthonormal}\right\}. \tag{4.13}$$

But this least infimum equals

$$\inf\left\{\sum_{i=1}^\infty \lambda_i(\phi_i, [-\Delta + v]\phi_i) | \phi_1, \phi_2, \ldots, \text{are orthonormal}, \right.$$
$$\left. 0 \le \lambda_i \le 1 \text{ and } \sum_{i=1}^\infty \lambda_i = N\right\}.$$

This is easy to verify.

Let $\psi \in \mathcal{W}_N$ and let $\gamma = \Sigma\lambda_i f_i\rangle\langle f_i$ be its one-particle density matrix (including spin and with the f^i-orthonormal. $0 \le \lambda_i \le 1$, $\Sigma\lambda_i = N$). Then

$$(\psi, [K+V]\psi) = \Sigma\lambda_i(f_i, [-\Delta + v]f_i).$$

Thus $E'(v) \ge E_N(v)$. But $E_N(v) = \tilde{E}(v)$ by inspection. ■

Remark. This proof gives a formula for $E'(v)$, namely, $E_N(v)$.

The situation is complicated, so let us summarize it. T_{KS} is defined only on $\mathscr{A}'_N$, the set of ρ's that come from ground states for some v. $\mathscr{A}'_N$ has a smaller subset, $\mathscr{A}''_N$, in which ρ comes from a determinantal ground state, $\mathscr{A}''_N$ includes. but is larger than, $\mathscr{A}'''_N$, the set of ρ's that come from nondegenerate ground states. (Note: By Theorem 3.2 any ρ comes from a unique v (up to constants). Thus, if ρ comes from a determinant in a degenerate ground eigenspace, then $\rho \notin \mathscr{A}'''_N$.) On $\mathscr{A}''_N$ we have

$$T_{\mathrm{KS}}(\rho) = T_{\mathrm{det}}(\rho) = \tilde{T}(\rho).$$

Elsewhere on $\mathscr{A}'_N$,

$$T_{\mathrm{det}}(\rho) > T_{\mathrm{KS}}(\rho) = \tilde{T}(\rho).$$

Thus, there are two choices for (4.8): either T_{KS} or T_{det}. On $\mathscr{B}_N$, the complement of $\mathscr{A}'_N$, T_{KS} is not defined (but $\tilde{T}$, T_{det}, and T are defined). The preferred functional here is $T(\rho)$ because it is convex and hence most manageable. On $\mathscr{B}_N$, $T(\rho)$ is probably strictly less than $\tilde{T}(\rho) \le T_{\mathrm{det}}(\rho)$—at least this is so when T has a continuous tangent functional (Theorem 3.10). Such points are dense (Theorem 3.14). In any case, since T, $\tilde{T}$, and T_{det} can be interchangeably used in (4.12); it makes no difference which is used as far as $E'(v)$ is concerned.

5. Some Density Functionals That Are Bounds

In this section we forego the abstract functional theory of the previous sections and instead expound a different philosophy. Rather than pursuing "the correct density functional," which seems to be uncomputable, we shall content ourselves here with finding upper and lower bounds to the various quantities of interest in terms of $\rho(x)$. This latter program can provide rigorous bounds on ground state energies that, while they may not always be extremely accurate, do have a proper place in our conceptual scheme. Some of these bounds will be briefly displayed here; the interested reader is referred to the original papers for proofs.

It should be remembered that if one has bounds on two quantities (e.g., T and I; see below) and even if these bounds are optimal, then, in general, the sum of the bounds is *not* optimal for the sum of the two quantities (e.g., $T+I$).

A. *Kinetic Energy Lower Bound*

Lieb and Thirring (LT) [21] (also see Ref. 16) proved (for fermions in three dimensions) that if $\psi \mapsto \rho$, then (for all N)

$$T(\psi) \ge K^c (4\pi)^{-2/3} q^{-2/3} \int \rho(x)^{5/3}\, dx, \tag{5.1}$$

where K^c is the "classical" value $(3/5)(6\pi^2)^{2/3}$. LT conjectured that (5.1) holds in three dimensions with the $(4\pi)^{-2/3}$ deleted. [Note: Although an analog of (5.1) holds in all dimensions, the corresponding constant is definitely less than K^c in one and two dimensions.] In Ref. 22 (also see Ref. 16) $(4\pi)^{-2/3}$ was replaced by $1.496(4\pi)^{-2/3}$.

Incidentally, the statement $T(\psi) \geq Kq^{-2/3} \int \rho^{5/3}$ for all ψ, all N, and some K is equivalent to the following [21]. Let v be any nonpositive potential in $L^{5/2}(\mathbb{R}^3)$ and let $e_1 \leq e_2 \leq \cdots$, be the negative eigenvalues (if any) of $-\Delta + v(x)$ counting degeneracy, but not counting the q-fold degeneracy. Then

$$\sum e_i \geq -L \int |v(x)|^{5/2}\, dx$$

with $K = (3/5)(2/5L)^{2/3}$.

B. Kinetic Energy Upper Bound

There is, of course, no upper bound for $T(\psi)$ in terms of ρ. March and Young (MY) [17] proposed that for all $\rho \in \mathscr{I}_N$ there is a *determinantal* ψ, with $\psi \mapsto \rho$, such that

$$T(\psi) \leq q^{-2/n} K^{c} \int \rho(x)^{(n+2)/n}\, dx + \int [\nabla \rho^{1/2}(x)]^2\, dx, \tag{5.2}$$

where n is the dimension and $K^c = \pi^2/3$ for $n = 1$. (Compare (5.2) with Theorem 1.2.) They proved (5.2) for $n = 1$, but their proof for $n > 1$ has an error. Equation (5.2) for $n > 1$ is still an *open problem*. The MY construction for $n = 1$ motivated the construction in the proof of Theorem 1.2.

C. Lower Bound for the Indirect Part of the Coulomb Repulsion

Let Γ be a density matrix (which may be a pure state, $\Gamma = \psi\rangle\langle\psi)$ with $\Gamma \mapsto \rho$. Let

$$I(\Gamma) = \operatorname{Tr}\left\{\Gamma \sum_{1 \leq i < j \leq N} |x_i - x_j|^{-1}\right\} \tag{5.3}$$

be the Coulomb repulsive energy. The *indirect part* of this energy, $E(\Gamma)$, is *defined* by

$$I(\Gamma) = D(\rho) + E(\Gamma), \tag{5.4}$$

with

$$D(\rho) = \tfrac{1}{2} \iint \rho(x)\rho(y)|x-y|^{-1}\, dx\, dy \tag{5.5}$$

being the *direct part*.

In Ref. 23 it was shown that

$$E(\Gamma) \geq -C \int \rho(x)^{4/3}\, dx \tag{5.6}$$

with $C = 8.5$. In Ref. 24 this was improved to $C = 1.68$. The sharp (i.e., best) C in (5.6) is not known, but it is larger than 1.23.

It is well known that in any pure, *determinantal* state, $E(\Gamma)<0$. For other states, $E(\Gamma)$ can be positive. Indeed, for any fixed ρ there is no upper bound for $E(\Gamma)$ (see Ref. 24).

There is no q-dependence in (5.6) and, indeed, (5.6) holds for all statistics (i.e., C does not depend on statistics). This is explained in Ref. 24. The Dirac approximation has $C_D q^{-1/3}$ in (5.6) with $C_D = 3(6/\pi)^{1/3}/4 = 0.93$, but this q dependence is an artifact of the particular q-dependent determinantal ψ used to evaluate E from (5.4).

It should be noted that the bound

$$I(\Gamma) \geq D(\rho) - C \int \rho^{4/3} \tag{5.7}$$

is *not convex* in ρ. It is not even positive. These two faults lead to absurd conclusions when the right side of (5.7) is used in Thomas–Fermi–Dirac theory (see Ref. 25).

Since $\Gamma \mapsto \rho$ is linear,

$$\tilde{I}(\rho) = \inf\{I(\Gamma) | \Gamma \mapsto \rho\} \tag{5.8}$$

is convex in ρ. In other words, an optimal positive, convex lower bound must exist. Any reader who is devoted to abstract density functional theory, in the spirit of Sec. 3 or (5.8), should try to guess a plausible form for $\tilde{I}(\rho)$. (Proving it is another matter.) It will quickly be seen that $\tilde{I}(\rho)$ must be extremely complicated, and to say that it is "nonlocal" is an understatement. To see this, consider $N=2$ and ρ consisting of two "bumps," ρ_1 and ρ_2, very far apart. As long as $\int \rho_1 = \int \rho_2 = 1$, $\tilde{I}(\rho) - D(\rho) \approx 0$, independently of ρ_1 and ρ_2. But when $\int \rho_1 > 1$, $\int \rho_2 < 1$, then $\tilde{I}(\rho) - D(\rho)$ depends heavily on ρ_1 but not on ρ_2. The reason is that in the former case the two electrons can be far apart in the two bumps; in the latter case the two electrons must partly be close together in the first bump.

A problem that is physically more relevant and that illustrates the hidden complexity of density functional theory is the following problem about induced dipolar (or Van der Waals) forces raised in Ref. 25. When two atoms are a distance R apart, and R is large, there is an attraction $-R^{-6}$ (neglecting retardation effects). This attraction comes from the Coulomb repulsion, but it is not a static effect. The atomic dipole moment is almost zero. (There are, in fact, tiny dipole moments, but these are opposite in sign by symmetry, and hence repulsive. They must exist by the Feynmann–Hellman theorem: dE/dR = electric potential at the nucleus. I thank C. Herring for this remark.) There is almost no static dipole moment because to create one would cost a polarization energy αd^2. The attractive energy is $-d^2R^{-3}$ and, if $R^{-3}<\alpha$, $d=0$ for minimum energy. The cause of the $-R^{-6}$ energy is more subtle, but it has a semiclassical basis: The electrons in each atom *move in phase* while maintaining the spherical symmetry about each atom. The energy cost is then αd^4 and the minimum energy occurs when $2\alpha d^2 = R^{-3}$. Thus, the $-R^{-6}$ attraction comes from the fact that the electron cloud cannot be thought of as a simple "fluid." This effect is somehow built into $\tilde{I}(\rho)$, but an explicit form of $\tilde{I}(\rho)$ that will produce this effect has yet to be displayed.

D. A Variational Principle

$E(v)$, given by (3.5), satisfies (by definition) the well-known variational principle

$$E(v) \leq (\psi, H_v\psi). \tag{5.9}$$

Can an upper bound for $E(v)$ be given in terms of ρ alone? If (5.2) were true, then, for any $\rho \in \mathscr{I}_N$,

$$E(v) \leq \text{right side of (5.2)} + D(\rho) + \int v\rho. \tag{5.10}$$

[See the remark about $E(\Gamma)$ for determinants in Sec. 5C.]

An upper bound for $E(v)$ can, indeed, be given in terms of the one-particle density *matrices* $\gamma(z, z')$ as follows [26]:

Let γ be any admissible one-particle density matrix ($0 \leq \gamma \leq I$, $\operatorname{Tr} \gamma = N$). (Note: γ includes spin. It was called $\bar{\gamma}$ in Sec. 2). Then

$$E(v) \leq \operatorname{Tr} \gamma(-\Delta + v(x)) + \tfrac{1}{2} \int K_2(z, z')|x - x'|^{-1}\, dz\, dz', \tag{5.11}$$

where $\int dz = \sum_\sigma \int dx$ and

$$K_2(z, z') = \gamma(z, z)\gamma(z', z') - |\gamma(z, z')|^2. \tag{5.12}$$

The form (5.11) is well known *if* γ came from a pure state $\psi\rangle\langle\psi$ with ψ being a determinant. The point about (5.11) is that it holds for *all* admissible γ. Incidentally, the minimum of (5.11) over all admissible γ occurs when γ comes from a determinantal ψ. In other words, *the best* Hartree–Fock function minimizes (5.11), but (5.11) is interesting precisely because this HF function is unknown.

E. Thomas–Fermi Theory

This theory (see Ref. 25 for an exposition) does not yield bounds and therefore does not properly belong here. However, it illustrates the usefulness of the bounds in Secs. 5A–C.

The TF functional is

$$\mathscr{E}^{\mathrm{TF}}(\rho) = K^{\mathrm{c}} q^{-2/3} \int \rho^{5/3} + D(\rho) + \int v\rho, \tag{5.13}$$

while the TF Weizsaecker functional $\mathscr{E}^{\mathrm{TFW}}(\rho)$ is the right-hand side of (5.10). If $-C \int \rho^{4/3}$ is added to the right-hand side of (5.13), the result is TF Dirac theory.

The TF energy for N particles is defined by

$$E^{\mathrm{TF}} = \inf\left\{\mathscr{E}^{\mathrm{TF}}(\rho) \,\Big|\, \int \rho = N\right\}, \tag{5.14}$$

and similarly for E^{TFW} and E^{TFD}.

Now suppose that v is an atomic or molecular potential, that is,

$$v(x) = -\sum_{j=1}^{k} z_j |x - R_j|^{-1}, \tag{5.15}$$

with the $z_j > 0$. It is a fact [14] that under the scaling $z_j \to \lambda z_j$ and $N \to \lambda N$, as $\lambda \to \infty$

$$E^{\mathrm{TF}}/E(v) \to 1, \tag{5.16}$$

where $E(v)$ is the true ground state energy. Note that (5.16) also holds if E^{TF} is replaced by E^{TFW} or E^{TFD} (see Ref. 25).

Thus we see that if the conjecture in Sec. 5A holds, then, combining (5.1) with (5.7), TFD theory is a lower bound that is asymptotically exact. Similarly, if (5.2) holds, then, as remarked in Sec. 5C TFW theory is an upper bound that is asymptotically exact.

F. Two-Body Density Matrices

If one is willing to go beyond the one-body density ρ or one-body density matrix γ and consider the two-body reduced density matrix $\gamma^{(2)}$, then $E(v)$ is directly and exactly expressible in terms of $\gamma^{(2)}$, since H_v has only one- and two-body terms. The problem is that it is very difficult to decide when a given $\gamma^{(2)}$ is, in fact, the reduction of an admissible N-body density matrix Γ. This is called the *N-representability problem* and it has not been solved. (This is to be compared with the fact that there is a simple necessary and sufficient condition for a one-body γ to be N-representable; see Sec. 2).

It is possible, however, to find some necessary conditions and some sufficient conditions for $\gamma^{(2)}$ to be N-representable. Using these, bounds on $E(v)$ can be derived. Since this approach is outside the scope of this article, we refer the reader to the excellent review of Percus [27].

Appendix: Proofs of Theorems 1.3, 3.3, and 4.4

The following proof of Theorem 1.3 is due to H. Brezis (pivate communication.)

Proof. For simplicity of presentation we take $N = 2$ and $q = 1$ (no spin). Therefore we have $\psi(x, y)$ and

$$F(x) = \left(\int |\psi(x, y)|^2 \, dy\right)^{1/2} = [\rho(x)/2]^{1/2}.$$

Suppose that $\psi_n \to \psi$ in H^1 $(\mathbb{R}^3 \times \mathbb{R}^3)$; that is, $\psi_n \to \psi$ and $\nabla\psi_n \to \nabla\psi$ in L^2. We want to show that $F_n \to F$ in $L^2(\mathbb{R}^3)$ and $\nabla F_n \to \nabla F$ in $L^2(\mathbb{R}^3)$. The former is trivial:

By the Schwarz inequality,

$$\left(\int |\psi_n(x,y)|^2\,dy \int |\psi(x,y)|^2\,dy\right)^{1/2} \geq \tfrac{1}{2}\int \psi_n(x,y)^*\psi(x,y)\,dy + \tfrac{1}{2}\int \psi(x,y)^*\,\psi_n(x,y)\,dy.$$

Therefore

$$\int |F_n(x)-F(x)|^2\,dx \leq \int dx \int |\psi_n(x,y)-\psi(x,y)|^2\,dy.$$

and the right-hand side converges to zero.

The proof that $\nabla F_n \to \nabla F$ is the difficult one. It is sufficient to prove convergence for *some* subsequence n_j, $j=1,2,\ldots$. (If $\nabla F_n \not\to \nabla F$, then there is some subsequence and some $\varepsilon>0$ such that $\|\nabla F_{n_j}-\nabla F\|>\varepsilon$. But then this subsequence clearly does not have a subsequence which converges to ∇F.) Now since $\psi_n \to \psi$ in H^1 there is some subsequence and some function $G \in H^1$ such that

$$|\psi_{n_j}(x,y)| \leq G(x,y) \text{ and } |\nabla\psi_{n_j}(x,y)| \leq G(x,y),$$

a.e. in $\mathbb{R}^6$. (The proof of this fact is the same as the first half of the proof of the Riesz–Fischer lemma that L^1 is complete.) Henceforth, we shall replace n_j by n. We shall also assume, for simplicity, that $F_n(x)$ and $F(x)>0$ for all x (otherwise, an approximation argument can be used).

Now

$$2\nabla F_n(x) = A_n(x)/F_n(x) + \text{c.c.}$$

with

$$A_n(x) = \int B_n(x,y)\,dy \qquad \text{and} \qquad B_n(x,y) = \psi_n(x,y)^*\nabla\psi_n(x,y).$$

As we saw, $F_n \to F$ in L^2, so, by passing to a subsequence, we can assume $F_n(x) \to F(x)$ a.e. Furthermore, a.e.

$$|B_n(x,y)| \leq G(x,y)^2 \in L^1(\mathbb{R}^6).$$

By passing to a subsequence we can assume $\nabla\psi_n \to \nabla\psi$ and $\psi_n \to \psi$ a.e. Thus, by dominated convergence, $B_n \to B$ in $L^1(\mathbb{R}^6)$. For this subsequence $|B_n(x,y)-B(x,y)| \to 0$, a.e. ($\mathbb{R}^6$). Then, for a.e. x, $|B_n(x,y)-B(x,y)| \to 0$ a.e. y. Thus, by dominated convergence $\int dy\,|B_n(x,y)-B(x,y)| \to 0$, a.e. x. In other words, for some subsequence, $\nabla F_n(x) \to \nabla F(x)$, a.e.

Finally, we note that, by the Schwarz inequality,

$$|\nabla F_n(x)|^2 \leq \int |\nabla\psi_n(x,y)|^2\,dy \leq \int G(x,y)^2\,dy \equiv C(x)^2.$$

Since C is a fixed L^2 function, $\nabla F_n \to \nabla F$ in L^2 by dominated convergence. ∎

Proof of Theorem 3.3. Let ψ_j (with $\psi_j \mapsto \rho$) be a minimizing sequence for $\tilde{F}(\rho)$. The ψ_j are obviously bounded in $H^1(\mathbb{R}^{3N})$, so, by the Banach–Alaoglu theorem, there is a $\psi \in H^1(\mathbb{R}^{3N})$ such that $\psi_j \rightharpoonup \psi$ weakly in $H^1(\mathbb{R}^{3N})$. Obviously, ψ has the same symmetry as the ψ_j. It is well known that under weak limits positive quadratic forms decrease. Thus

$$\tilde{F}(\rho) = \lim\,(\psi_j, H_0\psi_j) \geq (\psi, H_0\psi).$$

If we can show that $\psi \mapsto \rho$, we are done. To do so it is sufficient to prove that $\psi_j \to \psi$ strongly because if $\psi \mapsto \tilde{\rho}$ we have, by the easy part of Theorem 1.3 that $\rho^{1/2} = \rho_j^{1/2} \to \tilde{\rho}^{1/2}$ in L^2, so that $\tilde{\rho} = \rho$.

Strong convergence will be proved by showing that $\int |\psi|^2 = 1$. Let S be the characteristic function of some bounded set in $\mathbb{R}^{3N}$. By the Rellich–Kondrachov theorem [28] there is a subsequence (which can be chosen independent of S) of the ψ_j such that $S\psi_j$ converges strongly (in L^2) to $S\psi$. Pick $\varepsilon > 0$ and let χ be the characteristic function of a bounded set in $\mathbb{R}^3$ such that

$$\varepsilon > \int \rho(1-\chi) \equiv \int |\psi_j|^2 \sum_j [1-\chi(x_i)].$$

But

$$\sum [1-\chi(x_i)] \geq 1 - S,$$

where $S = \Pi\chi(x_i)$. Thus, $\int |\psi_j|^2 S \geq 1-\varepsilon$. Since $|\psi_j|^2 S \to \int |\psi|^2 S$, we have that $\int |\psi|^2 \geq \int |\psi|^2 S \geq 1-\varepsilon$ for all $\varepsilon > 0$. ■

Remark. The symmetry of ψ was not needed in this proof provided one generalizes definition (1.6) to

$$\rho(x) = \sum_\sigma \sum_{i=1}^N \int |\psi(z_1, \ldots, (x, \sigma_i), \ldots, z_N)|^2\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_N. \quad \text{(A.1)}$$

The following proof of Theorem 4.4 is due to B. Simon (private communication). It is closely related to the proof of Theorem 3.3 just given.

Proof. Without loss, replace H_0 by $h^2 = H_0 + 1$ in the definitions. h^{-1} is a bounded operator. We can assume that $g_n \equiv F_{\mathrm{DM}}(\rho_n) < \infty$, $g \equiv \lim g_n$ exists, and $\operatorname{Tr} h\Gamma_n h = \operatorname{Tr} \Gamma_n h^2 \leq g_n + 1/n$ with $\Gamma_n \mapsto \rho_n$. Thus, $y_n \equiv h\Gamma_n h$ is uniformly bounded in the trace norm. The dual of the compact operators, com, is the trace class operators t, and $\gamma \in t$ takes $A \in$ com into $\operatorname{Tr} \gamma A$. A sequence $\gamma_n \in t$ converges to $\gamma \in t$, in the weak* topology, if and only if $\operatorname{Tr} \gamma_n A \to \operatorname{Tr} \gamma A$ for all $A \in$ com.

The Banach–Alaoglu theorem states that a norm-closed ball of finite radius in t is compact in the weak* topology. For us this means that there exists y with $\operatorname{Tr} y < \infty$, so that, for a subsequence, $\operatorname{Tr} y_n A \to \operatorname{Tr} yA$ for every compact A. Clearly, $y \geq 0$ and therefore $\underline{\lim} \operatorname{Tr} y_n \geq \operatorname{Tr} y$. Also, y obviously has the correct (Pauli) symmetry. If we can show that $\Gamma \equiv h^{-1} y h^{-1}$ (which is in trace class) satisfies $\Gamma \mapsto \rho$, we are done. To do this we shall show that if $\Gamma \mapsto \rho'$, then $\int (\rho_n - \rho')f \to 0$ for any $f \in L^\infty$. This would mean that $\rho_n \rightharpoonup \rho'$ weakly in L^1. But since $\rho_n' \rightharpoonup \rho$ in L^1, $\rho' = \rho$.

As in the proof of Theorem 3.3, for any $\varepsilon > 0$ there is a χ (=characteristic function of a bounded set in $\mathbb{R}^3$) such that

$$\int \rho(1-\chi) < \varepsilon \qquad \text{and} \qquad \int \rho'(1-\chi) < \varepsilon.$$

Since $\rho_n \rightharpoonup \rho$, $\int \rho_n(1-\chi) < \varepsilon$ for n sufficiently large. If

$$\phi_n(x_1, \ldots, x_N) = \Gamma_n(x_1, \ldots, x_N; x_1, \ldots, x_N)$$

(after summing on spins), and similarly for ϕ, we have (as in Theorem 3.3)

$$\int \phi_n(1-S) < \varepsilon \qquad \text{and} \qquad \int \phi(1-S) < \varepsilon$$

where $S = \Pi\chi(x_i)$. In view of this, it is sufficient to show that

$$\int_n P \to \int \phi P \qquad \text{with} \qquad P = S \sum_i f(x_i).$$

Let $P = P(x_1, \ldots, x_N)$ be any bounded functions of compact support and let M_P be the operator (in L^2) of multiplication by P. It is a fact that $A_P \equiv h^{-1}M_P h^{-1}$ is compact. (This is essentially the same as the Rellich–Kondrachov theorem used in Theorem 3.3.) Therefore

$$\operatorname{Tr} \Gamma_n M_P = \operatorname{Tr} y_n A_P \to \operatorname{Tr} y A_P = \operatorname{Tr} \Gamma M_P.$$

■

Acknowledgment

This work was partially supported by U.S. National Science Foundation grant No. PHY-7825390-A02. This paper is a revised version of a paper with the same title that appeared in *Physics as Natural Philosophy: Essays in Honor of Laszlo Tisza on his 75th Birthday*, H. Feshbach and A. Shimony, Eds. (M.I.T. Press, Cambridge, 1982), pp. 111–149.

Bibliography

[1] L. H. Thomas, Proc. Camb. Phil. Soc. **23**, 542 (1927).
[2] E. Fermi, Rend. Accad. Naz. Lincei **6**, 602 (1927).
[3] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964).
[4] M. M. Morell, R. G. Parr and M. Levy, J. Chem. Phys. **62**, 549 (1975).
[5] R. G. Parr, S. Gadre and L. J. Bartolotti, Proc. Natl. Acad. Sci. USA **76**, 2522 (1979).
[6] R. A. Donnelly and R. G. Parr, J. Chem. Phys. **69**, 4431 (1978).
[7] H. Englisch and R. Englisch, "Hohenberg–Kohn theorem and non-v-representable densities," Physica A, to be published.
[8] T. L. Gilbert, Phys. Rev. B **6**, 211 (1975).
[9] J. E. Harriman, Phys. Rev. A **6**, 680 (1981).

[10] M. Levy, Proc. Natl. Acad. Sci. USA **76**, 6062 (1979).
[11] M. Levy, Phys. Rev. A **26**, 1200 (1982).
[12] S. M. Valone, J. Chem. Phys. **73**, 1344 (1980); *ibid.* **73**, 4653 (1980).
[13] A. S. Bamzai and B. M. Deb, Rev. Mod. Phys. **53**, 95 (1981). *Erratum*, **53**, 593 (1981).
[14] E. H. Lieb and B. Simon, Adv. Math. **23**, 22 (1977). See also Thomas–Fermi theory revisited, Phys. Rev. Lett. **31**, 681 (1973). See also Refs. 16 and 25.
[15] M. Hoffmann-Ostenhof, and T. Hoffmann-Ostenhof, Phys. Rev. A **16**, 1782 (1977).
[16] E. H. Lieb, Rev. Mod. Phys. **48**, 553 (1976).
[17] N. H. March and W. H. Young, Proc. Phys. Soc. **72**, 182 (1958).
[18] M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, New York, 1978), Vol. 4.
[19] S. Mazur, Studia Math. **4**, 70 (1933).
[20] R. B. Israel, *Convexity in the Theory of Lattice Gases* (Princeton U.P., Princeton NJ, 1979).
[21] E. H. Lieb and W. E. Thirring, "Inequalities for the moments of the eigenvalues of the Schrödinger hamiltonian and their relation to Sobolev inequalities," in *Studies in Mathematical Physics*, E. H. Lieb, B. Simon, and A. S. Wightman, Eds. (Princeton U.P., Princeton, NJ, 1976). See also Phys. Rev. Lett. 687 (1975); *Errata*, **35**, 1116 (1975).
[22] E. H. Lieb, Am. Math. Soc. Proc. Symp. Pure Math. **36**, 241 (1980).
[23] E. H. Lieb, Phys. Lett. A **70**, 444 (1979).
[24] E. H. Lieb and S. Oxford, Int. J. Quantum Chem. **19**, 427 (1981).
[25] E. H. Lieb, Rev. Mod. Phys. **53**, 603 (1981); *Errata*, **54**, 311 (1982).
[26] E. H. Lieb, Phys. Rev. Lett. **46**, 457 (1981); *Erratum*. **47**, 69 (1981).
[27] J. K. Percus, Int. J. Quantum Chem. **13**, 89 (1978).
[28] R. A. Adams, *Sobolev Spaces* (Academic Press, New York, 1975).
[29] W. Fenchel, Can. J. Math. **1**, 23 (1949).
[30] W. Kohn and L. J. Sham, Phys. Rev. A **140** 1133 (1965).

Received October 19, 1982
Accepted for publication March 11, 1983