

# The stability of matter

Elliott H. Lieb\*

*Department of Mathematics and Department of Physics, Princeton University, Princeton, New Jersey 08540*

A fundamental paradox of classical physics is why matter, which is held together by Coulomb forces, does not collapse. The resolution is given here in three steps. First, the stability of *atom* is demonstrated, in the framework of nonrelativistic quantum mechanics. Next the Pauli principle, together with some facts about Thomas-Fermi theory, is shown, to account for the stability (i.e., saturation) of *bulk* matter. Thomas-Fermi theory is developed in some detail because, as is also pointed out, it is the asymptotically correct picture of heavy atoms and molecules (in the  $Z \rightarrow \infty$  limit). Finally, a rigorous version of screening is introduced to account for *thermodynamic* stability.

## CONTENTS

Introduction	553
I. The Stability of Atoms	554
II. Extension of the Uncertainty Principle to Many Fermions	556
III. Thomas-Fermi Theory	557
IV. The Stability of Bulk Matter	562
V. The Thermodynamic Limit	563
References	568

## INTRODUCTION

Some features of the physical world are so commonplace that they hardly seem to deserve comment. One of these is that ordinary matter, either in the form of atoms or in bulk, is held together with Coulomb forces and yet is stable. Nowadays we regard this truly remarkable phenomenon as a consequence of quantum mechanics, but it is far from obvious how the conclusion follows from the premise. It is not necessary to ponder the question very long before realizing that it is a subtle one and that the answer is not to be found in any textbook.

Although the Schrödinger equation is half a century old, it was only in the last few years that the proof of stability was completed. The aim of this paper is to present the full story in a simple and coherent way, highlighting only the main physical and mathematical ideas.

The sense of profound unease about the problem just before the dawn of quantum mechanics is exemplified by this quotation (Jeans, 1915):

"... there would be a very real difficulty in supposing that the law  $1/r^2$  held down to zero values of  $r$ . For the force between two charges at zero distance would be infinite; we should have charges of opposite sign continually rushing together and, when once together, no force would be adequate to separate them... Thus the matter in the universe would tend to shrink into nothing or to diminish indefinitely in size... We should however probably be wrong in regarding a molecule as a cluster of electrons and positive charges. A more likely suggestion, put forward by Larmor and others is that the molecule may consist, in part at least, of rings of electrons in rapid orbital motion."

Jeans' words strike a contemporary chord, especially since one aspect of the problem that worried him has not yet been fully resolved. This is that electrons and nuclei have a magnetic dipole-dipole interaction whose energy goes as  $r^{-3}$ . Although the angular average of this interaction vanishes, the interaction can cause the collapse that Jeans feared, even with Schrödinger mechanics. A proper quantum electrodynamics is needed to describe the dipolar interaction at very small distances. For that reason spin dependent forces will be ignored in this paper; only nonrelativistic quantum mechanics will be considered.

It is difficult to find a reliable textbook answer even to the question: How does quantum mechanics prevent the collapse of an atom? One possibility is to say that the Schrödinger equation for the hydrogen atom can be solved and the answer seen explicitly. This is hardly satisfactory for the many-electron atom or for the molecule. Another possible answer is the Heisenberg uncertainty principle. This, unfortunately, is a false argument, as shown in Sec. I. There is, however, a much better uncertainty principle, formulated by Sobolev, which does adequately describe the intuitive fact that a particle's kinetic energy increases sufficiently fast, as the wave function is compressed, to prevent collapse. (See Kato, 1951).

The next question to consider is well stated in this quotation from Ehrenfest (in Dyson, 1967):

"We take a piece of metal. Or a stone. When we think about it, we are astonished that this quantity of matter should occupy so large a volume. Admittedly, the molecules are packed tightly together, and likewise the atoms within each molecule. But why are the atoms themselves so big?... Answer: only the Pauli principle, 'No two electrons in the same state.' That is why atoms are so unnecessarily big, and why metal and stone are so bulky."

Dyson then goes on to say that without the Pauli principle

"We show that not only individual atoms but matter in bulk would collapse into a condensed high-density phase. The assembly of any two macroscopic objects would release energy comparable to that of an atomic bomb."

Two distinct facts are involved here. One is that matter is stable (or saturates), meaning that the ground

\*Work partially supported by U. S. National Science Foundation grant MCS 75-21684.

state energy is bounded below by a constant times the first power, and not a higher power, of the particle number. This was proved for the first time by Dyson and Lenard (Dyson and Lenard, 1967, and Lenard and Dyson, 1968), in a beautiful series of papers. Their method is quite complicated, however, and a simpler proof is given in Sec. IV. In addition, they used sufficiently many inequalities that their estimate (for hydrogen atoms) is about  $-10^{14}$  Ry/particle. We will obtain a bound of about  $-23$  Ry/particle. The second fact is that matter would definitely not be stable if electrons were bosons (Dyson, 1967). The energy would increase at least as fast as  $-N^{7/5}$ .

Therefore, Ehrenfest's surmise that the Pauli principle plays a crucial role in preventing collapse is correct. The problem is to display the essence of the Pauli principle in a clear, succinct and mathematically precise way. Unless this is done the physics of stability will remain unclear.

The key fact is developed in Sec. II: If  $\rho(x)$  is the one-particle density of any fermion wave function then the total kinetic energy is bounded below by (constant)  $\int \rho(x)^{5/3} dx$ .<sup>1</sup> This inequality may be termed the uncertainty principle for fermions. It is simple yet powerful enough to establish stability.

Given this bound, it is then necessary to show how the kinetic energy eventually overcomes the  $r^{-1}$  Coulomb singularity. It turns out that Thomas-Fermi (TF) theory is exactly what is needed for this purpose because, as Teller discovered in 1962, atoms do not bind in TF theory. Thus TF theory immediately implies saturation. The necessary facts about TF theory are developed in Sec. III.

There is also another good reason for understanding TF theory in detail. The theory used to be regarded as an uncertain approximation in atomic physics, but it is now known that it is more than that. It happens to be an asymptotically correct theory of atoms and molecules as the nuclear charges tend to infinity. In short, TF theory and the theory of the hydrogen atom constitute two opposite, but rigorous foundations for the many electron problem.

After putting together the results of Sec. II and III in Sec. IV, and thereby proving the stability of bulk matter, we address the third main topic of this paper in Sec. V. Does a sensible thermodynamic limit exist for matter? The problem here centers around the long range  $r^{-1}$  nature of the Coulomb potential, not the short range singularity. Put another way, the question is that if matter does not implode, how do we know that it does not explode? Normally systems with potentials that fall off less slowly than  $r^{-3-\epsilon}$  for some  $\epsilon > 0$  cannot be expected to have a thermodynamic limit. The crucial physical fact was discovered by Newton in 1687: outside an isotropic distribution of charge, all the charge appears to be concentrated at the center. This fact is the basis for screening, but to use it a geometric fact about the packing of balls will be needed. Quantum mechanics as such plays almost no role in Sec. V.

<sup>1</sup> $\int f(x)dx$ , or simply  $\int f$ , always denotes a three-dimensional integral.

The content of this paper can be summarized as follows:

- (i) Atoms are stable because of an uncertainty principle,
- (ii) Bulk matter is stable because of a stronger uncertainty principle that holds only for fermions;
- (iii) Thermodynamics exists because of screening.

My hope is that the necessary mathematics, which is presented as briefly as possible, will not obscure these simple physical ideas.

This paper is based on research carried out over the past few years, and it was my good fortune to have had the benefit of collaboration with J. L. Lebowitz, B. Simon, and W. E. Thirring. Without their insights and stimulation probably none of this would have been carried to fruition. Secs. II and IV come from Lieb and Thirring (1975), Sec. III from Lieb and Simon (1977), and Sec. V from Lieb and Lebowitz (1972).

Lectures given in 1976 at the Centro Internazionale Matematico Estivo in Bressanone were the impetus for writing this paper. The bibliography is not intended to be scholarly, but I believe no theorem or idea has been quoted without proper credit.

I am doubly grateful to S. B. Treiman. He kindly invited me to submit this paper to *Reviews of Modern Physics*, and he also generously devoted much time to reading the manuscript and made many valuable suggestions to improve its clarity.

## I. THE STABILITY OF ATOMS

By the phrase "stability of an atom" is meant that the ground state energy of an atom is finite. This is a weaker notion than the concept of  $H$  stability of matter, to be discussed in Sec. IV, which means that the ground state energy of a many-body system is not merely bounded below but is also bounded by a constant times the number of particles. This, in turn, is different from thermodynamic stability discussed in Sec. V.

Consider the Hamiltonian for the hydrogenic atom:

$$H = -\Delta - Z|x|^{-1} \quad (1)$$

(using units in which  $\hbar^2/2=1$ ,  $m=1$ , and  $|e|=1$ ).  $H$  acts on  $L^2(\mathbb{R}^3)$ , the square integrable functions on 3-space. Why is the ground state energy finite, i.e., why is

$$\langle \psi, H\psi \rangle \geq E_0 \langle \psi, \psi \rangle \quad (2)$$

for some  $E_0 > -\infty$ ? The obvious elementary quantum mechanics textbook answer is the *Heisenberg uncertainty principle* (Heisenberg, 1927): If the *kinetic energy* is defined by

$$T_\psi \equiv \int |\nabla \psi(x)|^2 dx, \quad (3)$$

and if

$$\langle |x|^2 \rangle_\psi \equiv \int |x|^2 |\psi(x)|^2 dx,$$

then when

$$\begin{aligned} \langle \psi, \psi \rangle &= \|\psi\|_2^2 = \int |\psi(x)|^2 dx = 1, \\ T_\psi \langle |x|^2 \rangle_\psi &\geq \frac{9}{4}. \end{aligned} \quad (4)$$

The intuition behind applying the Heisenberg uncertainty principle (4) to the ground state problem (2) is that if the electron tries to get within a distance  $R$  of the nucleus, the kinetic energy  $T_\psi$  is at least as large as  $R^{-2}$ . Consequently  $\langle \psi, H\psi \rangle \geq R^{-2} - Z/R$ , and this has a minimum  $-Z^2/4$  for  $R = 2/Z$ .

The above argument is *false*! The Heisenberg uncertainty principle says no such thing, despite the endless invocation of the argument. Consider a  $\psi$  consisting of two parts,  $\psi = \psi_1 + \psi_2$ .  $\psi_1$  is a narrow wave packet of radius  $R$  centered at the origin with  $\int |\psi_1|^2 = \frac{1}{2}$ .  $\psi_2$  is spherically symmetric and has support in a narrow shell of mean radius  $L$  and  $\int |\psi_2|^2 = \frac{1}{2}$ . If  $L$  is large then, roughly,  $\int |x|^2 |\psi(x)|^2 dx \sim L^2/2$ , whereas  $\int |x|^{-1} |\psi(x)|^2 dx \sim 1/2R$ . Thus, from (4) we can conclude *only* that  $T_\psi > 9/2L^2$  and hence that  $\langle \psi, H\psi \rangle \geq 9/2L^2 - Z/2R$ . With this wave function, and using *only* the Heisenberg uncertainty principle, we can make  $E_0$  arbitrarily negative by letting  $R \rightarrow 0$ .

A more colorful way to put the situation is this: an electron cannot have both a sharply defined position and momentum. If one is willing to place the electron in two widely separated packets, however, say here and on the moon, then the Heisenberg uncertainty principle *alone* does not preclude each packet from having a sharp position and momentum.

Thus, while Eq. (4) is correct it is a pale reflection of the power of the operator  $-\Delta$  to prevent collapse. A better uncertainty principle (i.e., a lower bound for the kinetic energy in terms of some integral of  $\psi$  which does not involve derivatives) is needed, one which reflects more accurately the fact that if one tries to compress a wave function *anywhere* then the kinetic energy will increase. This principle was provided by Sobolev (1938) and for some unknown reason his inequality, which is simple and goes directly to the heart of the matter, has not made its way into the quantum mechanics textbooks where it belongs. Sobolev's inequality in three dimensions [unlike (4) its form is dimension dependent] is

$$T_\psi = \int |\nabla \psi(x)|^2 dx \geq K_s \left\{ \int \rho_\psi(x) dx \right\}^{1/3} = K_s \|\rho_\psi\|_3, \quad (5)$$

where

$$\rho_\psi(x) = |\psi(x)|^2 \quad (6)$$

is the density and

$$K_s = 3(\pi/2)^{4/3} \approx 5.478$$

is known to be the best possible constant. Equation (5) is nonlinear in  $\rho$ , but that is unimportant.

A rigorous derivation of (5) would take too long to present but it can be made plausible as follows (Rosen, 1971):  $K_s$  is the minimum of

$$K^\psi = \frac{\int |\nabla \psi(x)|^2 dx}{\left\{ \int |\psi(x)|^6 dx \right\}^{1/3}}.$$

Let us accept that a minimizing  $\psi$  exists (this is the hard part) and that it satisfies the obvious variational equation

$$-(\Delta \psi)(x) - \alpha \psi(x)^5 = 0$$

with  $\alpha > 0$ . Assuming also that there is a minimizing  $\psi$

which is nonnegative and spherically symmetric (this can be proved by a rearrangement inequality), one finds by inspection that

$$\psi(x) = (3/\alpha)^{2/3} (1 + |x|^2)^{-1/2}.$$

When this is inserted into the expression for  $K^\psi$  the result is  $K_s = 3(\pi/2)^{4/3}$ . The minimizing  $\psi$  is not square integrable, but that is of no concern.

Now let us make a simple calculation to show how good (5) really is. For any  $\psi$

$$\langle \psi, H\psi \rangle \geq K_s \left( \int \rho_\psi(x) dx \right)^{1/3} - Z \int |x|^{-1} \rho_\psi(x) dx \equiv h(\rho), \quad (7)$$

and hence when  $\langle \psi, \psi \rangle = 1$

$$\langle \psi, H\psi \rangle \geq \min \left\{ h(\rho) : \rho(x) \geq 0, \int \rho = 1 \right\}. \quad (8)$$

The latter calculation is trivial (for any potential) since gradients are not involved. One finds that the solution to the variational equation is  $\bar{\rho}(x) = \alpha[|x|^{-1} - R^{-1}]^{1/2}$  for  $|x| \leq R$  and  $\bar{\rho}(x) = 0$  for  $|x| \geq R$ , with  $R = K_s \pi^{-4/3} Z^{-1}$ . Then

$$h(\bar{\rho}) = Z^2(\pi/2)^{4/3}/K_s = -\frac{4}{3} Z^2 \text{ Ry}.$$

(Recall that one Rydberg =  $\text{Ry} = \frac{1}{4}$  in these units.) Thus, Eq. (5) leads easily to the conclusion

$$E_0 \geq -\frac{4}{3} Z^2 \text{ Ry} \quad (9)$$

and this is an excellent lower bound to the correct  $E_0 = -Z^2 \text{ Ry}$ , especially since no differential equation had to be solved.

In anticipation of later developments, a weaker, but also useful, form of Eq. (5) can be derived. By Hölder's inequality<sup>2</sup>

$$\int \rho(x)^{5/3} dx \leq \left\{ \int \rho(x)^3 dx \right\}^{1/3} \left\{ \int \rho(x) dx \right\}^{2/3} \quad (10)$$

and, since we always take  $\int |\psi|^2 = 1$ ,

$$T_\psi \geq K_s \int \rho_\psi(x)^{5/3} dx. \quad (11)$$

Note that there is now an exponent 1 outside the integral. Although  $K_s$  is the best constant in (5) it is not the best constant in (11). Call the latter  $K_1$ .  $K_1$  is the minimum of

$$\frac{\int |\nabla \psi(x)|^2 dx}{\int \rho(x)^{5/3} dx}$$

subject to  $\int \rho(x) dx = 1$ . This leads to a nonlinear Schrödinger equation whose numerical solution yields (J. F. Barnes, private communication)

$$K_1 \approx 9.578.$$

In any event

$$K_1 > K^c \equiv \left(\frac{3}{5}\right)(6\pi^2)^{2/3} \approx 9.116,$$

<sup>2</sup>Hölder's inequality states that

$$\left| \int f(x)g(x) dx \right| \leq \left\{ \int |f(x)|^p dx \right\}^{1/p} \left\{ \int |g(x)|^q dx \right\}^{1/q}$$

when  $p^{-1} + q^{-1} = 1$  and  $p \geq 1$ . To obtain (10) take  $f = \rho$ ,  $g = \rho^{2/3}$ ,  $p = 3$ ,  $q = \frac{3}{2}$ .

and hence

$$T_\psi \geq K^c \int \rho_\psi(x)^{5/3} dx \text{ when } \int |\psi(x)|^2 dx = 1. \quad (12)$$

$K^c$  is much bigger than  $K_s$ ; it is the classical value and will be encountered again in Sec. II and in Sec. III, where its significance will be clarified.

We can repeat the minimization calculation analogous to Eq. (8) using the bound (12) and the functional

$$h^c(\rho) = K^c \int \rho(x)^{5/3} dx - Z \int |x|^{-1} \rho(x) dx.$$

(We could, of course, use the better constant  $K_1$ .) This time

$$\bar{\rho}(x) = \left\{ \frac{3}{5} (Z/K^c) (|x|^{-1} - R^{-1}) \right\}^{3/2} \quad (13)$$

for  $|x| \leq R$ .  $R$  is determined by  $\int \rho = 1$  and one finds that  $R = (K^c/Z)(4/\pi^2)^{2/3}$  and

$$E_0 \geq -(9Z^2/5K^c)(\pi^2/4)^{2/3} = -3^{1/3} Z^2 \text{ Ry}. \quad (14)$$

The quantity  $3^{1/3}$  is only 8.2% greater than  $\frac{4}{3}$ .

The Sobolev inequality (5) or its variant (12) is, for our purposes, a much better uncertainty principle than Heisenberg's—indeed it is also fairly accurate. We now want to extend (12) to the  $N$ -particle case in order to establish the stability of bulk matter. The important new fact that will be involved is that the  $N$  particles are *fermions*; that is to say the  $N$ -particle wave function is an antisymmetric function of the  $N$ -space, spin variables.

## II. EXTENSION OF THE UNCERTAINTY PRINCIPLE TO MANY FERMIONS

A well known elementary calculation is that of the lowest kinetic energy,  $T^V$  of  $N$  fermions in a cubic box of volume  $V$ . For large  $N$  one finds that

$$T^V \approx q^{-2/3} K^c V \rho^{5/3}, \quad (15)$$

where  $\rho = N/V$  and  $q$  is the number of spin states available to each particle ( $q = 2$  for electrons). Equation (15) is obtained by merely adding up the  $N/q$  lowest eigenvalues of  $-\Delta$  with Dirichlet ( $\psi = 0$ ) boundary conditions on the walls of the box. The important feature of (15) is that it is proportional to  $N^{5/3}$  instead of  $N$ , as would be the case if the particles were not fermions. The extra factor  $N^{2/3}$  is essential for the stability of matter; if electrons were bosons, matter would not be stable.

Equation (15) suggests that Eq. (12), with a factor  $q^{-2/3}$  ought to extend to the  $N$ -particle case if  $\rho(x)$  is interpreted properly. The idea is old, going back to Lenz (1932), who got it from Thomas–Fermi theory. The proof that something like (12) is not only an approximation but is also a lower bound is new.

To say that the  $N$  particles are *fermions* with  $q$  spin states means that the  $N$ -particle wave function  $\psi(x_1, \dots, x_N; \sigma_1, \dots, \sigma_N)$  defined for  $x_i \in \mathbb{R}^3$  and  $\sigma_i \in \{1, 2, \dots, q\}$  is *antisymmetric* in the pairs  $(x_i, \sigma_i)$ . The norm is given by

$$\langle \psi, \psi \rangle = \sum_{\sigma_i=1}^q \int |\psi(x_1, \dots, x_N; \sigma_1, \dots, \sigma_N)|^2 dx_1 \dots dx_N.$$

Define

$$T_\psi = N \sum_{\sigma_i=1}^q \int |\nabla_1 \psi(x_1, \dots, x_N; \sigma_1, \dots, \sigma_N)|^2 dx_1 \dots dx_N \quad (16)$$

to be the usual kinetic energy of  $\psi$  and define

$$\rho_\psi(x) = N \sum_{\sigma_i=1}^q \int |\psi(x, x_2, \dots, x_N; \sigma_1, \dots, \sigma_N)|^2 dx_2 \dots dx_N \quad (17)$$

to be the single particle density, i.e., the probability of finding a particle at  $x$ . The analog of (12) is the following (Lieb and Thirring, 1975):

**Theorem 1.** If  $\langle \psi, \psi \rangle = 1$  then

$$T_\psi \geq (4\pi q)^{-2/3} K^c \int \rho_\psi(x)^{5/3} dx. \quad (18)$$

Apart from the annoying factor  $(4\pi)^{-2/3} \approx 0.185$ , (18) says that the intuition behind considering (15) as a lower bound is correct. We believe that  $(4\pi)^{-2/3}$  does not belong in (18) and hope to eliminate it someday. Recent work (Lieb, 1976) has improved the constant by a factor  $(1.83)^{2/3} = 1.496$ , so we are now off from the conjectured constant  $q^{-2/3} K^c$  only by the factor 0.277.

The proof of Theorem 1 is not long but it is slightly tricky. It is necessary first to investigate the negative eigenvalues of a one-particle Schrödinger equation when the potential is nonpositive.

**Theorem 2.** Let  $V(x) \leq 0$  be a potential for the one-particle, three dimensional Schrödinger operator  $H = -\Delta + V(x)$  on  $L^2(\mathbb{R}^3)$ . For  $E < 0$  let  $N_E(V)$  be the number of eigenstates of  $H$  with energies  $\leq E$ . Then

$$N_E(V) \leq (4\pi)^{-1} (2|E|)^{-1/2} \int |V(x) - E/2|_-^2 dx, \quad (19)$$

where  $|f(x)|_- = |f(x)|$  if  $f(x) \leq 0$  and  $|f(x)|_- = 0$  otherwise.

**Corollary.** If  $e_1 \leq e_2 \leq \dots \leq 0$  are the negative eigenvalues of  $H$  (if any) then

$$\sum_{j=1}^{\infty} |e_j| \leq \frac{4\pi}{15\pi^2} \int |V(x)|^{5/2} dx. \quad (20)$$

**Proof.**  $\sum |e_j| = \int_0^\infty N_{-\alpha}(V) d\alpha$ . Insert (19) and do the  $\alpha$  integration first and then the  $x$  integration. The result is (20). ■

We believe the factor  $(4\pi)$  does not belong in (20).

**Proof of Theorem 2.** From the Schrödinger equation  $H\psi = e\psi$  it is easy to deduce that  $N_E(V)$  is equal to the number of eigenvalues which are  $\geq 1$  of the positive definite Birman–Schwinger operator (Birman, 1961; Schwinger, 1961)

$$B_E(V) = |V|^{1/2} (-\Delta - E)^{-1} |V|^{1/2}. \quad (21)$$

Essentially Eq. (21) comes from the fact that if  $H\psi = e\psi$  then  $(-\Delta - e)\psi = |V|\psi$ . If one defines  $|V|^{1/2}\psi \equiv \phi$ , then  $B_E\phi = \phi$ . Thus  $B_E$  has an eigenvalue 1 when  $e$  is an eigenvalue. However,  $B_E$  is a compact positive semidefinite operator on  $L^2(\mathbb{R}^3)$  for  $E < 0$  and, as an operator,  $B_E$  is monotone increasing in  $E$ . Thus, if  $B_E$  has  $k$  eigenvalues  $\geq 1$ , there exist  $k$  numbers  $e_1 \leq e_2 \leq \dots \leq e_k \leq E$  such that

$B_{e_j}$  has eigenvalue 1.

Consequently  $N_E(V) \leq \text{Tr} B_E(V)^2$ . On the other hand,  $N_E(V) \leq N_{E/2}(-|V-E/2|_-)$  by the variational principle (draw a graph of  $V(x) - E/2$ ). Thus, since  $B_E(V)$  has a kernel,  $B_E(x, y) = |V(x)|^{1/2} \exp\{-|E|^{1/2}|x-y|\} [4\pi|x-y|]^{-1} \times |V(y)|^{1/2}$ , one has that

$$\begin{aligned} N_E(V) &\leq \text{Tr} B_{E/2}(-|V-E/2|_-)^2 \\ &= (4\pi)^{-2} \int \int dx dy |V(x) - E/2|_- |V(y) - E/2|_- \\ &\quad \times \exp\{-(2|E|)^{1/2}|x-y|\} |x-y|^{-2}. \end{aligned} \quad (22)$$

Equation (19) results from applying Young's inequality<sup>3</sup> to Eq. (22). Alternatively, one can do the convolution integral by Fourier transforms and note that the Fourier transform of the last factor has a maximum at  $p=0$ , where it is  $4\pi(2|E|)^{-1/2}$ . ■

Using (20), which is a statement about the energy levels of a single particle Hamiltonian, we can, surprisingly, prove Theorem 1, which refers to the kinetic energy of  $N$  fermions.

*Proof of Theorem 1.*  $\psi$  and hence  $\rho_\psi(x)$  are given. Consider the non-positive single particle potential  $V(x) \equiv -\alpha\rho_\psi(x)^{2/3}$  where  $\alpha$  is given by  $(2/3\pi)q\alpha^{3/2}=1$ . Next consider the following  $N$ -particle Hamiltonian:

$$\tilde{H}_N = \sum_{i=1}^N h_i; \quad h_i = -\Delta_i + V(x_i)$$

on  $L^2(\mathbb{R}^3; \mathbb{C}^q)^N$ . If  $E_0$  is the fermion ground state energy of  $\tilde{H}_N$ , we have that  $E_0 \geq q \sum e_j$ , where the  $e_j$  are the negative eigenvalues of the single particle Hamiltonian  $h$ . (We merely fill the lowest negative energy levels  $q$  times until the  $N$  particles are accounted for; if there are  $k$  such levels and if  $N < kq$  then  $E_0 > q \sum e_j$ . If  $N > kq$ , the surplus particles can be placed in wave packets far away from the origin with arbitrarily small kinetic energy.) On the other hand,  $E_0 \leq \langle \psi, \tilde{H}_N \psi \rangle = T_\psi - \alpha \int \rho_\psi(x)^{5/3} dx$  by the variational principle. If these two inequalities are combined together with (20), which says that  $\sum e_j \geq -(4/15\pi)\alpha^{5/2} \int \rho_\psi(x)^{5/3} dx$ , then (18) is the result. ■

It might not be too much out of place to explain at this point why  $K^c$  is called the classical constant. The name does not stem from its antiquity, as in the ideal gas kinetic energy (15), but rather from classical mechanics—more precisely the semiclassical approximation to quantum mechanics. This intuitive idea is valuable.

As the proof of Theorem 1 shows, the constant in (18) for  $T_\psi$  is simply related to the constant in (20) for the

sum of the eigenvalues. The point is that the semiclassical approximation to this sum is

$$\sum |e_j| \approx (15\pi^2)^{-1} \int |V(x)|^{5/2} dx,$$

and this, in turn, would yield (18) without the  $(4\pi)^{-2/3}$  factor. The semiclassical approximation is obtained by saying that a region of volume  $(2\pi)^3$  in the six-dimensional phase space  $(p, x)$  can accommodate one eigenstate. Hence, integrating over the set  $\theta(H)$ , in which  $H(p, x) = p^2 + V(x)$  is negative,

$$\begin{aligned} \sum e_j &\approx (2\pi)^{-3} \int_{\theta(H)} dx dp \{p^2 + V(x)\} \\ &= (2\pi)^{-3} \int dx 4\pi \int_0^{|V(x)|^{1/2}} p^2 dp \{p^2 + V(x)\} \\ &= -(15\pi^2)^{-1} \int |V(x)|^{5/2} dx. \end{aligned}$$

If a coupling constant  $g$  is introduced, and if  $V$  is replaced by  $gV$ , then it is a theorem that the semiclassical approximation is asymptotically exact as  $g \rightarrow \infty$  for any  $V$  in  $L^{5/2}(\mathbb{R}^3)$ .

Theorem 1 gives a lower bound to the kinetic energy of fermions which is crucial for the  $H$  stability of matter as developed in Sec. IV. To appreciate the significance of Theorem 1 it should be compared with the one-particle Sobolev bound (12). Suppose that  $\rho(x) = 0$  outside some fixed domain  $\Omega$  of volume  $V$ . Then since

$$\int_\Omega \rho(x)^{5/3} dx \geq \left\{ \int_\Omega \rho(x) dx \right\}^{5/3} \left\{ \int_\Omega 1 \right\}^{-2/3} = N^{5/3} V^{-2/3}$$

by Hölder's inequality, one sees that  $T_\psi$  grows at least as fast as  $N^{5/3}$ . Using Eq. (12) alone, one would only be able to conclude that  $T_\psi$  grows as  $N$ . This distinction stems from the Pauli principle, i.e., the antisymmetric nature of the  $N$ -particle wave function. As we shall see, this  $N^{5/3}$  growth is essential for the stability of matter because without it the ground state energy of  $N$  particles with Coulomb forces would grow at least as fast as  $-N^{7/5}$  instead of  $-N$ .

The Fermi pressure is needed to prevent a collapse, but to learn how to exploit it we must first turn to another chapter in the theory of Coulomb systems, namely Thomas-Fermi theory.

### III. THOMAS-FERMI THEORY

The statistical theory of atoms and molecules was invented independently by Thomas and Fermi (Thomas, 1927; Fermi, 1927). For many years the TF theory was regarded as an uncertain approximation to the  $N$ -particle Schrödinger equation and much effort was devoted to trying to determine its validity (e.g., Gombás, 1949). It was eventually noticed numerically (Sheldon, 1955) that molecules did not appear to bind in this theory, and then Teller (1962) proved this to be a general theorem.

It is now understood that TF theory is really a large  $Z$  theory (Lieb and Simon, 1977); to be precise it is exact in the limit  $Z \rightarrow \infty$ . For finite  $Z$ , TF theory is qualita-

<sup>3</sup>Young's inequality states that

$$\begin{aligned} \left| \iint f(x)g(x-y)h(y) dx dy \right| &\leq \left\{ \int |f(x)|^p dx \right\}^{1/p} \\ &\quad \times \left\{ \int |g(x)|^q dx \right\}^{1/q} \\ &\quad \times \left\{ \int |h(x)|^r dx \right\}^{1/r} \end{aligned}$$

when  $p^{-1} + q^{-1} + r^{-1} = 2$  and  $p, q, r \geq 1$ . For (22) take  $p=r=2$  and  $q=1$ .

tively correct in that it adequately describes the bulk of an atom or molecule. It is not precise enough to give binding. Indeed, it should not do so because binding in TF theory would imply that the cores of atoms bind, and this does not happen. Atomic binding is a fine quantum effect. Nevertheless, TF theory deserves to be well understood because it is exact in a limit; the TF theory is to the many-electron system as the hydrogen atom is to the few-electron system. For this reason the main features of the theory are presented here, mostly without proof.

A second reason for our interest in TF theory is this: in the next section the problem of the  $H$  stability of matter will be reduced to a TF problem. The knowledge that TF theory is  $H$  stable (this is a corollary of the no binding theorem) will enable us to conclude that the true quantum system is  $H$  stable.

The Hamiltonian for  $N$  electrons with  $k$  static nuclei of charges  $z_i > 0$  and locations  $R_i$  is

$$H_N = \sum_{i=1}^N -\Delta_i - V(x_i) + \sum_{1 \leq i < j \leq N} |x_i - x_j|^{-1} + U(\{z_j, R_j\}_{j=1}^k), \quad (23)$$

where

$$V(x) = \sum_{j=1}^k z_j |x - R_j|^{-1}, \quad (24a)$$

and

$$U(\{z_j, R_j\}_{j=1}^k) = \sum_{1 \leq i < j \leq k} z_i z_j |R_i - R_j|^{-1}. \quad (24b)$$

The nuclear-nuclear repulsion  $U$  is, of course, a constant term in  $H_N$  but it is included for two reasons:

(i) We wish to consider the dependence on the  $R_i$  of

$$E_N^Q(\{z_j, R_j\}_{j=1}^k) \equiv \text{the ground state energy of } H_N. \quad (25)$$

(ii) Without  $U$  the energy will not be bounded by  $N$ .

The nuclear kinetic energy is not included in  $H_N$ . For the  $H$ -stability problem we are only interested in finding a lower bound to  $E_N^Q$ , and the nuclear kinetic energy adds a positive term. In other words,

$$\inf_{\{R_i\}} E_N^Q(\{z_j, R_j\}_{j=1}^k)$$

is smaller than the ground state energy of the true Hamiltonian [defined in Eq. (58)] in which the nuclear kinetic energy is included. Later on when we do the proper thermodynamics of the whole system we shall have to include the nuclear kinetic energy.

The problem of estimating  $E_N^Q$  is as old as the Schrödinger equation. The TF theory, as interpreted by Lenz (1932), reads as follows: For fermions having  $q$  spin states ( $q=2$  for electrons) define the TF energy functional:

$$\begin{aligned} \mathcal{G}(\rho) = & q^{-2/3} K^c \int \rho(x)^{5/3} - \int V(x) \rho(x) \\ & + \frac{1}{2} \iint \rho(x) \rho(y) |x - y|^{-1} dx dy + U(\{z_j, R_j\}_{j=1}^k) \end{aligned} \quad (26)$$

for non-negative functions  $\rho(x)$ . Then for  $\lambda \geq 0$

$$E_\lambda^{\text{TF}} \equiv \inf \left\{ \mathcal{G}(\rho) : \int \rho(x) dx = \lambda \right\} \quad (27)$$

is the TF energy for  $\lambda$  electrons ( $\lambda$  need not be an integer, of course). When  $\lambda = N$ , the minimizing  $\rho$  is supposed to approximate the  $\rho_\psi$  given by (17), wherein  $\psi$  is the true ground state wave function, and  $E_N^{\text{TF}}$  is supposed to approximate  $E_N^Q$ .

The intuitive idea behind TF theory is this: If  $\psi$  is any fermion wave function and  $T_\psi$  and  $\rho_\psi$  are given by Eqs. (16) and (17), then the first term in (26) is supposed to approximate  $T_\psi$ . This is based on the box kinetic energy (15). The last three terms in (26) represent, respectively, the electron-nuclear, electron-electron, and nuclear-nuclear Coulomb energy.  $E_\lambda^{\text{TF}}$  in (27) is then the "ground state energy" of (26).

The second and fourth terms on the right side of (26) are exact but the first and third are not. The first is to some extent justified by the kinetic energy inequality, Theorem 1; the third term will be discussed later. In any event, Eqs. (26) and (27) define TF theory.

It would be too much to try to reproduce here the details of our analysis of TF theory. A short summary of some of the main theorems will have to suffice.

The first question is whether or not  $E_\lambda^{\text{TF}}$  (which, by simple estimates using Young's and Hölder's inequalities, can be shown to be finite for all  $\lambda$ ) is a minimum as distinct from merely an infimum. The distinction is crucial because the TF equation [the Euler-Lagrange equation for (26) and (27)]

$$\frac{5}{3} K^c q^{-2/3} \rho^{2/3}(x) = \max\{\phi(x) - \mu, 0\} \quad (28)$$

with

$$\phi(x) = V(x) - \int \rho(y) |x - y|^{-1} dy \quad (29)$$

has a solution with  $\int \rho = \lambda$  if and only if there is a minimizing  $\rho$  for  $E_\lambda^{\text{TF}}$ . The basic theorem is as follows.

**Theorem 3.** If  $\lambda \leq Z \equiv \sum_{j=1}^k z_j$  then

(i)  $\mathcal{G}(\rho)$  has a minimum on the set  $\int \rho(x) dx = \lambda$ .  
(ii) This minimizing  $\rho$  (call it  $\rho_\lambda^{\text{TF}}$ ) is unique and satisfies (28) and (29).  $\mu$  is non-negative, and  $-\mu$  is the chemical potential, i.e.,

$$-\mu = dE_\lambda^{\text{TF}}/d\lambda. \quad (30)$$

(iii) There is no other solution to (28) and (29) (for any  $\mu$ ) with  $\int \rho = \lambda$  other than  $\rho_\lambda^{\text{TF}}$ .

(iv) When  $\lambda = Z$ ,  $\mu = 0$ . Otherwise  $\mu > 0$ , i.e.,  $E_\lambda^{\text{TF}}$  is strictly decreasing in  $\lambda$ .

(v) As  $\lambda$  varies from 0 to  $Z$ ,  $\mu$  varies continuously from  $+\infty$  to 0.

(vi)  $\mu$  is a convex, decreasing function of  $\lambda$ .

(vii)  $\phi_\lambda^{\text{TF}}(x) > 0$  for all  $x$  and  $\lambda$ . Hence when  $\lambda = Z$

$$\frac{5}{3} K^c q^{-2/3} \rho_Z^{\text{TF}}(x)^{2/3} = \phi_Z^{\text{TF}}(x).$$

If  $\lambda > Z$  then  $E_\lambda^{\text{TF}}(\lambda)$  is not a minimum and (28) and (29) have no solution with  $\int \rho = \lambda$ . Negative ions do not exist in TF theory. Nevertheless,  $E_\lambda^{\text{TF}}$  exists and  $E_\lambda^{\text{TF}} = E_Z^{\text{TF}}$  for  $\lambda \geq Z$ .

The proof of Theorem 3 is an exercise in functional

analysis. Basically, one first shows that  $\mathcal{E}(\rho)$  is bounded below so that  $E_\lambda^{\text{TF}}$  exists. The Banach–Alaoglu theorem is used to find an  $L^{5/3}$  weakly convergent sequence of  $\rho$ 's such that  $\mathcal{E}(\rho)$  converges to  $E_\lambda^{\text{TF}}$ . Then one notes that  $\mathcal{E}(\rho)$  is weakly lower semicontinuous so that a minimizing  $\rho$  exists under the subsidiary condition that  $\int \rho \leq \lambda$ . The uniqueness comes from an important property of  $\mathcal{E}(\rho)$ , namely that it is *convex*. This also implies that the minimizing  $\rho$  satisfies  $\int \rho = \lambda$ .

A major point to notice is that a solution of the TF equation is obtained as a by-product of minimizing  $\mathcal{E}(\rho)$ ; a direct proof that the TF equation has a solution would be very complicated.

Only in the case  $\lambda \leq Z$  is  $\rho_\lambda^{\text{TF}}(x)$  positive for all  $x$ , when  $\lambda < Z$ ,  $\mu > 0$  and, since  $\phi_\lambda^{\text{TF}}(x)$  goes to zero as  $|x|$  goes to infinity, Eq. (28) implies that  $\rho_\lambda^{\text{TF}}(x)$  vanishes outside some bounded set.

Apart from the details presented in Theorem 3, the main point is that TF theory is well defined. In particular the density  $\rho^{\text{TF}}$  is unique—a state of affairs in marked contrast to that of Hartree–Fock theory (Hartree, 1927–28; Fock, 1930; Slater, 1930; Lieb and Simon, 1973).

The TF density  $\rho_\lambda^{\text{TF}}$  has the following properties:

*Theorem 4.* If  $\lambda \leq Z$  then

$$(i) \left(\frac{5}{3}\right) K^c q^{-2/3} \rho_\lambda^{\text{TF}}(x)^{2/3} \sim z_i |x - R_i|^{-1} \quad (31)$$

near each  $R_i$ .

(ii) In the neutral case,  $\lambda = Z = \sum_{j=1}^k z_j$ ,

$$|x|^6 \rho_\lambda^{\text{TF}}(x) \rightarrow (3/\pi)^3 \left[\frac{5}{3} K^c q^{-2/3}\right]^3 \quad (32)$$

as  $|x| \rightarrow \infty$ , irrespective of the distribution of the nuclei.

(iii)  $\phi_\lambda^{\text{TF}}(x)$  and  $\rho_\lambda^{\text{TF}}(x)$  are real analytic in  $x$  away from all the  $R_i$ , on all of 3-space in the neutral case and on  $\{x: \phi_\lambda^{\text{TF}}(x) > \mu\}$  in the positive ionic case.

Equation (32) is especially remarkable: at large distances one loses all knowledge of the nuclear charges and configuration. Property (i) recalls the singularity found in the minimization of  $h^c(\rho)$  [see Eq. (13)].

Equation (31) can be seen from (28) and (29) by inspection. Equation (32) is more subtle but it is consistent with the observation that (28) and (29) can be rewritten (when  $\mu = 0$ ) as

$$-(4\pi)^{-1} \Delta \phi_\lambda^{\text{TF}}(x) = -\left\{\left(\frac{3}{5}\right) q^{2/3} \phi_\lambda^{\text{TF}}(x) / K^c\right\}^{3/2}$$

away from the  $R_i$ . If it is assumed that  $\phi_\lambda^{\text{TF}}(x)$  goes to zero as a power of  $|x|$  then (32) follows. This observation was first made by Sommerfeld (1932). The proof that a power law falloff actually occurs is somewhat subtle and involves potential theoretic ideas such as that used in the proof of Lemma 8.

As pointed out earlier, the connection between TF theory and the Schrödinger equation is best seen in the limit  $Z \rightarrow \infty$ . Let the number,  $k$ , of nuclei be held fixed, but let  $N \rightarrow \infty$  and  $z_i \rightarrow \infty$  in such a way that the degree of ionization  $N/Z$  is constant, where

$$Z = \sum_{j=1}^k z_j.$$

To this end we make the following definition: Fix  $\{z_j, R_j\}_{j=1}^k$  and  $\lambda$ . It is not necessary to assume that  $\lambda$

$\leq Z$ . For each  $N=1, 2, \dots$  define  $a_N$  by  $\lambda a_N = N$ . In  $H_N$  (23) replace  $z_j$  by  $z_j a_N$  and  $R_j$  by  $R_j a_N^{-1/3}$ . This means that the nuclei come together as  $N \rightarrow \infty$ . If they stay at fixed positions then that is equivalent, in the limit, to isolated atoms, i.e., it is equivalent to starting with all the nuclei infinitely far from each other. Finally, for the nuclear configuration  $\{a_N z_j, a_N^{-1/3} R_j\}_{j=1}^k$  let  $\psi_N$  be the ground state wave function,  $E_N^Q$  the ground state energy, and  $\rho_N^Q(x)$  be the single particle density as defined by Eq. (17).<sup>4</sup>

It is important to note that there is a simple and obvious scaling relation for TF theory, namely

$$E_{\lambda a}^{\text{TF}}(\{a z_j, a^{-1/3} R_j\}_{j=1}^k) = a^{7/3} E_\lambda^{\text{TF}}(\{z_j, R_j\}_{j=1}^k) \quad (33)$$

and the densities for the two systems are related by

$$\rho_{\lambda a}^{\text{TF}}(a^{-1/3} x) = a^2 \rho_\lambda^{\text{TF}}(x) \quad (34)$$

for any  $a \geq 0$ . Hence, for the above sequence of systems parametrized by  $a_N$ ,

$$a_N^{-7/3} E_N^{\text{TF}}(\{a_N z_j, a_N^{-1/3} R_j\}_{j=1}^k) = E_\lambda^{\text{TF}}(\{z_j, R_j\}_{j=1}^k), \quad (35)$$

$$a_N^{-2} \rho_N^{\text{TF}}(a_N^{-1/3} x) = \rho_\lambda^{\text{TF}}(x) \quad (36)$$

for all  $N$ .

If, on the other hand, the nuclei are held fixed then one can prove that

$$\lim_{N \rightarrow \infty} a_N^{-7/3} E_N^{\text{TF}}(\{a_N z_j, R_j\}) = \sum_{j=1}^k E_{\lambda_j}^{\text{TF}}(z_j), \quad (37)$$

where  $E_{\lambda_j}^{\text{TF}}(z)$  is the energy of an isolated atom of nuclear charge  $z$ . The  $\lambda_j$  are determined by the condition that  $\sum_{j=1}^k \lambda_j = \lambda$  if  $\lambda \leq Z$  (otherwise,  $\sum_{j=1}^k \lambda_j = Z$ ) and that the chemical potentials of the  $k$  atoms are all the same. Another way to say this is that the  $\lambda_j$  minimize the right side of Eq. (37). With the nuclei *fixed*, the analog of (36) is

$$\lim_{N \rightarrow \infty} a_N^{-2} \rho_N^{\text{TF}}(a_N^{-1/3}(x - R_j)) = \rho_{\lambda_j}^{\text{TF}}(x). \quad (38)$$

The right side of Eq. (38) is the  $\rho$  for a single atom of nuclear charge  $z$  and electron charge  $\lambda_j$ . Equations (37) and (38) are a precise statement of the fact that isolated atoms result from fixing the  $R_j$ .

The TF energy for an isolated, *neutral atom* of nuclear charge  $Z$  is found numerically to be

$$E_Z^{\text{TF}} = -(2.21) q^{2/3} (K^c)^{-1} Z^{7/3}. \quad (39)$$

For future use, note that  $E_Z^{\text{TF}}$  is proportional to  $1/K^c$ . Thus, if one considers a TF theory with  $K^c$  replaced by some other constant  $\alpha > 0$ , as will be necessary in Sec. IV, then Eq. (39) is correct if  $K^c$  is replaced by  $\alpha$ .

*Theorem 5.* With  $a_N = N/\lambda$  and  $\{z_j, R_j\}_{j=1}^k$  fixed

- (i)  $a_N^{-7/3} E_N^Q(\{a_N z_j, a_N^{-1/3} R_j\}_{j=1}^k)$  has a limit as  $N \rightarrow \infty$ .
- (ii) This limit is  $E_\lambda^{\text{TF}}(\{z_j, R_j\}_{j=1}^k)$ .
- (iii)  $a_N^{-7/3} E_N^Q(\{a_N z_j, R_j\}_{j=1}^k)$  has a limit as  $N \rightarrow \infty$ . This limit is the right side of (37).
- (iv)  $a_N^{-2} \rho_N^Q(a_N^{-1/3} x; \{a_N z_j, a_N^{-1/3} R_j\}_{j=1}^k)$  also has a limit as

<sup>4</sup>If  $E_N^Q$  is degenerate,  $\psi_N$  can be *any* ground state wave function as far as Theorem 5 is concerned. If  $E_N^Q$  is not an eigenvalue, but merely  $\inf \text{spec } H_N$ , then it is possible to define an approximating sequence  $\psi_N$ , with  $\rho_N^Q$  still given by Eq. (17), in such a way that Theorem 5 holds. We omit the details of this construction here.

$N \rightarrow \infty$ . If  $\lambda \leq Z = \sum_{j=1}^k z_j$ , this limit is  $\rho_\lambda^{\text{TF}}(x)$  and the convergence is in weak  $L^1(\mathbf{R}^3)$ . If  $\lambda > Z$ , the limit is  $\rho_Z^{\text{TF}}(x)$  in weak  $L^1_{\text{loc}}(\mathbf{R}^3)$ .

(v) For fixed nuclei,  $a_N^{-2} \rho_N^Q(a_N^{-1/3}(x - R_j); \{a_N z_j, R_j\}_{j=1}^k)$  has a limit [in the same sense as (iv)] which is the right side of (38).

The proof of Theorem 5 does not use anything introduced so far. It is complicated, but elementary. One partitions 3-space into boxes with sides of order  $Z^{-1/3}$ . In each box the potential is replaced by its maximum (respectively, minimum) and one obtains an upper (respectively, lower) bound to  $E_N^Q$  by imposing Dirichlet ( $\psi=0$ ) (respectively, Neumann ( $\nabla\psi=0$ )) boundary conditions on the boxes. The upper bound is essentially a Hartree-Fock calculation. The  $-r^{-1}$  singularity near the nuclei poses a problem for the lower bound, and it is tamed by exploiting the concept of angular momentum barrier.

What Theorem 5 says, first of all, is that the true quantum energy has a limit on the order of  $Z^{7/3}$  when the ratio of electron to nuclear charge is held fixed. Second, this limit is given correctly by TF theory as is shown in Eq. (35). The requirement that the nuclei move together as  $Z^{-1/3}$  should be regarded as a refinement rather than as a drawback, for if the nuclei are fixed a limit also exists but it is an uninteresting one of isolated atoms.

Theorem 5 also says that the density  $\rho_N^Q$  is proportional to  $Z^2$  and has a scale length proportional to  $Z^{-1/3}$ . If  $\lambda > Z$ , Theorem 5 states that the surplus charge moves off to infinity and the result is a neutral molecule. This means that large atoms or molecules cannot have a negative ionization proportional to the total nuclear charge; at best they can have a negative ionization which is a vanishingly small fraction of the total charge. This result is physically obvious for electrostatic reasons, but it is nice to have a proof of it.

Theorem 5 also resolves certain "anomalies" of TF theory:

(a) In real atoms or molecules the electron density falls off exponentially, while in TF theory (Theorem 4) the density falls off as  $|x|^{-6}$ .

(b) The TF atom shrinks in size as  $Z^{-1/3}$  [cf. Eq. (36)] while real large atoms have roughly constant size.

(c) In TF theory there is no molecular binding, as we shall show next, but there is binding for real molecules.

(d) In real molecules the electron density is finite at the nuclei, but in TF theory it goes to infinity as  $z_j |x - R_j|^{-3/2}$  (Theorem 4).

As Theorem 5 shows, TF theory is really a theory of heavy atoms or molecules. A large atom looks like a stellar galaxy, poetically speaking. It has a core which shrinks as  $Z^{-1/3}$  and which contains most of the electrons. The density (on a scale of  $Z^2$ ) is not finite at the nucleus because, as the simplest Bohr theory shows, the S-wave electrons have a density proportional to  $Z^3$  which is infinite on a scale of  $Z^2$ . Outside the core is a mantle in which the density is proportional to (cf. Theorem 4)  $(3/\pi)^3 [(\frac{5}{3})K^2 2^{2/3}]^3 Z^2 / (Z^{1/3} |x|)^6$ , which is independent of  $Z$ ! This density is correct to infinite distances on a length scale  $Z^{-1/3}$ . The core and the man-

tle contain 100% of the electrons as  $Z \rightarrow \infty$ . The third region is a transition region to the outer shell, and while it may contain many electrons, it contains only a vanishingly small fraction of them. The fourth region is the outer shell in which chemistry and binding takes place. TF theory has nothing to say about this region. The fifth region is the one in which the density drops off exponentially.

Thus, TF theory deals only with the core and the mantle in which the bulk of the energy and the electrons reside. There ought not to be binding in TF theory, and indeed there is none, because TF energies are proportional to  $Z^{7/3}$  and binding energies are of order one. The binding occurs in the fourth layer.

An important question is what is the next term in the energy beyond the  $Z^{7/3}$  term of TF theory. Several corrections have been proposed: (e.g., Dirac, 1930; Von Weizsäcker, 1935; Kirzhnits, 1957; Kompaneets and Pavlovskii, 1956; Scott, 1952). With the exception of the last, all these corrections are of order  $Z^{5/3}$ . Scott (as late as 1952!) said there should be a  $Z^{6/3}$  correction because TF theory is not able to treat correctly the innermost core electrons. Let us give a heuristic argument. Recall that in Bohr theory each inner electron alone has an energy proportional to  $Z^2$ . As these inner electrons are unscreened, their energies should be independent of the presence or absence of the electron-electron repulsion. In other words, the  $Z^2$  correction for a molecule should be precisely a sum of corrections, one for each atom. The atomic correction should be the difference between the Bohr energy and the  $Z^{7/3}$  TF energy for an atom in which the electron-electron repulsion is neglected. We already calculated the TF energy for such an "atom" in Eq. (14) (put  $Z=1$  there and then use scaling; also replace  $K^c$  by  $q^{-2/3}K^c$ ). Thus, for a neutral atom without electron-electron repulsion

$$\tilde{E}_Z^{\text{TF}} = -(3^{1/3}/4)q^{2/3}Z^{7/3}. \quad (40)$$

For the Bohr atom, each shell of energy  $-Z^2/4n^2$  has  $n^2$  states, so

$$\frac{Z}{q} = \frac{N}{q} = \sum_{n=1}^L n^2 + (L+1)^2 \phi = \frac{L^3}{3} + \frac{L^2}{2} + \frac{L}{6} + (L+1)^2 \phi$$

with  $0 \leq \phi \leq 1$  being the fraction of the  $(L+1)$ th shell that is filled. One finds  $L \approx (3Z/q)^{1/3} - \frac{1}{2} - \phi + o(1)$  and

$$E_Z^{\text{Bohr}} = -\frac{Z^2}{4}q \left\{ \phi + \sum_{n=1}^L 1 \right\} \approx \tilde{E}_Z^{\text{TF}} + \frac{q}{8}Z^2.$$

Thus, to the next order, the energy should be

$$E_N^Q(\{z_j, R_j\}_{j=1}^k) = E_N^{\text{TF}}(\{z_j, R_j\}_{j=1}^k) + \frac{1}{4} \sum_{j=1}^k z_j^2 + \text{lower order}, \quad (41)$$

since  $q=2$  for electrons. Note that  $E_N^{\text{TF}} \sim q^{2/3}$  while the Scott correction is proportional to  $q$ .

It is remarkable that Eq. (41) gives a precise conjecture about the next correction. It is simple to understand physically, yet we do not have the means to prove it.

The third main fact about TF theory is that there is no binding. This was proved by Teller in 1962. Considering the effort that went into the study of TF theory since its inception in 1927, it is remarkable that the no



binding phenomenon was not seriously noticed until the computer study of Sheldon in 1955. Teller's original proof involved some questionable manipulation with  $\delta$  functions and for that reason his result was questioned. His ideas were basically right, however, and we have made them rigorous.

**Theorem 6 (no binding).** If there are at least two nuclei, write the nuclear attraction  $V(x) = \sum_{j=1}^k z_j |x - R_j|^{-1}$  as the sum of two pieces,  $V = V^1 + V^2$  where  $V^1(x) = \sum_{j=1}^m z_j |x - R_j|^{-1}$  and  $1 \leq m < k$ . Let  $E_{\lambda}^{\text{TF},1}$  be the TF energy for the nuclei  $1, \dots, m$  (with  $U = \sum_{1 \leq i < j \leq m} z_i z_j |R_i - R_j|^{-1}$ , of course) and let  $E_{\lambda}^{\text{TF},2}$  be the same for the nuclei  $m+1, \dots, k$ . Given  $\lambda$ , let  $\lambda_1 \geq 0$  and  $\lambda_2 = \lambda - \lambda_1 \geq 0$  be chosen to minimize the sum of the energies of the separate molecules, i.e.,  $E_{\lambda_1}^{\text{TF},1} + E_{\lambda_2}^{\text{TF},2}$ . (If  $\lambda = Z = \sum_{j=1}^k z_j$  then by Theorem 3,  $\lambda_1 = \sum_{j=1}^m z_j$ .) Then

$$E_{\lambda}^{\text{TF}} \geq E_{\lambda_1}^{\text{TF},1} + E_{\lambda_2}^{\text{TF},2}. \quad (42)$$

Since the right side of Eq. (42) is the energy of two widely separated molecules, with the relative nuclear positions unchanged within each molecule, Theorem 6 says that the TF energy is unstable under every decomposition of the big molecule into smaller molecules. In particular, a molecule is unstable under decomposition into isolated atoms, and Theorem 9 is a simple consequence of this fact. One would suppose that if  $\lambda$  and the  $z_j$  are fixed, but the  $R_j$  are replaced by  $\alpha R_j$  then

$$E_{\lambda}^{\text{TF}}(\{z_j, \alpha R_j\}_{j=1}^k) \text{ is monotone decreasing in } \alpha.$$

In other words, the "pressure" is always positive. This is an unproved *conjecture*, but it has been proved (Balázs, 1967) in the case  $k=2$  and  $z_1 = z_2$ .

An interesting side remark is the following.

**Theorem 7.** If the TF energy (26), (27) is redefined by excluding the repulsion term  $U$  in (26), then the inequality in (42) is reversed.

Thus, the nuclear repulsion is essential for the no binding theorem 6.

Another useful fact for some further developments of the theory, especially the TF theory of solids and the TF theory of screening (Lieb and Simon, 1977) is the following lemma (also attributed to Teller), which is used to prove the main no binding theorem 6.

**Lemma 8.** Fix  $\{R_j\}_{j=1}^k$  and fix  $\mu \geq 0$  in the TF equation (28) but not  $\{z_j\}_{j=1}^k$ . (This means that as the  $z_j$ 's are varied  $\lambda$  will vary, but always  $0 \leq \lambda \leq Z = \sum z_j$ . If  $\mu = 0$  then  $\lambda = Z$  always.) If  $\{z_j^1\}_{j=1}^k$  and  $\{z_j^2\}_{j=1}^k$  are two sets of  $z$ 's such that

$$z_j^1 \leq z_j^2 \text{ all } j, \text{ and } z_1^1 < z_1^2$$

and if  $\lambda_1$  and  $\lambda_2$  are the corresponding  $\lambda$ 's for the two sets, then for all  $x$

$$\phi_{\lambda_1}^{\text{TF}}(x) \leq \phi_{\lambda_2}^{\text{TF}}(x)$$

and hence

$$\rho_{\lambda_1}^{\text{TF}}(x) \leq \rho_{\lambda_2}^{\text{TF}}(x).$$

There is strict inequality when  $\mu = 0$ . In short, increasing some  $z_j$  increases the density everywhere, not just on the average.

The proof of Lemma 8 involves a beautifully simple potential theoretic argument which we cannot resist giving.

**Proof of Lemma 8.** We want to prove  $\phi_{\lambda_1}^{\text{TF}}(x) \leq \phi_{\lambda_2}^{\text{TF}}(x)$  for all  $x$  and will content ourselves here with proving only  $\leq$  when  $\mu = 0$ . Let  $B = \{x: \phi_{\lambda_1}^{\text{TF}}(x) > \phi_{\lambda_2}^{\text{TF}}(x)\}$ .  $B$  is an open set and  $B$  does not contain any  $R_i$  for which  $z_i^1 < z_i^2$  by the TF equation (29). Let  $\psi(x) = \phi_{\lambda_1}^{\text{TF}}(x) - \phi_{\lambda_2}^{\text{TF}}(x)$ . If  $x \in B$  then  $\psi(x) > 0$  and, by (28),  $\rho_{\lambda_1}^{\text{TF}}(x) \geq \rho_{\lambda_2}^{\text{TF}}(x)$ . For  $x \in B$ ,  $-(4\pi)^{-1} \Delta \psi(x) = \rho_{\lambda_2}^{\text{TF}}(x) - \rho_{\lambda_1}^{\text{TF}}(x) \leq 0$ , so  $\psi$  is *subharmonic* on  $B$  [i.e.,  $\psi(x) \leq$  the average of  $\psi$  on any sphere contained in  $B$  and centered at  $x$ ]. Hence  $\psi$  has its maximum on the boundary of  $B$  or at  $\infty$ , at all of which points  $\psi = 0$ . Therefore  $B$  is the empty set. ■

In the  $\mu = 0$  case it is easy to show how Theorem 6 follows from Lemma 8.

**Proof of Theorem 6 when  $\lambda = \sum_{j=1}^k z_j$ .** The proof when  $\lambda < \sum z_j$  uses the same ideas but is more complicated. Since  $\lambda = \sum z_j$  then  $\lambda_1 = \sum_{j=1}^m z_j$ ,  $\lambda_2 = \sum_{j=m+1}^k z_j$  and  $\mu = 0$  for all three systems. For  $\alpha > 0$  let  $f(\alpha) = E^{\text{TF}}(\alpha z_1, \dots, \alpha z_m, z_{m+1}, \dots, z_k; R_1, \dots, R_k) - E^{\text{TF}}(\alpha z_1, \dots, \alpha z_m; R_1, \dots, R_m) - E^{\text{TF}}(z_{m+1}, \dots, z_k; R_{m+1}, \dots, R_k)$ , where the three  $E^{\text{TF}}$  are defined for neutral systems (i.e.,  $\mu = 0$  for all  $\alpha$ ). The goal is to show that  $f(1) \geq 0$ . Since  $f(0) = 0$ , it is enough to show that  $df(\alpha)/d\alpha \geq 0$ . From (26) and (27) it is true, and almost obvious, that

$$\begin{aligned} \frac{dE^{\text{TF}}}{d\alpha} &= - \int \rho^{\text{TF}}(y) |y - R_i|^{-1} dy + \sum_{j=1}^m z_j |R_i - R_j|^{-1} \\ &= \lim_{x \rightarrow R_i} \phi^{\text{TF}}(x) - z_i |x - R_i|^{-1}. \end{aligned}$$

This is the TF version of the Feynman-Hellmann theorem; notice how the nuclear-nuclear repulsion comes in here. Thus,

$$\frac{df(\alpha)}{d\alpha} = \sum_{i=1}^m \lim_{x \rightarrow R_i} z_i \eta_{\alpha}(x),$$

where  $\eta_{\alpha}(x) = \phi_1^{\text{TF}}(x) - \phi_2^{\text{TF}}(x)$  and  $\phi_1^{\text{TF}}$  is the potential for  $\{\alpha z_1, \dots, \alpha z_m, z_{m+1}, \dots, z_k; R_1, \dots, R_k\}$  and  $\phi_2^{\text{TF}}$  is the potential for  $\{\alpha z_1, \dots, \alpha z_m; R_1, \dots, R_m\}$ .  $\phi_1^{\text{TF}}(x) \geq \phi_2^{\text{TF}}(x)$  for all  $x$  by Lemma 8, and hence  $\eta_{\alpha}(x) \geq 0$ . ■

Theorem 6 has a natural application to the stability of matter problem. As will be shown in the next section, the TF energy (27) is, with suitably modified constants, a lower bound to the true quantum energy  $E_N^0$  for *all*  $Z$ . By Theorem 3 (iv) and Theorem 6 we have the following theorem.

**Theorem 9.** Fix  $\{z_j, R_j\}_{j=1}^k$  and let  $Z = \sum_{j=1}^k z_j$ . Then for all  $\lambda \geq 0$

$$E_{\lambda}^{\text{TF}} \geq E_Z^{\text{TF}} \geq -(2.21) q^{2/3} (K^c)^{-1} \sum_{j=1}^k z_j^{7/3}. \quad (43)$$

The latter constant, 2.21, is obtained by numerically solving the TF equation for a single, neutral atom (J. F. Barnes, private communication). By scaling, Eq. (43)

holds for an choice of  $K^c$  in the definition (26) of  $\mathcal{G}(\rho)$ .

Theorem 9 is what will be needed for the  $H$  stability of matter because it says that the TF system is  $H$  stable, i.e., the energy is bounded below by a constant times the nuclear particle number (assuming that the  $z_j$  are bounded, of course).

Another application of Theorem 6 that will be needed is the following strange *inversion of the role of electrons and nuclei* in TF theory. It will enable us to give a lower bound to the true quantum-mechanical *electron-electron* repulsion. This theorem has nothing to do with quantum mechanics per se; it is really a theorem purely about electrostatics even though it is derived from the TF no binding theorem.

**Theorem 10.** Suppose that  $x_1, \dots, x_N$  are any  $N$  distinct points in 3-space and define

$$V_X(y) = \sum_{j=1}^N |y - x_j|^{-1}. \quad (44)$$

Let  $\gamma > 0$  and let  $\rho(x)$  be any non-negative function such that  $\int \rho(x) dx < \infty$  and  $\int \rho(x)^{5/3} dx < \infty$ . Then

$$\begin{aligned} \sum_{1 \leq i < j \leq N} |x_i - x_j|^{-1} &\geq -\frac{1}{2} \int \int \rho(x) |x - y|^{-1} \rho(y) dx dy \\ &\quad + \int \rho(y) V_X(y) dy - (2.21)N/\gamma \\ &\quad - \gamma \int \rho(y)^{5/3} dy. \end{aligned} \quad (45)$$

*Proof.* Consider  $\mathcal{G}(\rho)$  (26) with  $q=1$ ,  $k=N$ ,  $K^c$  replaced by  $\gamma$ ,  $z_j \equiv 1$  and  $R_j \equiv x_j$ ,  $j=1, \dots, N$ . Let  $\lambda = \int \rho(x) dx$ . Then  $\mathcal{G}(\rho) \geq E_{\lambda}^{\text{TF}}$  (by definition) and  $E_{\lambda}^{\text{TF}} \geq -(2.21)N/\gamma$  by Theorem 9. The difference of the two sides in Eq. (45) is just  $\mathcal{G}(\rho) + (2.21)N/\gamma$ . ■

#### IV. THE STABILITY OF BULK MATTER

The various results of the last two sections can now be assembled to prove that the ground state energy (or infimum of the spectrum, if this not an eigenvalue) of  $H_N$  is bounded below by an extensive quantity, namely the total number of particles, independent of the nuclear locations  $\{R_j\}$ . This is called the  $H$  stability of matter to distinguish it from thermodynamic stability introduced in the next section. As explained before, the inclusion of the nuclear kinetic energy, as will be done in the next section, can only raise the energy.

The first proof of the  $N$  boundedness of the energy was given by Dyson and Lenard (Dyson and Lenard, 1967, Lenard and Dyson, 1968). Their proof is a remarkable analytic tour de force, but a chain of sufficiently many inequalities was used that they ended up with an estimate of something like  $-10^{14}$  Ry/particle. Using the results of the previous sections we will end up with  $-23$  Ry/particle [see Eq. (55)].

We have in mind, of course, that the nuclear charges  $z_j$ , if they are not all the same, are bounded above by some fixed charge  $z$ .

Take any fermion  $\psi(x_1, \dots, x_N; \sigma_1, \dots, \sigma_N)$  which is normalized and antisymmetric in the  $(x_i, \sigma_i)$ . Define the kinetic energy  $T_\psi$  and the single particle density  $\rho_\psi$  as in (16) and (17). We wish to compute a lower bound to

$$E_\psi^Q \equiv \langle \psi, H_N \psi \rangle \quad (46)$$

with  $H_N$  being the  $N$ -particle Hamiltonian given in (23) and  $\langle \psi, \psi \rangle = 1$ .

For the third term on the right side of (23) Theorem 10 can be used with  $\rho$  taken to be  $\rho_\psi$ . Then, for any  $\gamma > 0$

$$\begin{aligned} \left\langle \psi, \sum_{1 \leq i < j \leq N} |x_i - x_j|^{-1} \psi \right\rangle &\geq \frac{1}{2} \int \int \rho_\psi(x) |x - y|^{-1} \rho_\psi(y) dx dy \\ &\quad - (2.21)N\gamma^{-1} - \gamma \int \rho_\psi(y)^{5/3} dy. \end{aligned} \quad (47)$$

Notice how the first and second terms on the right side of (45) combine to give  $+\frac{1}{2}$  since

$$\left\langle \psi, \left\{ \int \rho_\psi(y) V_X(y) dy \right\} \psi \right\rangle = \int \int \rho_\psi(x) |x - y|^{-1} \rho_\psi(y) dx dy. \quad (48)$$

To control the kinetic energy in (23) Theorem 1 is used; the total result is then

$$\begin{aligned} E_\psi^Q &\geq \alpha \int \rho_\psi(x)^{5/3} dx - \int V(x) \rho_\psi(x) dx \\ &\quad + \frac{1}{2} \int \int \rho_\psi(x) |x - y|^{-1} \rho_\psi(y) dx dy \\ &\quad + U(\{z_j, R_j\}_{j=1}^k) - (2.21)N\gamma^{-1} \end{aligned} \quad (49)$$

with

$$\alpha = (4\pi q)^{-2/3} K^c - \gamma. \quad (50)$$

Restrict  $\gamma$ , which was arbitrary, so that  $\alpha > 0$ . Then, apart from the constant term  $-(2.21)N\gamma^{-1}$ , Eq. (49) is just  $\mathcal{G}_\alpha(\rho_\psi)$ , the Thomas-Fermi energy functional  $\mathcal{G}$  applied to  $\rho_\psi$ , but with  $q^{-2/3}K^c$  replaced by  $\alpha$ . Since  $\mathcal{G}_\alpha(\rho_\psi) \geq E_{\alpha, N}^{\text{TF}} \equiv \inf\{\mathcal{G}_\alpha(\rho) : \int \rho = N\}$  (by definition), and since the neutral case always has the lowest TF energy, as shown in Theorem 9, we have that

$$\mathcal{G}_\alpha(\rho_\psi) \geq -(2.21)\alpha^{-1} \sum_{j=1}^k z_j^{7/3}. \quad (51)$$

Thus we have proved the following:

**Theorem 11.** If  $\psi$  is a normalized, antisymmetric function of space and spin of  $N$  variables, and if there are  $q$  spin states associated with each particle then, for any  $\gamma > 0$  such that  $\alpha$  defined by Eq. (50) is positive,

$$\langle \psi, H_N \psi \rangle \geq -(2.21) \left\{ N\gamma^{-1} + \alpha^{-1} \sum_{j=1}^k z_j^{7/3} \right\}. \quad (52)$$

The optimum choice for  $\gamma$  is

$$\gamma = (4\pi q)^{-2/3} K^c \left[ \left( \sum_{j=1}^k \frac{z_j^{7/3}}{N} \right)^{1/2} + 1 \right]^{-1}$$

in which case

$$E_N^Q \geq -(2.21) \frac{(4\pi q)^{2/3} N}{K^c} \left\{ 1 + \left[ \sum_{j=1}^k \frac{z_j^{7/3}}{N} \right]^{1/2} \right\}^2. \quad (53)$$

This is the desired result, but some additional remarks are in order.

(1) Since  $[1 + a^{1/2}]^2 \leq 2 + 2a$ ,

$$E_N^Q \geq -(4.42)(4\pi q)^{2/3} (K^c)^{-1} \left\{ N + \sum_{j=1}^k z_j^{7/3} \right\}. \quad (54)$$

Thus, provided the nuclear charges  $z_j$  are bounded above by some fixed  $z$ ,  $E_N^Q$  is indeed bounded below by a constant times the total particle number  $N+k$ .

(2) Theorem 11 does not presuppose neutrality.

(3) For electrons,  $q=2$  and the prefactor in Eq. (53) is  $-(2.08)N$ . As remarked after Theorem 1, the unwanted constant  $(4\pi)^{2/3}$  has been improved to  $[4\pi/(1.83)]^{2/3}$ . Using this, the prefactor becomes  $-(1.39)N$ . If  $z_j=1$  (hydrogen atoms) and  $N=k$  (neutrality) then

$$E_N^Q \geq -(5.56)N = -(22.24)N \text{ Ry}. \quad (55)$$

(4) The power law  $z^{7/3}$  cannot be improved upon for large  $z$  because Theorem 5 asserts that the energy of an atom is indeed proportional to  $z^{7/3}$  for large  $z$ .

(5) It is also possible to show that matter is indeed bulky. This will be proved for any  $\psi$  and any nuclear configuration (not just the minimum energy configuration) for which  $E_\psi^Q \leq 0$ . The minimizing nuclear configuration is, of course, included in this hypothesis. Then

$$0 \geq E_\psi^Q = \frac{1}{2} T_\psi + \langle \psi, H'_N \psi \rangle,$$

where  $H'_N$  is Eq. (23) but with a factor  $\frac{1}{2}$  multiplying  $\sum_{i=1}^N \Delta_i$ . By Theorem 11,  $\langle \psi, H'_N \psi \rangle \geq 2E_N$ , where  $E_N$  is the right side of Eq. (53) (replace  $K^c$  by  $K^c/2$  there). Therefore, the first important fact is that

$$T_\psi \leq 4|E_N|,$$

and this is bounded above by the total particle number.

Next, for any  $p \geq 0$ , it is easy to check that there is a  $C_p > 0$  such that for any non-negative  $\rho(x)$ ,

$$\left\{ \int \rho(x)^{5/3} dx \right\}^{p/2} \int |x|^p \rho(x) dx \geq C_p \left\{ \int \rho(x) dx \right\}^{1+5p/6}.$$

It is easy to find a minimizing  $\rho$  for this and to calculate  $C_p$ :  $\rho(x)^{2/3} = 1 - |x|^p$  for  $|x| \leq 1$ ;  $\rho(x) = 0$ , otherwise.

Since  $T_\psi$  satisfies Eq. (18) we have that

$$\left\langle \psi, \sum_{i=1}^N |x_i|^p \psi \right\rangle = \int |x|^p \rho_\psi(x) dx \geq C'_p N (N^{5/3}/|E_N|)^{p/2},$$

with  $C'_p = C_p (K^c/4)^{p/2} (4\pi q)^{-p/3}$ .

If it is assumed that  $\sum z_j^{7/3}/N$  is bounded, and hence that  $(N^{5/3}/|E_N|)^{p/2} > AN^{p/3}$  for some  $A$ , we reach the conclusion that the radius of the system is at least of the order  $N^{1/3}$ , as it should be.

The above analysis did not use any specific property of the Coulomb potential, such as the virial theorem. It is also applicable to the more general Hamiltonian  $H_{n,k}$  in Eq. (58).

(6) The  $q$  dependence was purposely retained in Eq. (53) in order to say something about *bosons*. If  $q=N$ , then it is easy to see that the requirement of antisymmetry in  $\psi$  is *no restriction* at all. In this case then, one has simply

$$E_N^Q = \inf \text{spec } H_N$$

over all of  $L^2(\mathbb{R}^3)^N$ . Therefore

$$E_N^Q(\text{bosons}) \geq \frac{-(2.21)(4\pi)^{2/3}}{K^c} N^{5/3} \left\{ 1 + \left[ \sum_{j=1}^k \frac{z_j^{7/3}}{N} \right]^{1/2} \right\}^2. \quad (56)$$

It was shown by Dyson and Lenard (Dyson and Lenard, 1967) that

$$E_N^Q(\text{bosons}) \geq -(\text{const})N^{5/3},$$

and by Dyson (Dyson, 1967) that

$$E_N^Q(\text{bosons}) \leq -(\text{const})N^{7/5}. \quad (57)$$

Proving Eq. (57) was not easy. Dyson had to construct a rather complicated variational function related to the type used in the BCS theory of superconductivity. Therefore *bosons are not stable* under the action of Coulomb forces, but the exact power law is not yet known. Dyson has conjectured that it is  $\frac{7}{5}$ .

In any event, the essential point has been made that Fermi statistics is essential for the stability of matter. The uncertainty principle for one particle, even in the strong form (5), together with intuitive notions that the electrostatic energy ought not to be very great, are insufficient for stability. The additional physical fact that is needed is that the kinetic energy increases as the  $\frac{5}{3}$  power of the fermion density.

## V. THE THERMODYNAMIC LIMIT

Having established that  $E_N^Q$  is bounded below by the total particle number, the next question to consider is whether, under appropriate conditions,  $E_N^Q/N$  has a limit as  $N \rightarrow \infty$ , as expected. More generally, the same question can be asked about the free energy per particle when the temperature is not zero and the particles are confined to a box.

It should be appreciated that the difficulty in obtaining the lower bound to  $E_N^Q$  came almost entirely from the  $r^{-1}$  short range singularity of the Coulomb potential. Other potentials, such as the Yukawa potential, with the same singularity would present the same difficulty which would be resolved in the same way. The singularity was tamed by the  $\rho^{5/3}$  behavior of the fermion kinetic energy.

The difficulty for the thermodynamic limit is different. It is caused by the *long range*  $r^{-1}$  behavior of the Coulomb potential. In other words, we are faced with the problem of explosion rather than implosion. Normally, a potential that falls off with distance more slowly than  $r^{-3-\epsilon}$  for some  $\epsilon > 0$  does *not* have a thermodynamic limit. Because the charges have different signs, however, there is hope that a cancellation at large distances may occur.

An additional physical hypothesis will be needed, namely *neutrality*. To appreciate the importance of neutrality consider the case that the electrons have positive, instead of negative charge. Then  $E_N^Q > 0$  because every term in Eq. (23) would be positive. While the  $H$ -stability question is trivial in this case, the thermodynamic limit is not. If the particles are constrained to be in a domain  $\Omega$  whose volume  $|\Omega|$  is proportional to  $N$ , the particles will repel each other so strongly that they will all go to the boundary of  $\Omega$  in order to mini-

mize the electrostatic energy. The minimum electrostatic energy will be of the order  $+N^2|\Omega|^{-1/3} \sim +N^{5/3}$ . Hence no thermodynamic limit will exist.

When the system is neutral, however, the energy can be expected to be extensive, i.e.,  $O(N)$ . For this to be so, different parts of the system far from each other must be approximately independent, despite the long range nature of the Coulomb force. The fundamental physical, or rather electrostatic, fact that underlies this is *screening*; the distribution of the particles must be sufficiently neutral and isotropic locally so that according to Newton's theorem (13 below) the electric potential far away will be zero. The problem is to express this idea in precise mathematical form.

We begin by defining the Hamiltonian for the *entire* system consisting of  $k$  nuclei, each of charge  $z$  and mass  $M$ , and  $n$  electrons ( $\hbar^2/2=1$ ,  $m=1$ ,  $|e|=1$ ):

$$H_{n,k} = - \sum_{j=1}^n \Delta_j - \frac{1}{M} \sum_{j=n+1}^{n+k} \Delta_j - z \sum_{i=1}^n \sum_{j=n+1}^{n+k} |x_i - y_j|^{-1} \\ + \sum_{1 \leq i < j \leq n} |x_i - x_j|^{-1} + z^2 \sum_{n+1 \leq i < j \leq n+k} |y_i - y_j|^{-1}. \quad (58)$$

The first and second terms in Eq. (58) are, respectively, the kinetic energies of the electrons and the nuclei. The last three terms are, respectively, the electron-nuclear, electron-electron, and nuclear-nuclear Coulomb interactions. The electron coordinates are  $x_i$  and the nuclear coordinates are  $y_i$ . The electrons are fermions with spin  $\frac{1}{2}$ ; the nuclei may be either bosons or fermions.

The basic neutrality hypotheses is that  $n$  and  $k$  are related by

$$n = kz. \quad (59)$$

It is assumed that  $z$  is rational.

The thermodynamic limit to be discussed here can be proved under more general assumptions, i.e., we can have several kinds of negative particles (but they must all be fermions in order that the basic stability estimate of Sec. IV holds) and several kinds of nuclei with different statistics, charges, and masses. Neutrality must always hold, however. Short range forces and hard cores, in addition to the Coulomb forces, can also be included with a considerable sacrifice in simplicity of the proof. See (Lieb and Lebowitz, 1972).

$H_{n,k}$  acts on square integrable functions of  $n+k$  variables (and spin as well). To complete the definition of  $H_{n,k}$  we must specify boundary conditions: choose a domain  $\Omega$  (an open set, which need not be connected) and require that  $\psi=0$  if  $x_i$  or  $y_i$  are on the boundary of  $\Omega$ .

For each non-negative integer  $j$ , choose an  $n_j$  and a corresponding  $k_j$  determined by Eq. (59), and choose a domain  $\Omega_j$ . The symbol  $N_j$  will henceforth stand for the pair  $(n_j, k_j)$  and

$$|N_j| \equiv n_j + k_j.$$

We require that the densities

$$\rho_j \equiv |N_j| |\Omega_j|^{-1} \quad (60)$$

be such that

$$\lim_{j \rightarrow \infty} \rho_j = \rho. \quad (61)$$

$\rho$  is then the density in the thermodynamic limit. Here we shall choose the  $\Omega_j$  to be a sequence of *balls* of radii  $R_j$  and shall denote them by  $B_j$ .

It can be shown that the *same* thermodynamic limit for the energy and free energy holds for any sequence  $N_j, \Omega_j$  and depends *only* on the limiting  $\rho$  and  $\beta$ , and not on the "shape" of the  $\Omega_j$ , provided the  $\Omega_j$  go to infinity in some reasonable way.

The basic quantity of interest is the *canonical partition function*

$$Z(N, \Omega, \beta) = \text{Tr} \exp(-\beta H_{n,k}), \quad (62)$$

where the trace is on  $L^2(\Omega)^{|N|}$  and  $\beta = 1/T$ ,  $T$  being the temperature in units in which Boltzmann's constant is unity.

The *free energy per unit volume* is

$$F(N, \Omega, \beta) = -\beta^{-1} \ln Z(N, \Omega, \beta) / |\Omega| \quad (63)$$

and the problem is to show that with

$$F_j = F(N_j, \Omega_j, \beta), \quad (64)$$

then

$$\lim_{j \rightarrow \infty} F_j \equiv F(\rho, \beta) \quad (65)$$

exists. A similar problem is to show that

$$E(N, \Omega) \equiv |\Omega|^{-1} \inf_{\psi} \langle \psi, H_{n,k} \psi \rangle / \langle \psi, \psi \rangle, \quad (66)$$

the *ground state energy per unit volume*, has a limit

$$e(\rho) = \lim_{j \rightarrow \infty} E_j, \quad (67)$$

where

$$E_j = E(N_j, \Omega_j).$$

The proof we will give for the limit  $F(\rho, \beta)$  will hold equally well for  $e(\rho)$  because  $E_j$  can be substituted for  $F_j$  in all statements.

The basic strategy consists of two parts. The easiest part is to show that  $F_j$  is bounded below. We already know this for  $E_j$  by the results of Sec. IV. The second step is to show that in some sense the sequence  $F_j$  is decreasing. This will then imply the existence of a limit.

**Theorem 12.** Given  $N$ ,  $\Omega$ , and  $\beta$  there exists a constant  $C$  depending only on  $\rho = |N|/|\Omega|$  and  $\beta$  such that

$$F(N, \Omega, \beta) \geq C. \quad (68)$$

*Proof.* Write  $H_{n,k} = H_A + H_B$ , where

$$H_A = -\frac{1}{2} \left\{ \sum_{i=1}^n \Delta_i + \frac{1}{M} \sum_{j=n+1}^{n+k} \Delta_j \right\}$$

is half the kinetic energy. Then  $H_B \geq b|N|$ , with  $b$  depending only on  $z$ , by the results of Sec. IV (increasing the mass by a factor of 2 in  $H_B$  only changes the constant  $b$ ). Hence  $Z(N, \Omega, \beta) \leq \exp(-\beta b|N|) \text{Tr} \exp(-\beta H_A)$ . However,  $\text{Tr} \exp(-\beta H_A)$  is the partition function of an ideal gas and it is known by explicit computation that it is bounded above by  $e^{\beta d|N|}$  with  $d$  depending only on  $\rho = |N|/|\Omega|$  and  $\beta$ . Thus

$$F(N, \Omega, \beta) \geq (b-d)\rho. \quad \blacksquare$$

For the second step, two elementary but basic in-

equalities used in the general theory of the thermodynamic limit are needed and they will be described next.

A. *Domain partition inequality*: Given a domain  $\Omega$  and the particle numbers  $N = (n, k)$ , let  $\pi$  be a partition of  $\Omega$  into  $l$  disjoint domains  $\Omega^1, \dots, \Omega^l$ . Likewise  $N$  is partitioned into  $l$  integral parts (some of which may be zero):

$$N = N^1 + \dots + N^l.$$

Then for any such partition,  $\pi$ , of  $\Omega$  and  $N$

$$Z(N, \Omega, \beta) = \text{Tr} \exp(-\beta H_{n,k}) \geq \text{Tr}^\pi \exp(-\beta H_N^\pi). \quad (69)$$

Here  $\text{Tr}^\pi$  means trace over

$$\mathcal{H}^\pi \equiv L^2(\Omega^1)^{|N^1|} \otimes \dots \otimes L^2(\Omega^l)^{|N^l|},$$

and  $H_N^\pi$  is defined as in (58) but with Dirichlet ( $\psi=0$ ) boundary conditions for the  $N^i$  particles on the boundary of  $\Omega^i$  (for  $i=1, \dots, l$ ).

Simply stated, the first  $N^1$  particles are confined to  $\Omega^1$ , the second  $N^2$  to  $\Omega^2$ , etc. The interaction among the particles in different domains is still present in  $H_N^\pi$ . Equation (69) can be proved by the Peierls–Bogoliubov variational principle for  $\text{Tr} e^X$ . Alternatively, (69) can be viewed simply as the statement that the insertion of a hard wall, infinite potential on the boundaries of the  $\Omega^i$  only decreases  $Z$ ; the further restriction of a definite particle number to each  $\Omega^i$  further reduces  $Z$  because it means that the trace is then over only the  $H_N^\pi$ -invariant subspace,  $\mathcal{H}^\pi$ , of the full Hilbert space.

B. *Inequality for the interdomain interaction*: The second inequality is another consequence of the convexity of  $A \rightarrow \text{Tr} e^A$  (Peierls–Bogoliubov inequality):

$$\text{Tr} e^{A+B} \geq \text{Tr} e^A \exp \langle B \rangle, \quad (70)$$

where

$$\langle B \rangle \equiv \text{Tr} B e^A / \text{Tr} e^A. \quad (71)$$

Some technical conditions are needed here, but Eqs. (70) and (71) will hold in our application.

To exploit (70), first make the same partition  $\pi$  as in inequality A and then write

$$H_N^\pi = H_0 + W(X), \quad (72)$$

$$H_0 = H^1 + \dots + H^l, \quad (73)$$

with  $H^i$  being that part of the total Hamiltonian (58) involving only the  $N^i$  particles in  $\Omega^i$ , and  $H^i$  is defined with the stated Dirichlet boundary conditions on the boundary of  $\Omega^i$ .  $W(X)$ , with  $X$  standing for all the coordinates, is the interdomain Coulomb interaction. In other words,  $W(X)$  is that part of the last three terms on the right side of (58) which involves coordinates in different blocks of the partition  $\pi$ . Technically,  $W$  is a small perturbation of  $H_0$ .

With

$$A = -\beta H_0 \text{ and } B = -\beta W \quad (74)$$

in (70), we must calculate  $\langle W \rangle$ . Since  $e^A = e^{-\beta H_0}$  is a simple tensor product of operators on each  $L^2(\Omega^i)^{|N^i|}$ ,  $W$  is merely the *average* interdomain Coulomb energy in a canonical ensemble in which the Coulomb interaction is present in each subdomain but the  $l$  domains are independent of each other. In other words, let  $q^i(x)$ ,

$x \in \Omega^i$  denote the average *charge density* in  $\Omega^i$  for this ensemble of independent domains, namely

$$q^i(x) = \sum_{j=1}^{|N^i|} q_j \int_{|\Omega^i|^{-1}} \exp(-\beta H^i)(X^i, X^i) \hat{d}x_j / Z(N^i, \Omega^i, \beta) \quad (75)$$

with the following notation:  $X^i$  stands for the coordinates of the  $|N^i|$  particles in  $\Omega^i$ ,  $\hat{d}x_j$  means integration over all these coordinates (in  $\Omega^i$ ) with the exception of  $x_j$ , and  $x_j$  is set equal to  $x$ ;  $q_j$  is the charge ( $-1$  or  $+z$ ) of the  $j$ th particle;  $\exp(-\beta H^i)(X^i, Y^i)$  is a kernel ( $x$ -space representation) for  $\exp(-\beta H^i)$ .  $q^i(x)$  vanishes if  $x \notin \Omega^i$ .

With the definitions (75) one has that

$$\langle W \rangle = \sum_{i < j} \int_{\Omega^i} \int_{\Omega^j} q^i(x) q^j(y) |x - y|^{-1} dx dy. \quad (76)$$

Equation (70), together with (76) and (74), is the desired inequality for the interdomain interaction. It is quite general in that an analogous inequality holds for arbitrary two-body potentials. Neither specific properties of the Coulomb potential nor neutrality was used.

Now we come to the crucial point at which screening is brought in. The following venerable result from the *Principia Mathematica* is essential.

*Theorem 13 (Newton)*. Let  $\rho(x)$  be an integrable function on 3-space such that  $\rho(x) = \rho(y)$  if  $|x| = |y|$  (isotropy) and  $\rho(x) = 0$  if  $|x| > R$  for some  $R > 0$ . Let

$$\phi(x) = \int \rho(y) |x - y|^{-1} dy \quad (77)$$

be the Coulomb potential generated by  $\rho$ . Then if  $|x| \geq R$

$$\phi(x) = |x|^{-1} \int \rho(y) dy. \quad (78)$$

The important point is that an isotropic, *neutral* charge distribution generates zero potential outside its support, irrespective of how the charge is distributed radially.

Suppose that  $N^i$  is *neutral*, i.e., the electron number  $= z$  times the nucleon number for each subdomain in  $\Omega$ . Suppose also that the subdomain  $\Omega^i$  is a *ball* of radius  $R^i$  centered at  $a^i$ . Then since  $H^i$  is rotation invariant,  $q^i(x) = q^i(y)$  if  $|x - a^i| = |y - a^i|$ ,  $\int q^i(x) dx = 0$  (by neutrality) and  $q^i(x) = 0$  if  $|x - a^i| > R^i$ . Then, by Theorem 13, *every* term in Eq. (76) involving  $q^i$  vanishes, because when  $j \neq i$ ,  $q^j(y) = 0$  if  $|y - a^j| < R^j$  since  $\Omega^j$  is disjoint from  $\Omega^i$ . Consequently the average interdomain interaction,  $\langle W \rangle$ , *vanishes*.

In the decomposition,  $\pi$ , of  $\Omega$  into  $\Omega^1, \dots, \Omega^l$  and  $N$  into  $N^1, \dots, N^l$  we will arrange matters such that

- (i)  $\Omega^1, \dots, \Omega^{l-1}$  are balls,
- (ii)  $N^1, \dots, N^{l-1}$  are neutral,
- (iii)  $N^l = 0$ .

Then  $\langle W \rangle = 0$  and, using Eqs. (69) and (70)

$$Z(N, \Omega, \beta) \geq \text{Tr}^\pi \exp(-\beta H_N^\pi) \geq \prod_{i=1}^{l-1} Z(N^i, \Omega^i, \beta) e^{-\beta \langle W \rangle} = \prod_{i=1}^{l-1} Z(N^i, \Omega^i, \beta). \quad (79)$$

In addition to (i), (ii), (iii) it will also be necessary to arrange matters such that when  $\Omega$  is a ball  $B_K$  in the chosen sequence of domains, then the subdomains  $\Omega^1, \dots, \Omega^{l-1}$  in the partition of  $B_K$  are also smaller balls in the same sequence. With these requirements in mind the *standard sequence*, which depends on the limiting density  $\rho$ , is defined as follows:

- (1) Choose  $\rho > 0$ .
- (2) Choose any  $N_0$  satisfying the neutrality condition (59).
- (3) Choose  $R_0$  such that

$$28(4\pi/3)\rho R_0^3 = |N_0|. \quad (80)$$

- (4) For  $j \geq 1$  let

$$\begin{aligned} R_j &= (28)^j R_0, \\ N_j &= (28)^{3j-1} N_0 \end{aligned} \quad (81)$$

be the radius of the ball  $B_j$  and the particle number in that ball.

It will be noted that the density in all the balls except the first is

$$\rho_j = \rho, \quad j \geq 1, \quad (82)$$

while the density in the smallest ball is much bigger:

$$\rho_0 = 28\rho. \quad (83)$$

This has been done so that when a ball  $B_K, K \geq 1$  is packed with smaller balls in the manner to be described below, the density in each ball will come out right; the higher density in  $B_0$  compensates for the portion of  $B_K$  not covered by smaller balls. The radii increase geometrically, namely by a factor of 28.

The number 28 may be surprising until it is realized that the objective is to be able to pack  $B_K$  with balls of type  $B_{K-1}, B_{K-2}$ , etc., in such a way that as much as possible of  $B_K$  is covered and also that very little of  $B_K$  is covered by very small balls. If the ratio of radii were too close to unity, then the packing of  $B_K$  would be inefficient from this point of view. In short, if the number 28 is replaced by a much smaller number the analog of the following basic geometric theorem will not be true.

**Theorem 14 (Cheese theorem).** For  $j$  a positive integer define the integer  $m_j \equiv (27)^{j-1}(28)^{2j}$ . Then for each positive integer  $K \geq 1$  it is possible to pack the ball  $B_K$  of radius  $R_K$  (given by 81) with

$$\bigcup_{j=0}^{K-1} (m_{K-j} \text{ balls of radius } R_j).$$

"Pack" means that all the balls in the union are disjoint.

We will not give a proof of Theorem 14 here, but note that it entails showing that  $m_1$  balls of radius  $R_{K-1}$  can be packed in  $B_K$  in a cubic array, then that  $m_2$  balls of radius  $R_{K-2}$  can be packed in a cubic array in the interstitial region, etc.

Theorem 14 states that  $B_K$  can be packed with  $(28)^2$  balls of type  $B_{K-1}$ ,  $(27)(28)^4$  balls of type  $B_{K-2}$ , etc. If  $f_{K-j}$  is the fraction of the volume of  $B_K$  occupied by all the balls of radius  $R_j$  in the packing, then

$$f_j = m_j (R_{K-j}/R_K)^3 = \frac{1}{27} \gamma^j \quad (84)$$

with

$$\gamma = \frac{27}{28} < 1. \quad (85)$$

The packing is *asymptotically complete* in the sense that

$$\lim_{K \rightarrow \infty} \sum_{j=0}^{K-1} f_{K-j} = \frac{1}{27} \sum_{j=1}^{\infty} \gamma^j = 1. \quad (86)$$

It is also "geometrically rapid" because the fraction of  $|B_K|$  that is uncovered is

$$\sum_{j=K+1}^{\infty} f_j = \gamma^K. \quad (87)$$

The necessary ingredients having been assembled, we can now prove the following theorem.

**Theorem 15.** Given  $\rho$  and  $\beta > 0$ , the thermodynamic limits  $F(\rho, \beta)$  and  $e(\rho)$  (65, 67) exist for the sequence of balls and particle numbers specified by (80) and (81).

*Proof.* Let  $F_K$  given by Eq. (64) be the free energy per unit volume for the ball  $B_K$  with  $N_K$  particles in it. For  $K \geq 1$ , partition  $B_K$  into disjoint domains  $\Omega^1, \dots, \Omega^l$ , where the  $\Omega^i$  for  $i = 1, \dots, l-1$  designate the smaller balls referred to in Theorem 14, and  $\Omega^l$  (which is the "cheese" after the holes have been removed) is the remainder of  $B_K$ . The smaller balls are copies of  $B_j$ ,  $0 \leq j \leq K-1$ ; in each of these place  $N_j$  particles according to (81).  $N^l = 0$ . The total particle number in  $B_K$  is then

$$\begin{aligned} \sum_{j=0}^{K-1} N_j m_{K-j} &= N_0 \{ (27)^{K-1} (28)^{2K} + \sum_{j=1}^{K-1} (28)^{3j-1} (27)^{K-j-1} (28)^{2K-2j} \} \\ &= N_0 (28)^{3K-1} = N_K \end{aligned}$$

as it should be.

Use the basic inequality (79);  $\langle W \rangle = 0$  since all the smaller balls are neutral and  $\Omega^l$  contains no particles. Thus, taking logarithms and dividing by  $|B_K|$ , we have for  $K \geq 1$  that

$$F_K \leq \sum_{j=0}^{K-1} F_j f_{K-j} \quad (88)$$

with  $f_j = \gamma^j/27$  and  $\gamma = \frac{27}{28}$ . This inequality can be rewritten as

$$F_K = \sum_{j=0}^{K-1} \frac{F_j \gamma^{K-j}}{27} - d_K \quad (89)$$

with  $d_K \geq 0$ . Equation (89) is a *renewal equation* which can be solved explicitly by inspection:

$$F_K = -\gamma d_K - \sum_{j=1}^K \frac{d_j}{28} + \frac{F_0}{28}. \quad (90)$$

We now use the first step, Theorem 13, on the boundedness of  $F_K$ . Since  $F_K \geq C$ ,  $\sum_{j=1}^{\infty} d_j$  must be finite, for otherwise (90) would say that  $F_K \rightarrow -\infty$ . The convergence of the sum implies that  $d_K \rightarrow 0$  as  $K \rightarrow \infty$ . Hence the limit exists; specifically

$$F = \lim_{K \rightarrow \infty} F_K = -\sum_{j=1}^{\infty} \frac{d_j}{28} + \frac{F_0}{28}. \quad \blacksquare \quad (91)$$

Theorem 15 is the desired goal, namely the existence of the thermodynamic limit for the free energy (or ground state energy) per unit volume. There are, how-

ever, some additional points that deserve comment.

(A) For each given limiting density  $\rho$ , a particular sequence of domains, namely balls, and particle numbers was used. It can be shown that the same limit is reached for general domains, with some mild conditions on their shape including, of course, balls of different radii than that used here. The argument involves packing the given domains with balls of the standard sequence and vice versa. The proof is tedious, but standard, and can be found in (Lieb and Lebowitz, 1972).

(B) Here we have considered the thermodynamic limit for real matter, in which all the particles are mobile. There are, however, other models of some physical interest. One is *jellium* in which the positive nuclei are replaced by a *fixed, uniform background* of positive charge. With the aid of an additional trick the thermodynamic limit can also be proved for this model (Lieb and Narnhofer, 1975). Another, more important model is one in which the nuclei are *fixed point charges* arranged periodically in a lattice. This is the model of solid state physics. Unfortunately, local rotation invariance is lost and Newton's Theorem 13 cannot be used. This problem is still open and its solution will require a deeper insight into screening.

(C) An absolute physical requirement for  $\beta F(\rho, \beta)$ , as a function of  $\beta = 1/T$ , is that it be *concave*. This is equivalent to the fact that the specific heat is non-negative since (specific heat)  $= -\beta^2 \partial^2 \beta F(\rho, \beta) / \partial \beta^2$ . Fortunately it is true. From the definitions (57), (58) we see that  $\ln Z(N, \Omega, \beta)$  is convex in  $\beta$  for *every* finite system and hence  $\beta F(N, \Omega, \beta)$  is concave. Since the limit of a sequence of concave functions is always concave, the limit  $\beta F(\rho, \beta)$  is concave in  $\beta$ .

(D) Another absolute requirement is that  $F(\rho, \beta)$  be *convex* as a function of  $\rho$ . This is called *thermodynamic stability* as distinct from the lower bound  $H$  stability of the previous sections. It is equivalent to the fact that the compressibility is non-negative, since (compressibility) $^{-1} = \partial P / \partial \rho = \rho \partial^2 F(\rho, \beta) / \partial \rho^2$ . Frequently, in approximate theories (e.g., van der Waals' theory of the vapor-liquid transition, some field theories, or some theories of magnetic systems in which the magnetization per unit volume plays the role of  $\rho$ ), one introduces an  $F$  with a double bump. Such an  $F$  is nonphysical and never should arise in an exact theory.

For a finite system,  $F$  is defined only for integral  $N$ , and hence not for all real  $\rho$ . It can be defined for all  $\rho$  by linear interpolation, for example, but even so it can neither be expected to be, nor is it generally, convex, except in the limit. The idea behind the following proof is standard.

**Theorem 16.** The limit function  $F(\rho, \beta)$  is a convex function of  $\rho$  for each fixed  $\beta$ .  $E(\rho)$  is also a convex function of  $\rho$ .

*Proof:* This means that for  $\rho = \lambda \rho^1 + (1 - \lambda) \rho^2$ ,  $0 \leq \lambda \leq 1$ ,

$$F(\rho, \beta) \leq \lambda F(\rho^1, \beta) + (1 - \lambda) F(\rho^2, \beta) \quad (92)$$

and similarly for  $E(\rho)$ . As  $F$  is bounded *above* on bounded  $\rho$  intervals (this can be proved by a simple variational calculation), it is sufficient to prove (92) when  $\lambda = \frac{1}{2}$ . To avoid technicalities (which can be sup-

plied) and concentrate on the main idea, we shall here prove (92) when  $\rho^2$  and  $\rho^1$  are rationally related:  $a\rho^1 = b\rho^2$ ,  $a$  and  $b$  positive integers. Choose any neutral particle number  $M$  and define a sequence of balls  $B_j$  with radii as given in (81) and with  $28(4\pi/3)\rho R_0^3 = (a+b)|M|$ . For the  $\rho$  system take  $N_0 = (a+b)M$ ,  $N_j = (28)^{j-1}N_0$ ,  $j \geq 1$ . For the  $\rho^1$  (respectively,  $\rho^2$ ) system take  $N_0^1 = 2bM$ ,  $N_j^1 = (28)^{j-1}N_0^1$  [respectively,  $N_0^2 = 2aM$ ,  $N_j^2 = (28)^{j-1}N_0^2$ ]. Consider the  $\rho$  system. In the canonical partition  $\pi$  of  $B_K$  into smaller balls (Theorem 14) note that the number of balls  $B_j$  is  $m_{K-j}$  and this number is *even*. In *half* of these balls place  $N_j^1$  particles and in the other half place  $N_j^2$  particles,  $0 \leq j \leq K-1$ . Then in place of (88) we get

$$F_K(\rho) \leq \frac{1}{2} \sum_{j=0}^{K-1} f_{K-j} [F_j(\rho^1) + F_j(\rho^2)] \quad (93)$$

in an obvious notation. Inserting (89) on the right side of (93),

$$F_K(\rho) \leq \frac{1}{2} [F_K(\rho^1) + F_K(\rho^2)] + \frac{1}{2} (d_K^1 + d_K^2). \quad (94)$$

Since  $\lim_{K \rightarrow \infty} d_K^1 = 0$ , we can take the limit  $K \rightarrow \infty$  in Eq. (94) and obtain (92). ■

(E) The convexity in  $\rho^1$  and concavity in  $\beta$  of  $F(\rho, \beta)$  has another important consequence. Since  $F$  is bounded below (Theorem 13) and bounded above (by a simple variational argument) on bounded sets in the  $(\rho, \beta)$  plane, the convexity/concavity implies that it is *jointly continuous* in  $(\rho, \beta)$ . This, together with the monotonicity in  $K$  of  $F_K + \gamma d_K$  (see (90)), implies by a standard argument using Dini's theorem that *the thermodynamic limit is uniform* on bounded  $(\rho, \beta)$  sets. This uniformity is sometimes overlooked as a basic desideratum of the thermodynamic limit. Without it one would have to fix  $\rho$  and  $\beta$  *precisely* in taking the limit—an impossible task experimentally. With it, it is sufficient to have merely an increasing sequence of systems such that  $\rho_j \rightarrow \rho$  and  $\beta_j \rightarrow \beta$ . The same result holds for  $e(\rho)$ .

(F) An application of the uniformity of the limit for  $e(\rho)$  is the following. Instead of confining the particles to a box (Dirichlet boundary condition for  $H_{n,k}$ ) one could consider  $H_{n,k}$  defined on all of  $L^2(\mathbf{R}^3)|N|$ , i.e., no confinement at all. In this case

$$E_N^Q = \inf_{\psi} \langle \psi, H_{n,k} \psi \rangle / \langle \psi, \psi \rangle$$

is just the ground state energy of a neutral molecule and it is expected that  $E_N^Q/|N|$  has a limit. Indeed, this limit exists and it is simply

$$\lim_{n \rightarrow \infty} E_N^Q/|N| = \lim_{\rho \rightarrow 0} \rho^{-1} e(\rho).$$

There is no analog of this for  $F(\rho, \beta)$  because removing the box would cause the partition function to be infinite even for a finite system.

(G) The ensemble used here is the *canonical* ensemble. It is possible to define and prove the existence of the thermodynamic limit for the *microcanonical* and *grand canonical* ensembles and to show that all three ensembles are equivalent (i.e., that they yield the same values for all thermodynamic quantities, such as the pressure). (See Lieb and Lebowitz, 1972.)

(H) Charge neutrality was essentially for taming the long range Coulomb force. What happens if the system is *not neutral*? To answer this let  $N_j, \Omega_j$  be a sequence

of pairs of particle numbers and domains, but without (59) being satisfied. Let  $Q_j = zk_j - n_j$  be the net charge,  $\rho_j = |N_j|/|\Omega_j|$  as before, and  $\rho_j \rightarrow \rho$ . One expects that if (i)  $Q_j|\Omega_j|^{-2/3} \rightarrow 0$  then the same limit  $F(\rho, \beta)$  is achieved as if  $Q_j = 0$ .

On the other hand, if

(ii)  $Q_j|\Omega_j|^{-2/3} \rightarrow \infty$  then there is no limit for  $F(N_j, \Omega_j, \beta)$ . More precisely  $F(N_j, \Omega_j, \beta) \rightarrow \infty$  because the minimum electrostatic energy is too great. Both of these expectations can be proved to be correct.

The interesting case is if

(iii)  $\lim_{j \rightarrow \infty} Q_j|\Omega_j|^{-2/3} = \sigma$  exists. Then one expects a shape dependent limit to exist as follows. Assume that the  $\Omega_j$  are geometrically similar, i.e.,  $\Omega_j = \lambda_j \Omega_0$  with  $|\Omega_0| = 1$  and  $|N_j|\lambda_j^{-3} = \rho_j$  with  $\rho_j \rightarrow \rho$ . Let  $C$  be the electrostatic capacity of  $\Omega_0$ ; it depends upon the shape of  $\Omega_0$ . The capacity of  $\Omega_j$  is then  $C_j = C\lambda_j$ . From elementary electrostatics theory the expectation is that

$$\lim_{j \rightarrow \infty} F(N_j, \Omega_j, \beta) = F(\rho, \beta) + \sigma^2/2C. \quad (95)$$

Note that  $(Q_j^2/2C_j)|\Omega_j|^{-1} \rightarrow \sigma^2/2C$ .

Equation (95) can be proved for ellipsoids and balls. The proof is as complicated as the result is simple. With work, the proof could probably be pushed through for other domains  $\Omega_0$  with smooth boundaries.

The result (95) is amazing and shows how special the Coulomb force is. It says that the surplus charge  $Q_j$  goes to a thin layer near the surface. There, only its electrostatic energy, which overwhelms its kinetic energy, is significant. The bulk of  $\Omega_j$  is neutral and uninfluenced by the surface layer because the latter generates a constant potential inside the bulk. It is seldom that one has two strongly interacting subsystems and that the final result has no cross terms, as in Eq. (95).

(I) There might be a temptation, which should be avoided, to suppose that the thermodynamic limit describes a single phase system of uniform density. The temptation arises from the construction in the proof of Theorem 15 in which a large domain  $B_K$  is partitioned into smaller domains having essentially constant density. Several phases can be present inside a large domain. Indeed, if  $\beta$  is very large a solid is expected to form, and if the average density,  $\rho$ , is smaller than the equilibrium density,  $\rho_s$ , of the solid a dilute gas phase will also be present. The location of the solid inside the larger domain will be indeterminate.

From this point of view, there is an amusing, although expected, aspect to the theorem given in Eq. (95). Suppose that  $\beta$  is very large and that  $\rho < \rho_s$ . Suppose, also, that a surplus charge  $Q = \sigma V^{2/3}$  is present, where  $V$  is the volume of the container. In equilibrium, the surplus charge will never be bound to the surface of the solid, for that would give rise to a larger free energy than in (95).

(J) The inequality (53) of Sec. IV, together with known facts about the ideal gas, permit one to derive upper and lower bounds to the free energy and pressure for any neutral mixture of electrons and various nuclei. These bounds are absolutely rigorous and involve no approximation whatsoever (beyond the assumption of nonrelativistic Schrödinger mechanics with purely Coulomb forces).

If one has bounds on the free energy per unit volume

$$F^L(\rho, \beta) \leq F(\rho, \beta) \leq F^U(\rho, \beta), \quad (96)$$

then since the pressure  $P$  is equal to  $-F + \rho \partial F / \partial \rho$ , and since  $F$  is convex in  $\rho$ , one has that

$$\begin{aligned} P &\leq -F + \rho \max_{\epsilon > 0} \epsilon^{-1} \{F(\rho + \epsilon, \beta) - F(\rho, \beta)\}, \\ P &\geq -F + \rho \max_{\epsilon > 0} \epsilon^{-1} \{F(\rho, \beta) - F(\rho - \epsilon, \beta)\}. \end{aligned} \quad (97)$$

Inserting (96) into (97) yields bounds on  $P$ .

Equation (96) comes from bounds on  $Z$  [see Eq. (63)].

Using (70) and  $\rho = \rho_{\text{nuc}} + \rho_{\text{el}}$

$$F^U(\rho, \beta) = F_{\text{el}}^0(\rho_{\text{el}}, \beta) + F_{\text{nuc}}^0(\rho_{\text{nuc}}, \beta) + \langle W \rangle / |\Omega|, \quad (98)$$

where  $F^0$  is the ideal gas free energy, and  $\langle W \rangle / |\Omega|$  is the average total Coulomb energy per unit volume in the ideal gas state. This can easily be computed in terms of exchange integrals. To obtain  $F^L$ , choose  $0 < \gamma < 1$  and write  $H_{n,k} = (1 - \gamma)T_{\text{el}} + T_{\text{nuc}} + h(\gamma)$ , where  $T$  is the kinetic energy operator, and  $h(\gamma) = \gamma T_{\text{el}} + W$ .  $h(\gamma)$  is bounded below by  $A/\gamma \equiv [\text{right side of Eq. (53)}]/\gamma$ . Thus

$$Z \leq \exp[-\beta h(\gamma)] \text{Tr} \exp[-\beta((1 - \gamma)T_{\text{el}} + T_{\text{nuc}})]$$

and

$$\begin{aligned} F^L = F_{\text{nuc}}^0(\rho_{\text{nuc}}, \beta) + \max_{0 < \gamma < 1} \{ &(1 - \gamma)F_{\text{el}}^0(\rho_{\text{el}}, (1 - \gamma)\beta) \\ &+ \gamma^{-1}A/|\Omega| \}. \end{aligned} \quad (99)$$

A numerical evaluation of these bounds will be presented elsewhere.

As a final remark, the existence of the thermodynamic limit (and hence the existence of intensive thermodynamic variables such as the pressure) does not establish the existence of a unique thermodynamic state. In other words, it has not been shown that correlation functions, which always exist for finite systems, have unique limits as the volume goes to infinity. Indeed, unique limits might not exist if several phases are present. For well behaved potentials there are techniques available for proving that a state exists when the density is small, but these techniques do not work for the long range Coulomb potential. Probably the next chapter to be written in this subject will consist of a proof that correlation functions are well defined in the thermodynamic limit when  $\rho$  or  $\beta$  is small.

## REFERENCES

- Balázs, N., 1967, "Formation of stable molecules within the statistical theory of atoms," *Phys. Rev.* **156**, 42-47.
- Barnes, J. F., 1975, private communication.
- Birman, M. S., 1961, *Mat. Sb.* **55** (97), 125-174 ["The spectrum of singular boundary value problems," *Am. Math. Soc. Transl. Ser. 2* **53**, 23-80 (1966)].
- Dirac, P. A. M., 1930, "Note on exchange phenomena in the Thomas atom," *Proc. Camb. Phil. Soc.* **26**, 376-385.
- Dyson, F. J., 1967, "Ground-state energy of a finite system of charged particles," *J. Math. Phys.* **8**, 1538-1545.
- Dyson, F. J., and A. Lenard, 1967, "Stability of matter. I," *J. Math. Phys.* **8**, 423-434.
- Fermi, E., 1927, "Un metodo statistico per la determinazione di alcune priorità dell'atome," *Rend. Acad. Naz. Lincei* **6**,



- 602–607.
- Fock, V., 1930, "Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems," *Z. Phys.* **61**, 126–148; see also V. Fock, 1930, "Selfconsistent field" mit austausch für Natrium," *Phys.* **62**, 795–805.
- Gombás, P., 1949, *Die statistischen Theorie des Atomes und ihre Anwendungen* (Springer Verlag, Berlin).
- Hartree, D. R., 1927–1928, "The wave mechanics of an atom with a non-Coulomb central field. Part I. Theory and methods," *Proc. Camb. Phil. Soc.* **24**, 89–110.
- Heisenberg, W., 1927, "Über den anschaulichen Inhalt der quanten-theoretischen Kinematik und Mechanik," *Z. Phys.* **43**, 172–198.
- Jeans, J. H., 1915, *The Mathematical Theory of Electricity and Magnetism* (Cambridge University, Cambridge) 3rd ed., p. 168.
- Kato, T., "Fundamental properties of Hamiltonian operators of Schrödinger type," *Trans. Am. Math. Soc.* **70**, 195–211.
- Kirzhnits, D. A., 1957, *J. Exptl. Theoret. Phys. (USSR)* **32**, 115–123 [Engl. transl. "Quantum corrections to the Thomas-Fermi equation," *Sov. Phys. JETP* **5**, 64–71 (1957)].
- Kompaneets, A. S., and E. S. Pavlovskii, 1956, *J. Exptl. Theoret. Phys. (USSR)* **31**, 427–438 [Engl. transl. "The self-consistent field equations in an atom," *Sov. Phys. JETP* **4**, 328–336 (1957)].
- Lenard, A., and F. J. Dyson, 1968, "Stability of matter. II," *J. Math. Phys.* **9**, 698–711.
- Lenz, W., 1932, "Über die Anwendbarkeit der statistischen Methode auf Ionengitter," *Z. Phys.* **77**, 713–721.
- Lieb, E. H., 1976, "Bounds on the eigenvalues of the Laplace and Schrödinger operators," *Bull. Am. Math. Soc.*, in press.
- Lieb, E. H., and J. L. Lebowitz, 1972, "The constitution of matter: existence of thermodynamics for systems composed of electrons and nuclei," *Adv. Math.* **9**, 316–398. See also J. L. Lebowitz and E. H. Lieb, 1969, "Existence of thermodynamics for real matter with Coulomb forces," *Phys. Rev. Lett.* **22**, 631–634.
- Lieb, E. H., and H. Narnhofer, 1975, "The thermodynamic limit for jellium," *J. Stat. Phys.* **12**, 291–310; Erratum: *J. Stat. Phys.* **14**, No. 5 (1976).
- Lieb, E. H., and B. Simon, "On solutions to the Hartree-Fock problem for atoms and molecules," *J. Chem. Phys.* **61**, 735–736. Also, a longer paper in preparation.
- Lieb, E. H., and B. Simon, 1977, "The Thomas-Fermi theory of atoms, molecules and solids," *Adv. Math.*, in press. See also E. H. Lieb and B. Simon, 1973, "Thomas-Fermi theory revisited," *Phys. Rev. Lett.* **33**, 681–683.
- Lieb, E. H., and W. E. Thirring, 1975, "A bound for the kinetic energy of fermions which proves the stability of matter," *Phys. Rev. Lett.* **35**, 687–689; Errata: *Phys. Rev. Lett.* **35**, 1116. For more details on kinetic energy inequalities and their application, see also E. H. Lieb and W. E. Thirring, 1976, "Inequalities for the moments of the Eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities," in *Studies in Mathematical Physics: Essays in Honor of Valentine Bargmann*, edited by E. H. Lieb, B. Simon, and A. S. Wightman (Princeton University, Princeton).
- Rosen, G., 1971, "Minimum value for  $c$  in the Sobolev inequality  $\|\phi\|_3 \leq c \|\nabla\phi\|$ ," *SIAM J. Appl. Math.* **21**, 30–32.
- Schwinger, J., 1961, "On the bound states of a given potential," *Proc. Nat. Acad. Sci. (U.S.)* **47**, 122–129.
- Scott, J. M. C., 1952, "The binding energy of the Thomas Fermi atom," *Phil. Mag.* **43**, 859–867.
- Sheldon, J. W., 1955, "Use of the statistical field approximation in molecular physics," *Phys. Rev.* **99**, 1291–1301.
- Slater, J. C., 1930, "The theory of complex spectra," *Phys. Rev.* **34**, 1293–1322.
- Sobolev, S. L., 1938, *Mat. Sb.* **46**, 471. See also S. L. Sobolev, 1950, "Applications of functional analysis in mathematical physics," Leningrad; Am. Math. Soc. Transl. Monographs **7** (1963).
- Sommerfeld, A., 1932, "Asymptotische Integration der Differential-gleichung des Thomas-Fermischen Atoms," *Z. Phys.* **78**, 283–308.
- Teller, E., 1962, "On the stability of molecules in the Thomas-Fermi theory," *Rev. Mod. Phys.* **34**, 627–631.
- Thomas, L. H., 1927, "The calculation of atomic fields," *Proc. Camb. Phil. Soc.* **23**, 542–548.
- Von Weizsäcker, C. F., 1935, "Zur Theorie der Kernmassen," *Z. Phys.* **96**, 431–458.