Bound on the maximum negative ionization of atoms and molecules

Elliott H. Lieb
Departments of Mathematics and Physics, Princeton University, P.O. Box 708, Princeton, New Jersey 08544
(Received 14 November 1983)

It is proved that N_c, the number of negative particles that can be bound to an atom of nuclear charge z, satisfies $N_c < 2z+1$. For a molecule of K atoms, $N_c < 2Z+K$ where Z is the total nuclear charge. As an example, for hydrogen $N_c = 2$, and thus H^{--} is not stable, which is a result not proved before. The bound particles can be a mixture of different species, e.g., electrons and π mesons; statistics plays no role. The theorem is proved in the static-nucleus approximation, but if the nuclei are dynamical, a related, weaker result is obtained. The kinetic energy operator for the particles can be either $[p-eA(x)/c]^2/2m$ (nonrelativistic with magnetic field) or $\{[pc-eA(x)]^2+m^2c^4\}^{1/2}-mc^2$ (relativistic with magnetic field). This result is not only stronger than that obtained before, but the proof (at least in the atomic case) is simple enough to be given in an elementary quantum-mechanics course.

I. INTRODUCTION

One of the nonperiodic facts about the Periodic Table is that the number of electrons that can be bound to an atomic nucleus of charge ze is at most $z+1$, at least as far as present confirmed experimental data go. The theoretical proof of this fact, starting with the Schrödinger equation, is a challenge that has drawn the attention of several authors in recent years.[1–13]

The problem can obviously be extended in two ways: (i) The electrons, which are fermions, can be replaced by bosons or, more generally, by a mixture of particles of different species. (Because of spin, the two-electron problem is well known to be the same as the two-boson problem.) (ii) Instead of a single atom, a molecule can be considered. That these problems are difficult is shown by the fact that it was only recently proved by Ruskai[1] (for bosons) and later Sigal[2] and Ruskai[4] (for fermions) that the number of bound particles is not infinite.

The following is a summary of rigorous results to date. Our notation for the maximum particle number is N_c.

(1) Ruskai[1] proved that $N_c \le (\text{const})z^2$ for bosons. Recently, Sigal[3] proved for bosons that for every $\epsilon > 0$ there is a constant C_ϵ such that $N_c \le C_\epsilon z^{1+\epsilon}$.

(2) For fermions, Sigal[2] proved that $N_c \le cz$ with c being some constant. Ruskai[4] proved that $N_c \le (\text{const})z^{6/5}$. Sigal[3] improved his result to $N_c \le \alpha(z)z$ with $\alpha(z) \le 12$ and $\alpha(z) \to 2$ as $z \to \infty$.

(3) For bosons, Benguria and Lieb[5] proved that $N_c \ge \beta(z)z$ with $\beta(z) \to 1+\gamma$ as $z \to \infty$. Here, γ is some number satisfying $0 < \gamma < 1$ and is obtained by solving a Hartree equation. This equation was subsequently solved numerically by Baumgartner[6,11] with the result $\gamma = 0.21$. Thus, bosons strongly violate the $z+1$ rule; the Pauli exclusion principle plays a key role in the electron problem.

(4) In a related development, Benguria and Lieb[7] studied the Thomas–Fermi–von Weizsäcker (TFW) equation—a well-known density-functional equation which is supposed to imitate the Schrödinger equation for fermions. They proved that $N_c \le z + 0.73$ in the TFW model of an atom. (In TFW theory, the electronic charge is not quantized.) Thus, on the one hand, TFW theory really imitates the Schrödinger equation and, on the other hand, the TFW result supports the conjecture that $N_c \le z + \text{const}$ for the Schrödinger equation for arbitrarily large z. Earlier,[14] it had been shown that $Z < N_c < 2Z$ in TFW theory, even for a molecule. Here Ze is the total nuclear charge,

$$Z = \sum_{j=1}^{K} z_j \ , \tag{1.1}$$

for a molecule of K atoms with nuclear charges ez_j. Note that in Thomas–Fermi (TF) theory one always has $N_c = Z$ for any molecule.[14]

(5) Lieb, Sigal, Simon, and Thirring,[8] using Sigal's method,[3] have proved that $N_c/z \to 1$ as $z \to \infty$ in the fermion case.

(6) For $z=1$ (hydrogen), Hill[9] proved that three electrons cannot be bound in a quartet ($S = \frac{3}{2}$) state. There does not seem to be any proof of nonbinding for the doublet ($S = \frac{1}{2}$) state with $N=3$.

All of the above results are for a single atom. The Ruskai and Sigal methods can be extended to the molecular case; this was explicitly done for bosons.[1] The Benguria-Lieb result[7] in (4) extends to a molecule: $N_c < Z + 0.73$ K. Furthermore, all the results apply to the fixed-nucleus (sometimes called Born-Oppenheimer) approximation.

(7) Zhislin[10] proved that $N_c \ge z$ for an arbitrary mixture of particles (with any statistics) and including nuclear motion. This result extends to a molecule, $N_c \ge Z$.

In this paper it will be proved (Theorem 1) that

$$N_c < 2z+1 \tag{1.2}$$

for a fixed-nucleus atom and with any mixture of bound particles (with possibly different masses and statistics). Equation (1.2) holds if all the charges are $-e$, but a similar result holds with nonconstant but negative charges [see Eq. (2.11)]. If z is an integer, as in the physical case, (1.2) implies

$$N_c \le 2z \ . \tag{1.3}$$

This completes the story for hydrogen, i.e., $N_c=2$, since it is well known that two electrons can indeed be bound [see (6) above]. H^{--} is *not* stable. (Incidentally, I am not aware of any proof of the obvious assertion that if N electrons cannot be bound then $M>N$ electrons cannot be bound. Thus, even if the $S=\frac{1}{2}$ case in (6) above were settled, it would not immediately follow from this that $N_c=2$.) Equation (1.3) also states that two π^- mesons and a muon cannot be bound with $z=1$. Equation (1.2) implies that the critical z to bind two particles is at least 0.5. The exact value[12] is 0.9112.

For large z, Eq. (1.3) is hardly optimal in view of the conjecture that N_c-z is of order unity. For bosons, however, Eq. (1.3) gives the right order of magnitude since $N_c \geq 1.2z$ for large z [see (3) above].

In the case of a molecule of K atoms it will be proved that

$$N_c < 2Z+K \tag{1.4}$$

for fixed nuclei. Thus, for example, the hydrogen molecule cannot bind more than five particles. Again, this holds for arbitrary negative particles with common charge $-e$. For nonconstant charges see Eq. (2.10). A summary of the results of this paper is in Ref. 13.

As stated above, Theorem 1 does not require that the nuclear or the negative particle charges be integral. This generality may be relevant physically because, as pointed out to me by W. Thirring (private communication), particles in solids such as semiconductors may have nonintegral effective charges because of dielectric effects.

A remark should be made about the meaning of "fixed nuclei" in the molecular case. There are two possible interpretations.

Case A: The nuclei, of charges $z_1,\ldots,z_K>0$, have coordinates $R_1,\ldots,R_K$ which are arbitrary but which are fixed once and for all, independent of the particle number N.

Case B: For each particle number, the nuclear coordinates R_j are adjusted to minimize the total energy, namely,

$$\hat{E}_N \equiv \inf_{\underline{R}} E_N(\underline{R})+U(\underline{R}) , \tag{1.5}$$

where $\underline{R}$ denotes $\{R_1,\ldots,R_K\}$ and $E_N(\underline{R})$ is the electronic ground-state energy depending on $\underline{R}$, and $U(\underline{R})$ is the internuclear interaction. Usually $U(\underline{R})$ is the Coulomb energy

$$e^2 \sum_{1\leq i<j\leq K} z_i z_j \,|\, R_i-R_j \,|^{-1} ,$$

but for our purposes $U(\underline{R})$ can be anything.

The result in Eq. (1.4), and more generally in Theorem 1, holds for both Case A and Case B. Case B is, of course, more physical. However, Theorem 1 in Case A implies Theorem 1 in Case B; this is obvious since if binding does not occur for every choice of $\underline{R}$ it certainly does not occur for the minimizing $\underline{R}$. It is Case A (fixed nuclei) that we shall actually consider henceforth.

The case of truly dynamical nuclei cannot be treated as definitively by the method presented here. Nevertheless, some information about this system is contained in Sec. VI. Section VI also contains the extension of the previously cited results to the case of smeared, but spherical, nuclear charge densities. These results are also shown to hold in the Hartree-Fock theory.

Finally, it will also be proved that Eqs. (1.2) and (1.4) hold if the particle kinetic energy operator, which is

$$p^2/2m = -(\hbar^2/2m)\Delta \tag{1.6}$$

in the usual Schrödinger equation, is replaced by the relativistic expression

$$(p^2c^2+m^2c^4)^{1/2}-mc^2=(-c^2\hbar^2\Delta+m^2c^4)^{1/2}-mc^2 . \tag{1.7}$$

Another variation for which (1.2) and (1.4) hold is the inclusion of a magnetic field in either Eq. (1.6) or (1.7):

$$p^2 \rightarrow [\vec{p}-e\vec{A}(x)/c]^2 , \tag{1.8}$$

where $\vec{A}$ is a bounded vector potential. One can even have different $\vec{A}$'s for different particles. Other generalizations of the kinetic energy are also possible, as well as generalizations to potentials other than the Coulomb potential (see the remark in Appendix A).

Not only are the results proved here stronger than that obtained previously [except for (5) above], but the proof itself is much simpler. For the atomic case with one species of fermions or bosons the proof is so short that it can be given in an elementary quantum-mechanics course. In order to display the basic idea as clearly as possible, this atomic case will be treated first in Sec. IV. The method of proof borrows heavily from the proof in Ref. 14, Theorem 7.23 of $N_c<2Z$ for the TFW theory. (In Ref. 14, N_c was called λ_c.) As mentioned there, the basic idea for the TFW proof in the atomic case is due to Benguria.

II. PRELIMINARIES AND NOTATION

The Hamiltonian for N particles is

$$H_N=\sum_{i=1}^{N}[T_i-q_iV(x_i)]+\sum_{1\leq i<j\leq N} q_iq_j \,|\, x_i-x_j \,|^{-1} , \tag{2.1}$$

where

$$V(x)=\sum_{j=1}^{K} z_j \,|\, x-R_j \,|^{-1} \tag{2.2}$$

is the electric potential produced by K fixed nuclei of charges

$$\underline{z}=\{z_1,\ldots,z_K\}$$

located at $\underline{R}=\{R_1,\ldots,R_K\}$ with $R_i\in\mathbb{R}^3$. Units are used in which the electron charge e is unity, and $\hbar=1$. We assume $z_i>0$, all $i=1,\ldots,K$. The number $-q_i$ is the charge of the ith particle, and we assume $q_i\geq 0$ for $i=1,\ldots,N$.

The operator T_i is the kinetic energy operator for the ith particle and it is assumed to have one of the following three forms:

$$T_i^{(1)}=-\Delta_i/2m_i , \tag{2.3}$$

$$T_i^{(2)}=[i\vec{\nabla}_i+\vec{A}_i(x)]^2/2m_i\ , \tag{2.4}$$

$$T_i^{(3)}=c\{[i\vec{\nabla}_i+\vec{A}_i(x)]^2+m_i^2c^2\}^{1/2}-m_ic^2\ . \tag{2.5}$$

Here, $m_i>0$ is the mass of the ith particle, $c\vec{A}_i(x)/q_i$ is the vector potential applied to the ith particle, and c is the speed of light. $\vec{A}_i(x)$ is assumed to be bounded and to go to zero as $|x|\to\infty$. Equation (2.3) is obviously a special case of (2.4) but, for simplicity, it is treated separately. It must be noted, however, that if the form $T_i^{(3)}$ is used even for just one particle (for example, i), then every nuclear charge z_j must satisfy $\alpha q_i z_j\le 2/\pi$, where $\alpha=e^2/\hbar c$ is the fine-structure constant. The reason is that the single-particle operator $T^{(3)}-e^2qV(x)$ is bounded below (as a quadratic form) if and only if $e^2qz_j\le 2/\pi$ for every j. The situation is not changed by the addition of the third term in Eq. (2.1). (See Ref. 15.)

The ith particle can have any one of the three forms of T_i, independently. The m_i, q_i, and $\vec{A}_i$ need not be related for different i. There is one proviso, however. If several particles are of the same type (bosons or fermions) then that group must, of course, have the same m_i, q_i, and T_i. Spin can be included in the usual way. H_N is spin independent. The easiest way to treat spin is to think of the spin coordinate as merely labeling a particle type. Thus, for spin-$\frac{1}{2}$ electrons, there are two kinds of fermions: those with spin up and those with spin down.

The ground-state energy, is defined by

$$E_N=\inf \mathrm{Spec}(H_N) \tag{2.6}$$

$$=\inf_{\psi}\frac{\langle\psi|H_N|\psi\rangle}{\langle\psi|\psi\rangle}\ , \tag{2.7}$$

the latter being the variational principle. The admissible ψ in Eq. (2.7) must, of course, satisfy the required statistics for the various groups of particles. E_N need not be an eigenvalue, and it will not be if the system is not bound. When a particle is added to the system it can always be placed at infinity (i.e., arbitrarily far away and with arbitrarily small kinetic and potential energy), and we thus have the relation

$$E_{N,j}\ge E_N,\quad j=1,\dots,N\ , \tag{2.8}$$

where N,j denotes the $(N-1)$-particle system with the jth particle removed.

Definition: The N-particle system is said to be *bound* if and only if E_N is an eigenvalue, i.e.,

$$H_N\psi=E_N\psi \tag{2.9}$$

for some $\psi\in L^2(\mathbb{R}^{3N})$.

In this definition it is not required that $E_{N,j}$ be an eigenvalue for any j or that $E_N<E_{N,j}$ for any j. If $E_N<E_{N,j}$ for some j then E_N is automatically an eigenvalue. The word "bound" is *not applicable* to a system that merely has an eigenvalue in the continuum (i.e., E_N is not an eigenvalue but H_N has an eigenvalue greater than $E_{N,j}$ for some j); such a system would only be metastable. Our main result is the following.

Theorem 1: Let $Z=\sum_{j=1}^K z_j$ as before. Let $Q=\sum_{i=1}^N q_i$ be the negative of the total particle charge and let q be the maximum of the q_i. Then, if binding occurs, the following must be satisfied:

$$Q<2Z+qK\ . \tag{2.10}$$

In the atomic case ($K=1$), Eq. (2.10) can be replaced by the slightly stronger requirement

$$Q<2Z+\sum_{i=1}^N q_i^2/Q\ . \tag{2.11}$$

III. STRATEGY OF THE PROOF

Given Eq. (2.9) select one variable, for example, j, and a positive function of one variable, denoted by $1/\phi(x)$, to be determined appropriately. Multiply Eq. (2.9) by

$$\psi^*(x_1,\dots,x_N)/\phi(x_j)$$

and integrate over all the variables, and then take the real part. On the left-hand side there will be four terms:

$$h_j=\langle[\psi/\phi(x_j)]\,|\,H_{N,j}\,|\,\phi\rangle\ , \tag{3.1}$$

$$t_j=\mathrm{Re}\langle[\psi/\phi(x_j)]\,|\,T_j\,|\,\psi\rangle\ , \tag{3.2}$$

$$-a_j=-q_j\langle[\psi/\phi(x_j)]\,|\,V(x_j)\,|\,\psi\rangle\ , \tag{3.3}$$

$$r_j=\left\langle[\psi/\phi(x_j)]\left|\sum_{\substack{i\\ i\neq j}}q_iq_j\,|\,x_i-x_j\,|^{-1}\right|\psi\right\rangle\ . \tag{3.4}$$

[The right-hand sides of Eqs. (3.1), (3.3), and (3.4) are automatically real.] On the right-hand side there will be E_NI_j with

$$I_j=\langle[\psi/\phi(x_j)]\,|\,\psi\rangle\ . \tag{3.5}$$

Let us assume, provisionally, that all these terms are finite.

Because of Eq. (2.8)

$$E_NI_j\le h_j\ . \tag{3.6}$$

To see this, let X_j denote the variables $x_1,\dots,x_N$ with x_j excluded. Consider

$$P_j(X_j,X_j')=\int\psi(x_j,X_j)[\psi^*(x_j,X_j')/\phi(x_j)]d^3x_j\ , \tag{3.7}$$

so that

$$\Gamma_j(X_j,X_j')\equiv P_j(X_j,X_j')/I_j$$

is a properly normalized density matrix. (The positivity of ϕ is crucial here.) Moreover, Γ_j satisfies the correct statistics in the X_j variables (since ψ does) and thus, by the variational principle,

$$h_{N,j}/I_j=\mathrm{Tr}(H_{N,j}\Gamma_j)\ge E_{N,j}\ . \tag{3.8}$$

This, together with Eq. (2.8), proves (3.6). Thus, binding cannot occur if

$$t_j-a_j+r_j>0\ . \tag{3.9}$$

Let $\phi(x)$ be any function of the form

$$\phi(x)=\int|x-y|^{-1}d\mu(y)+C\ , \tag{3.10}$$

where $d\mu$ is a (positive) measure with $0<\int d\mu<\infty$ and $C\ge 0$ is a constant. In our application we shall choose ϕ

to be of the same form as V but with different coefficients $\mu_j > 0$:

$$\phi(x) = \sum_{j=1}^{k} \mu_j |x - R_j|^{-1} . \tag{3.11}$$

It will be proved that for every j

$$t_j > 0 \tag{3.12}$$

for any ψ and ϕ and choice of T_j [see Eqs. (2.3)—(2.5)]. Consequently, binding cannot occur if, for any $j=1,\ldots,N$,

$$r_j \geq a_j . \tag{3.13}$$

Actually, we shall not prove Eq. (3.13) for any particular j, but shall prove instead that when Eq. (2.10) or (2.11) is violated then

$$R \equiv \sum_{j=1}^{N} r_j \geq A \equiv \sum_{j=1}^{N} a_j . \tag{3.14}$$

[In fact, we shall actually prove that $R > A$ and, therefore, that strict inequality is not really needed in Eq. (3.12). The reason that $R > A$ is given in the second part of Appendix A.] Clearly, Eq. (3.14) implies that Eq. (3.13) holds for some j, and thus Theorem 1 will be proved.

To prove (3.12) we note that the X_j integration in (3.2) can be done after the x_j integration. Therefore, (3.12) will be true if it holds for each X_j, i.e., if it holds for any function of the one variable x_j. Thus, if $f(x) \in L^2(R^3)$ we want to prove that

$$t \equiv \mathrm{Re} \int f^*(x)[(Tf)(x)/\phi(x)]d^3x > 0 . \tag{3.15}$$

When $T = p^2/2m$, (3.15) was first proved for f spherically symmetric and real by Benguria, and then for f real by Lieb. This was given in Ref. 14, Lemma 7.21. Baumgartner[11] found a more direct proof and also extended (3.15) to complex f. Baumgartner's proof easily extends to $T^{(2)}$, but the proof for $T^{(3)}$ is very different and is given in Appendix A.

In Appendix A a proof of (3.15) under carefully stated conditions on f is given. The following technical point, which is also discussed in Appendix A, has to be considered: We assumed that all the quantities in Eqs. (3.1)—(3.5) are finite. By the condition stated after Eq. (3.10), r_j and a_j are automatically finite. Conceivably, h_j, t_j, and I_j could be infinite. If so, this can be remedied by replacing $\phi(x)$ by $\phi(x)+C$ and then letting $C \to 0$ at the end. This procedure is also discussed in Appendix A.

Since one of our stated goals is to present a proof of Theorem 1 in the atomic case that is simple enough to be given in an elementary quantum-mechanics course, let us temporarily suspend any reservations about technicalities and give the following proof of (3.15), following Baumgartner's method,[11] when $T = p^2/2m$.

The key fact is that $\Delta\phi \leq 0$. Given f and ϕ, define $g(x) = f(x)/\phi(x)$. We then require that

$$2mt = -\mathrm{Re} \int g^*(x)\Delta[g(x)\phi(x)]d^3x > 0 . \tag{3.16}$$

By partial integration,

$$\begin{aligned} 2mt &= \mathrm{Re} \int \vec{\nabla} g^*(x) \cdot \vec{\nabla}[\phi(x)g(x)]d^3x \\ &= \mathrm{Re} \int \vec{\nabla} g^*(x) \cdot [\phi(x)\vec{\nabla} g(x) + g(x)\vec{\nabla}\phi(x)]d^3x \\ &= \int \phi(x) |\vec{\nabla} g(x)|^2 d^3x + \tfrac{1}{2} \int \vec{\nabla}\phi(x) \cdot [g(x)\vec{\nabla} g^*(x) + g^*(x)\vec{\nabla} g(x)]d^3x \\ &= \int \phi(x) |\vec{\nabla} g(x)|^2 d^3x + \tfrac{1}{2} \int \vec{\nabla}\phi(x) \cdot \vec{\nabla}[|g(x)|^2]d^3x \\ &= \int \phi(x) |\vec{\nabla} g(x)|^2 d^3x - \tfrac{1}{2} \int |g(x)|^2 \Delta\phi(x) d^3x > 0 . \end{aligned} \tag{3.17}$$

[Note that we have greater than 0 instead of greater than or equal to 0 because $\int \phi |\vec{\nabla} g|^2 > 0$ since $\phi(x) > 0$ for all x; if $\vec{\nabla} g(x) \equiv 0$ then $f = (\text{const})\phi$, but $\phi \notin L^2$ and $f \in L^2$.]

In the rest of this paper (except Appendix A) we shall assume that (3.15) holds and shall concentrate on proving that Eq. (3.14) holds. As mentioned before, all r_j and a_j are necessarily real and finite.

IV. ATOMS WITH IDENTICAL PARTICLES

Take the nuclear coordinate $R_1 = 0$, $z_1 \equiv z$ and let $\phi(x) = 1/|x|$. Assume that the particles are identical (bosons or fermions) so that $q_j = q$, $r_j = r$, and $a_j = a$ are independent of j. (Note: If the particles are fermions they are allowed to have spin, in which case $\int dx$ in the following should be understood as $\sum_\sigma \int dx$, where σ is the spin variable.)

We denote $\{x_1,\ldots,x_N\}$ by X and assume that ψ is normalized. Taking $j=1$ we have

$$a = zq \int [|\psi(X)|^2 |x_1| / |x_1|]d^{3N}X = zq , \tag{4.1}$$

$$r = q^2 \sum_{j=2}^{N} \int |\psi(X)|^2 |x_1| \, |x_1 - x_j|^{-1} d^{3N}X \tag{4.2}$$

$$= q^2(N-1) \int |\psi(X)|^2 |x_1| \, |x_1 - x_2|^{-1} d^{3N}X . \tag{4.3}$$

In going from Eq. (4.2) to (4.3) the fact that $|\psi|^2$ is symmetric was used. This symmetry also implies that the integral in Eq. (4.3) is not changed if $|x_1|$ is replaced by $|x_2|$. Thus,

$$r = \tfrac{1}{2} q^2(N-1) \int |\psi(X)|^2 [|x_1| + |x_2|] \times |x_1 - x_2|^{-1} d^{3N}X \geq \tfrac{1}{2} q^2(N-1) \tag{4.4}$$

since

$$|x_1| + |x_2| \geq |x_1 - x_2|$$

by the triangle inequality.

Therefore,

$$r-a \geq \tfrac{1}{2}q^2(N-1)-zq \tag{4.5}$$

and, by Eq. (3.13), binding will not occur if $Q \geq 2z+q$ with $Q=Nq$. This proves Theorem 1 in this special case.

V. GENERAL CASE

Our goal here is to prove (3.14) if (2.11) (atomic case) or (2.10) (molecular case) is violated. In the atomic case without particle symmetry, we cannot use symmetry to go from Eq. (4.2) to (4.3). In the molecular case, even if particle symmetry exists, the obvious choice $\phi(x)=V(x)$ will not work and we shall have to use ϕ as given in Eq. (3.11).

In the atomic case

$$\phi(x)=1/|x|=V(x)/z$$

does work. Assume ψ to be normalized. For each j, Eqs. (4.1) and (4.2) become

$$a_j=zq_j\ , \tag{5.1}$$

$$r_j=\sum_{i\,(\neq j)} q_iq_j\int |\psi(X)|^2\,|x_j|\,|x_i-x_j|^{-1}d^{3N}X\ . \tag{5.2}$$

Now sum these over j to obtain [recalling Eq. (3.14)]

$$R-A=\tfrac{1}{2}\sum_{\substack{i,j\\ i\neq j}} q_iq_j\int |\psi(X)|^2(|x_i|+|x_j|)\times|x_i-x_j|^{-1}d^{3N}X-Qz \tag{5.3}$$

$$\geq \tfrac{1}{2}\sum_{\substack{i,j\\ i\neq j}} q_iq_j-Qz=\tfrac{1}{2}Q^2-\tfrac{1}{2}\sum_i q_i^2-Qz\ . \tag{5.4}$$

Again, the triangle inequality $|x|+|y|\geq|x-y|$ has been used. Clearly, if Eq. (2.11) is violated then $R\geq A$ and this proves Theorem 1 in the atomic case.

In the molecular case the following is needed. It does not depend on the dimension d being 3.

Lemma 1: Let $\psi(X)$ be any normalized function in $L^2(R^{Nd})$ (without any particular symmetry). For $j=1,\ldots,N$ let

$$\rho_j(x)=\int |\psi(x,X_j)|^2d^{Nd-d}X_j \tag{5.5}$$

be the one-particle density for particle j. Let $g_1(x),\ldots,g_N(x)$ be any given functions of one variable such that $g_j(x_j)\psi(X)$ is in $L^2(\mathbb{R}^{Nd})$ and define

$$\langle g_j\rangle=\int \rho_j(x)g_j(x)d^dx=\int |\psi(X)|^2g_j(x_j)d^{Nd}X\ . \tag{5.6}$$

Then

$$\mathrm{Re}\int |\psi(X)|^2\sum_{1\leq i<j\leq N} g_i^*(x_i)g_j(x_j)d^{Nd}X \geq \tfrac{1}{2}\left|\sum_{j=1}^N\langle g_j\rangle\right|^2-\tfrac{1}{2}\sum_{j=1}^N\langle|g_j|^2\rangle\ . \tag{5.7}$$

This is proved simply by noting that

$$\mathrm{Re}\sum_{\substack{i,j\\ i<j}} g_i^*g_j=\tfrac{1}{2}\left|\sum_i g_i\right|^2-\tfrac{1}{2}\sum_i|g_i|^2$$

and then using the Schwarz inequality on the first term.

To apply Lemma 1 to the general case, let

$$\tilde{\rho}(x)\equiv\sum_{j=1}^N q_j\rho_j(x) \tag{5.8}$$

be the negative of the single particle *charge* density for ψ [see Eq. (5.5)]. Let $\phi(x)$ be the potential in Eq. (3.11). We then have

$$A=\sum_{j=1}^N a_j=\sum_{s=1}^K z_s\gamma_s\ , \tag{5.9}$$

with

$$\gamma_s=\int\tilde{\rho}(x)\,|x-R_s|^{-1}\phi(x)^{-1}d^3x\ . \tag{5.10}$$

$$R=\sum_{j=1}^N r_j=\sum_{1\leq i<j\leq N} q_iq_j\int|\psi(X)|^2\,|x_i-x_j|^{-1}\times[\phi(x_i)^{-1}+\phi(x_j)^{-1}]d^{3N}X\ . \tag{5.11}$$

Let us write [following an idea in Ref. 11)]

$$\phi(x)^{-1}+\phi(y)^{-1}=[\phi(x)\phi(y)]^{-1}[\phi(x)+\phi(y)]=\sum_{j=1}^K\mu_j[\phi(x)\,|x-R_j|\,]^{-1}[\phi(y)\,|y-R_j|\,]^{-1}(|x-R_j|+|y-R_j|)\ . \tag{5.12}$$

Again, noting that

$$|x-R_j|+|y-R_j|\geq|x-y|$$

we have, upon inserting Eq. (5.12) in (5.11),

$$R\geq\sum_{s=1}^K\mu_s\int|\psi(X)|^2\sum_{1\leq i<j\leq N} g_i^s(x_i)g_j^s(x_j)d^{3N}X\ , \tag{5.13}$$

with

$$g_i^s(x)=q_i[\phi(x)\,|x-R_s|\,]^{-1}\ . \tag{5.14}$$

Using Lemma 1 for each s we have

$$R\geq\tfrac{1}{2}\sum_{s=1}^K\mu_s\left[\gamma_s^2-\sum_{i=1}^N q_i^2\int\rho_i(x)\phi(x)^{-2}\times|x-R_s|^{-2}d^3x\right]\ . \tag{5.15}$$

However,

$$\mu_s/[\,|x-R_s|\,\phi(x)]\le 1\ ,$$

and $q_i\le q$ (by definition). Hence,

$$R\ge \tfrac{1}{2}\sum_{s=1}^{N}(\mu_s\gamma_s^2-q\gamma_s)\ . \tag{5.16}$$

We shall have $R\ge A$ if

$$\sum_{s=1}^{K}\gamma_s\{\mu_s\gamma_s-q-2z_s\}\ge 0\ , \tag{5.17}$$

and our aim is to choose the $\{\mu_s\}$ so that Eq. (5.17) is satisfied when condition (2.10) is violated. (Note that γ_s depends on $\{\mu_s\}$.) To this end, let

$$\delta_s\equiv\mu_s\gamma_s/Q\ , \tag{5.18}$$
$$\beta_s\equiv(2z_s+q)/(2Z+qK)\ .$$

Note that [see Eqs. (5.8) and (5.10)]

$$\sum_{s=1}^{K}\delta_s=\sum_{s=1}^{K}\beta_s=1\ . \tag{5.19}$$

Suppose we can choose $\{\mu_s\}$ such that

$$\delta_s=\beta_s,\quad s=1,\ldots,K\ . \tag{5.20}$$

Then the left-hand side of Eq. (5.17) becomes

$$\sum_{s=1}^{K}\gamma_s\delta_s[Q-(2Z+qK)]\ ,$$

and this is non-negative if condition (2.10) is violated.

Thus, showing that the K equations (5.20) in K unknowns have a solution proves Theorem 1. This is done in Appendix B.

VI. THREE GENERALIZATIONS OF THEOREM 1

A. Smeared nuclei

Suppose that the nuclear charge densities, instead of being points, are smeared into *spherically symmetric* distributions about R_j, namely, $z_j|x-R_j|^{-1}$ in Eq. (2.2) is replaced by

$$V_j(x)=\int|x-y-R_j|^{-1}d\mu_j(y)\ , \tag{6.1}$$

where $d\mu_j$ is a spherically symmetric (positive) measure with $\int d\mu_j(y)=z_j$. Then *Theorem 1 continues to hold without modification.* The proof is as before—with the same $\phi(x)$. One merely has to note that

$$V_j(x)\le z_j|x-R_j|^{-1} \tag{6.2}$$

for all x and, hence, a_j, defined by Eq. (3.3), is not greater than it would be for the point nucleus.

B. Dynamical nuclei

Suppose that the K nuclear coordinates $R_1,\ldots,R_K$ are dynamical variables and the Hamiltonian is

$$\tilde{H}_N=H_N+T_{\rm nuc}+U(\underline{R})\ , \tag{6.3}$$

and

$$T_{\rm nuc}=\sum_{i=1}^{K}T_i\ , \tag{6.4}$$

where H_N is given by Eq. (2.1) and each T_i is one of the operators given in Eqs. (2.3)—(2.5). $U(\underline{R})$ is a potential that depends on the nuclear coordinates, as explained after Eq. (1.5). If $U(\underline{R})$ is translation invariant, the eigenvalue equation

$$\tilde{H}_N\psi=\tilde{E}_N\psi\ , \tag{6.5}$$

with $\tilde{E}_N=\inf\mathrm{Spec}\,(\tilde{H}_N)$ and $\psi=\psi(X,\underline{R})$, would not generally be expected to have a solution in $L^2(\mathbb{R}^{3N+3K})$; if there are no magnetic fields present then it certainly would not. With magnetic fields present one cannot simply remove the center of mass motion, and the situation is complicated. To avoid technical complications it will be assumed that $U(\underline{R})$ also contains one-body terms which serve to contain the nuclei, and therefore that Eq. (6.5) indeed has an L^2 solution. Physically, this is no real restriction because the confining potential could, for example, be an infinitely high walled box of arbitrarily large size. Because the *negative* particles are *not confined* it is still true that $\tilde{E}_N\le\tilde{E}_{N,j}$.

In this case a weakened form of Theorem 1 holds. Let us define $E^{\infty}_{N,j}$ to be the inf Spec $(\tilde{H}^{\infty}_{N,j})$, with negative particle j removed as before, but with all the nuclear masses set equal to infinity. Alternatively,

$$E^{\infty}_{N,j}=\inf_{\underline{R}}E_{N,j}(\underline{R})+U(\underline{R})\ , \tag{6.6}$$

where $E_{N,j}(\underline{R})$ is the ground-state energy of $H_{N,j}$ as defined in Eq. (2.7). Clearly,

$$E^{\infty}_{N,j}<\tilde{E}_{N,j}=\inf\mathrm{Spec}(\tilde{H}_{N,j})\ . \tag{6.7}$$

Theorem 2: Suppose that the system is bound [i.e., Eq. (6.5) holds] and the binding energy satisfies

$$\tilde{E}_N-\tilde{E}_{N,j}\le E^{\infty}_{N,j}-\tilde{E}_{N,j} \tag{6.8}$$

for all $j=1,\ldots,N$. Then Eq. (2.10) holds in the molecular case $K>1$ and Eq. (2.11) holds in the atomic case $K=1$.

In the physical situation, the right-hand side of Eq. (6.9) is numerically small, but it is a challenge to eliminate this condition.

To prove Theorem 2 we multiply Eq. (6.5) by

$$\psi^*(X,\underline{R})/\phi(x_j,\underline{R})$$

and integrate over all the variables. Here $\phi(x,\underline{R})$ is as in the proof of Theorem 1, namely, Eq. (3.11) with *fixed constants* μ_j. [It is important that the μ_j do not depend on $\underline{R}$, otherwise the dependence of ϕ on each R_j would not have the form of Eq. (3.10), and thus the positivity of the integrated $T_{\rm nuc}$ term might be lost.]

Let us write (for any fixed j, $1\le j\le N$)

$$\tilde{H}_N=\tilde{H}^{\infty}_{N,j}+T_{\rm nuc}+T_j-q_jV(x_j,\underline{R})+\sum_{i\,(\ne j)}|x_i-x_j|^{-1}\ . \tag{6.9}$$

The first term, $\tilde{H}^{\infty}_{N,j}=\tilde{H}^{\infty}_{N,j}(X_j,\underline{R})$, satisfies (as an operator) $\tilde{H}^{\infty}_{N,j}\ge E^{\infty}_{N,j}$. Therefore, moving this term to the right-hand side of Eq. (6.5) we obtain less than or equal to 0 as before [using Eq. (6.8)].

The term involving T_j is positive as before. The T_{nuc} term is positive since $\phi(x_j,\underline{R})$, with all variables except R_i fixed, has the correct form [as given by Eq. (3.10)] as a function of R_i. The third term is (after summing on j) the same as in Eq. (5.9) but with

$$\gamma_s = \int\int |\psi(X,\underline{R})|^2 \sum_{j=1}^{N} [q_j |x_j - R_s|^{-1}/\phi(x_j,\underline{R})] \times d^{3N}X\, d^{3K}R \ . \tag{6.10}$$

In the atomic case, $\phi(x,R) = |x-R|^{-1}$ and the proof proceeds exactly as before. In the molecular case, Eq. (5.16) holds with γ_s given by Eq. (6.10). [An obvious modification of Lemma 1 and Eqs. (5.11)—(5.15) is needed to include the $\underline{R}$ dependence; the basic observation is that the Schwarz inequality used in Lemma 1 is still applicable.]

To complete the proof in the molecular case, we require that Eq. (5.20) have a solution with the new definition of γ_s, Eq. (6.10). The proof in Appendix B is easily modified in terms of the appropriate matrix M [which has strictly positive elements and satisfies Eq. (B8)]:

$$M_{sj} = Q^{-1}\int\int |\psi(X,\underline{R})|^2 \times \sum_{k=1}^{N} q_k |x_k - R_s|^{-1} |x_k - R_j|^{-1} \times \phi(x_k,\underline{R})^{-2} d^{3N}X\, d^{3K}R \ . \tag{6.11}$$

C. Hartree and Hartree-Fock theories

Theorem 1 applies to the Schrödinger equation, but the conclusion remains true in both the Hartree and Hartree-Fock (HF) (either restricted or unrestricted) approximations. Here, the proof for unrestricted HF theory will be given; the proof for the Hartree theory is very similar. It will also be assumed that all the N charges, masses, and kinetic energy operators are identical. We take $q_i = 1$.

In the HF theory, the ground-state energy is defined as in Eq. (2.7), but with ψ restricted to the class of determinantal functions:

$$\psi = (N!)^{-1/2} \det u_i(x_j,\sigma_j) \tag{6.12}$$

and the u_i are orthonormal. As in the earlier definition, the N-particle system is said to be bound if and only if there is a ψ that actually minimizes the energy expression (2.7). As before, $E_N \le E_{N,j}$. The reader is referred to Ref. 16 for details; in particular, for the proof that binding occurs if $N < Z+1$ (when $T = T^{(1)}$).

If there is a minimum, the u_i (after possibly an $N\times N$ unitary transformation) satisfy the N coupled HF equations:

$$hu_i = \epsilon_i u_i \ , \tag{6.13}$$

where h is the single-particle operator

$$h = T - V(x) + U - K \ . \tag{6.14}$$

Here, T is one of the operators in Eqs. (2.3)—(2.5), U is the direct part of the Coulomb repulsion, and K is the exchange part.

Since ψ minimizes the energy, each $\epsilon_i \le 0$ in Eq. (6.13). The reason is that the dependence of the numerator in Eq. (2.7) on any one u_i is constant plus quadratic, and the latter term is just ϵ_i. If $\epsilon_1 > 0$, for example, then it is easy to see that the $(N-1)$-particle ψ composed of $u_2, \ldots, u_N$ would have an energy strictly below E_N, and this contradicts $E_{N,1} \ge E_N$.

Let $\phi(x)$ be as in Eq. (3.11), multiply Eq. (6.13) by $u_i^*(x)/\phi(x)$, integrate over x, and sum over spins. Then sum over i. The right-hand side is nonpositive. On the left-hand side, the terms involving T are positive, as before. The $V(x)$ term is

$$A = N\int |\psi(X)|^2 V(x_1)/\phi(x_1)\, d^{3N}X \ .$$

The repulsion term, $U-K$, is

$$R = \tfrac{1}{2}N(N-1)\int |\psi(X)|^2 |x_1 - x_2|^{-1} \times [\phi(x_1)^{-1} + \phi(x_2)^{-1}]\, d^{3N}X \ .$$

(A summation on spins is understood in these two integrals.) The rest of the analysis proceeds as before.

ACKNOWLEDGMENTS

This work was partially supported by U.S. National Science Foundation Grant No. PHY-81-16101-A01.

APPENDIX A: KINETIC-ENERGY INEQUALITY

1. Proof of Eq. (3.15)

Our first goal is to prove Eq. (3.15) when T is one of the three forms in Eqs. (2.3)—(2.5). The potential ϕ is defined by

$$\phi(x) = \int |x-y|^{-1} d\mu(y) + C \ , \tag{A1}$$

with $d\mu$ a positive measure, $0 < \int d\mu < \infty$, and $C \ge 0$. The function f satisfies

$$f \text{ and } Tf \in L^2(R^3) \ . \tag{A2}$$

Note that if $C > 0$ then $1/\phi(x)$ is a bounded function and $f(x)/\phi(x)$ is automatically in L^2.

The following Lemma 2 validates the assertions used in Sec. III provided

$$\psi(X)/\phi(x_j) \in L^2(R^{3N}) \ .$$

Subsection II below shows how to deal with the case $\psi/\phi \notin L^2$ by taking $C > 0$ and then letting $C \to 0$. [Technical remark: In the nonrelativistic case a solution to Eq. (2.9) automatically satisfies $T_j\psi \in L^2(R^{3N})$ and hence $Tf \in L^2(R^3)$. The analogous statement is known to hold in the relativistic case if $e^2 q_j z_i < \frac{1}{2}$ (for all i), which is less than the critical value $2/\pi$. Thus, there is possibly a minor technical gap in the relativistic case.] Lemma 2 in the relativistic case was originally proved only for $\vec{A}(x) \equiv 0$. The extension to $\vec{A}(x) \not\equiv 0$ follows from Lemma 3, which is a joint work with Michael Loss. Lemma 3 shows that Lemma 2 (relativistic) follows automatically from Lemma 2 (nonrelativistic). Nevertheless, the original proof of Lemma 2 for $T^{(3)}$ with $\vec{A}(x) \equiv 0$ is not

without interest and is therefore given below.

Lemma 2: Assume Eqs. (A1), (A2), and

$$f(x)/\phi(x)\in L^2(\mathbb{R}^3) .$$

Let T be any one of the operators in Eqs. (2.3)—(2.5), with $\vec{A}(x)$ bounded in Eqs. (2.4) and (2.5). Then

$$h(x)\equiv f^*(x)(Tf)(x)/\phi(x) \tag{A3}$$

is in L^1 and

$$t=\mathrm{Re}\int h(x)d^3x>0 . \tag{A4}$$

Proof: At first assume $C>0$ so that $1/\phi$ is bounded. By a simple density argument, we can suppose $f\in C_0^\infty$ (infinitely differentiable functions of compact support). We can also assume [by replacing $d\mu$ by $\exp(-\epsilon x^2)*d\mu$, for example, and then using dominated convergence] that ϕ and $1/\phi\in C^\infty\cap L^\infty$. Unfortunately, the limiting argument just cited will only allow us to conclude that $t\geq 0$. The proof that $t>0$ is given at the very end.

$T^{(1)}=-\Delta$: The proof given in Eqs. (3.16) and (3.17) is completely rigorous for such functions. Another proof is given below in connection with $T^{(3)}[\vec{A}(x)\equiv 0]$.

$T^{(2)}=[i\vec{\nabla}+\vec{A}(x)]^2$: We set $g(x)=f(x)/\phi(x)$ as before and follow the same manipulation as in Eq. (3.17), but replacing $\vec{\nabla}$ by $\vec{\nabla}-i\vec{A}(x)$. The only essential difference is that we get an extra term

$$i\int \vec{\nabla}\phi(x)\cdot\vec{A}(x)\,|g(x)|^2d^3x ,$$

but this vanishes when the real part is taken. Strict positivity follows from the fact that $(i\vec{\nabla}-\vec{A})g$ cannot vanish identically.

$T^{(3)}=(-\Delta+1)^{1/2}-1$: Here we have to work in momentum space. $T^{(3)}f$ is defined to be the function whose Fourier transform $(T^{(3)}f)\hat{}$ is

$$[(k^2+1)^{1/2}-1]\hat{f}(k) ,$$

with

$$\hat{f}(k)=\int \exp(ik\cdot x)f(x)d^3x .$$

Again, let $g=f/\phi$. Since $(ab)\hat{}=\hat{a}*\hat{b}$, and since $(|x|^{-1})\hat{}=4\pi/k^2$, we have

$$2\pi^2 t=\mathrm{Re}\int d\mu(y)\int \hat{g}^*(k)\,|k-q|^{-2}e^{iy\cdot(k-q)} \times m(q)\hat{g}(q)d^3k\,d^3q + (C/4\pi)\int |g(k)|^2 m(k)d^3k , \tag{A5}$$

with $m(q)=(q^2+1)^{1/2}-1$. Clearly, the second term on the right-hand side of Eq. (A5) is positive. As for the first term, it is sufficient to prove strict positivity for each y but, since

$$\exp[iy\cdot(k-q)]$$

is a product function, it suffices to prove that the kernel

$$K(k,q)=|k-q|^{-2}[m(k)+m(q)] \tag{A6}$$

is positive definite.

Let us temporarily return to $T^{(1)}=-\Delta$. In this case $m(k)$ in Eq. (A6) is replaced by k^2. However,

$$k^2+q^2=|k-q|^2+2k\cdot q \tag{A7}$$

and thus

$$K(k,q)=1+2k\cdot q\,|k-q|^{-2} . \tag{A8}$$

The kernel 1 in Eq. (A8) is clearly positive semidefinite. For the second term, note that $|k-q|^{-2}$ is positive definite and $k\cdot q$ is a product function; thus, K is positive definite and we have a second proof for the $T^{(1)}$ case.

Returning to $T^{(3)}$, let us write

$$K(k,q)=[n(k)+n(q)]^{-1}\,|k-q|^{-2}[n(k)+n(q)] \times[m(k)+m(q)] , \tag{A9}$$

with $n(k)=(k^2+1)^{1/2}$. The last two factors in Eq. (A9) are

$$B(k,q)=k^2+q^2+2[n(k)-1][n(q)-1] . \tag{A10}$$

Now the first factor can be written as an integral over product functions, namely,

$$\int_0^\infty ds\,\exp\{-s[n(k)+n(q)]\} ,$$

and this does not affect the positive definiteness. As for the rest we have two terms, namely,

$$|k-q|^{-2}(k^2+q^2) ,$$

which is positive definite as we just proved for $T^{(1)}$, and

$$|k-q|^{-2}[n(k)-1][n(q)-1] .$$

But this last term is a positive definite kernel times a product function, so it too is positive definite. This concludes the proof that $t>0$ if $C>0$ and $\vec{A}(x)\equiv 0$ for $T^{(3)}$.

Thus far we have proved that $t>0$ when (i) $f\in C_0^\infty$, $\phi\in C^\infty$ and (ii) $C>0$. In all cases we found that $t=Q(g,g)$ with Q being a positive definite quadratic form and $g=f/\phi$. Let us first remove condition (i). There exist sequences f^n and $d\mu^n$ such that $\phi^n\to\phi$ pointwise almost everywhere and $f^n\to f$, $Tf^n\to Tf$, and

$$g^n=f^n/\phi^n\to g=f/\phi$$

in L^2. Then

$$t=\lim_{n\to\infty}t^n=\lim_{n\to\infty}Q(g^n,g^n)\geq Q(g,g)>0 .$$

Finally, we want to let $C\to 0$ (if that is the case at hand). With f and ϕ fixed, let $\phi_C=\phi+C$, $g_C=f/(\phi+C)$. With $Y_C=\phi/(\phi+C)$, we have that $0<Y_C\leq 1$, and $Y_C\to 1$ pointwise almost everywhere. Then $t_C\to t$ by dominated convergence. Also, $g_C\to g=f/\phi$ and $\hat{g}_C\to\hat{g}$ in L^2. Again,

$$t=\lim_{C\to 0}t_C=\lim_{C\to 0}Q(g_C,g_C)\geq Q(g,g)>0 .$$

This completes the proof for $T^{(1)}$, $T^{(2)}$, and $T^{(3)}$ $[\vec{A}(x)\equiv 0]$. The general case $T^{(3)}$ $[\vec{A}(x)\not\equiv 0]$ follows from the $T^{(2)}$ case and Lemma 3, in which C is the multiplication operator $1/\phi(x)$ and B is the operator $T^{(2)}$, and the integral representation

$$(B+1)^{1/2}-1=\pi^{-1}B\int_1^\infty (x-1)^{1/2}x^{-1}(x+B)^{-1}dx\ , \tag{A11}$$

Q.E.D.

Lemma 3 (in collaboration with M. Loss): Let H be a Hilbert space with an inner product $(\cdot,\cdot)$ and let B and C be non-negative, self-adjoint linear operators with domain $D(B),D(C)$. Suppose that

(i) $(B+x)^{-1}$: $D(C)\to D(C)$, all $x>0$

(ii) $\mathrm{Re}(B\phi,C\phi)>0$ (respectively, ≥ 0), all $0\neq\phi\in D(B)\cap D(C)$

(iii) $$g(\lambda)=\int d\mu(y)(s\lambda+t)(\lambda+y)^{-1} \tag{A12}$$

with $s,t\geq 0$, $s+t>0$, and $\mu\not\equiv 0$ a non-negative Borel measure on R with $\mu\{(-\infty,0]\}=0$ and

$$\int d\mu(y)(1+y)^{-1}<\infty\ .$$

Then

$$\mathrm{Re}(g(B)\phi,\ C\phi)>0 \text{ (respectively, } \geq 0), \text{ all } \phi\in D(g(B))\cap D(C)\ . \tag{A13}$$

Proof: First, consider the special case

$$g(\lambda)=g_x(\lambda)=(s\lambda+t)(\lambda+x)^{-1}$$

for some fixed $x>0$. Then $G_x=g_x(B)$ is bounded and we want to prove that $I=(G_x\phi,C\phi)$ satisfies $\mathrm{Re}I>0$ (respectively, ≥ 0) for all $0\neq\phi\in D(C)$. Let $\psi=(B+x)^{-1}\phi$, whence

$$0\neq\psi\in D(B)\cap D(C)\ .$$

Since $(B+x)\psi\in D(C)$ then ψ and $B\psi\in D(C)$. Thus,

$$I=((sB+t)\psi,\ C(B+x)\psi)=s(B\psi,\ CB\psi)+sx(B\psi,\ C\psi)+t(\psi,\ CB\psi)+tx(\psi,\ C\psi)\ .$$

Since $(\psi,\ CB\psi)=(C\psi,\ B\psi)=(B\psi,\ C\psi)^*$, we have

$$\mathrm{Re}I\geq(sx+t)\mathrm{Re}(B\psi,\ C\psi)>0 \text{ (respectively, } \geq 0)\ .$$

Now let $I=(g(B)\phi,\ C\phi)$ with

$$\phi\in D(g(B))\cap D(C)$$

and, for $0<\epsilon<1$,

$$I_\epsilon=(g^\epsilon(B)\phi,\ C\phi)$$

with

$$g^\epsilon(\lambda)=\int d\mu^\epsilon(y)(s\lambda+t)/(\lambda+y)$$

and where μ^ϵ is μ restricted to the interval $(\epsilon,1/\epsilon)$. Clearly, $g^\epsilon(B)$ is bounded and

$$I_\epsilon=\int d\nu(\phi,C\phi;\lambda)\int d\mu^\epsilon(y)(s\lambda+t)/(\lambda+y)\ ,$$

where $\nu(\phi,\phi';\cdot)$ is the spectral measure of B associated with ϕ,ϕ'. We want to show that

$$I_\epsilon=\int d\mu^\epsilon(y)\int d\nu(\phi,C\phi;\lambda)(s\lambda+t)/(\lambda+y)\ .$$

Fubini's Theorem cannot be used since ν is not positive. However, by the polarization identity,

$$d\nu(\phi,C\phi;\lambda)=\tfrac{1}{4}\,d\nu(a,a;\lambda)-\tfrac{1}{4}\,d\nu(b,b;\lambda)+\tfrac{1}{4}id\nu(c,c;\lambda)-\tfrac{1}{4}id\nu(d,d;\lambda)\ ,$$

where $a=\phi+C\phi$, $b=\phi-C\phi$, $c=\phi-iC\phi$, and $d=\phi+iC\phi$. Each of these four measures is non-negative and, since $g^\epsilon(B)$ is bounded, each integral is finite. Thus, we can exchange the order of integration and

$$I_\epsilon=\int d\mu^\epsilon(x)M(x)$$

with

$$M(x)=(G_x\phi,C\phi)$$

and $\mathrm{Re}M(x)>0$ (respectively, ≥ 0). Therefore, $\lim_{\epsilon\to 0}\mathrm{Re}I_\epsilon>0$ (respectively, ≥ 0). On the other hand,

$$I-I_\epsilon=([g(B)-g^\epsilon(B)]\phi,\ C\phi)\ .$$

Since $\phi\in D(g(B))$, it is easy to see that $[g(B)-g_\epsilon(B)]\phi\to 0$. Thus,

$$I=\lim_{\epsilon\to 0}I_\epsilon\ .$$

Q.E.D.

Remark: Suppose that $\tilde{g}(\lambda)$ is another function with the same kind of representation as in Eq. (A12). Then, starting with Eq. (A13) and with the pair $g(B),C$ instead of B,C, one can apply Lemma 3 to $\tilde{g}(C)$ and deduce that

$$\mathrm{Re}(g(B)\phi,\ \tilde{g}(C)\phi)>0 \text{ (respectively, } \geq 0), \text{ all } \phi\in D(g(B))\cap D(\tilde{g}(C))\ . \tag{A14}$$

It is merely necessary to verify that for all $x>0$,

$$(C+x)^{-1}\text{: } D(g(B))\to D(g(B))\ .$$

This implies the following generalization of the results of this paper:

(i) The relativistic kinetic energy (with magnetic field) can be generalized to any function g of $[\vec{p}-\vec{A}(x)]^2$ that has the form of Eq. (A12).

(ii) The Coulomb potential $1/|x|$ can be replaced (everywhere) in Eq. (2.1) by $v(x)=1/w(|x|)$ for any function w with the representation

$$w(|x|)=\int d\mu(y)(s|x|)(|x|+y)^{-1}\ . \tag{A15}$$

With $s>0$ and $\mu\geq 0$. For example, $1/|x|\to|x|^{-p}$, $0<p<1$ is allowed. It is easy to check that $C=w(|x|)$ satisfies

$$(C+\lambda)^{-1}\text{: } D(T^{(i)})\to D(T^{(i)}),\ i=1,2,3$$

for $\lambda>0$. It is also necessary to check that the "triangle inequality"

$$w(|x|)+w(|z|)\geq w(|x-z|)$$

holds, and this is easily seen to be the case from Eq. (A15). [It is the requirement of the triangle inequality that dictates $s|x|$, instead of $s|x|+t$, in Eq. (A15).]

2. Eliminating infinity

After Eq. (3.5) we made the assumption that all quantities in Eqs. (3.1)—(3.5) were finite. Conceivably, this need not be true with ϕ given by Eq. (3.11). To remedy this defect replace ϕ in Eq. (3.11) by

$$\phi_C(x)=\phi(x)+C, \quad C>0 .$$

Then all quantities are finite. Denote R (respectively, A) with ϕ_C by R_C (respectively, A_C). Binding cannot occur if $R_C>A_C$ for any C (here, the fact that $t>0$ is ignored).

As $C\to 0$, R_C and A_C have finite limits R and A which, by dominated convergence, are the R and A given in Eq. (3.14). Thus, it suffices to show that $R>A$ when condition (2.10) or (2.11) is violated. In the earlier proof in Secs. IV and V it was shown that $R\geq A$ by using the triangle inequality

$$|x_i-x_j|\leq |x_i|+|x_j| .$$

This inequality will now be investigated more closely to show that, in fact, $R>A$.

If we look at Eq. (5.11), for example, we have, after integrating over the variables other than $x_i=x$ and $x_j=y$, an expression of the form

$$L\equiv \int f(x,y)(|x|+|y|)|x-y|^{-1}d^3x\,d^3y , \tag{A16}$$

and we wish to show that

$$L>M\equiv \int f(x,y)d^3x\,d^3y>0 . \tag{A17}$$

Note that $f(x,y)$ is a non-negative *function* in L^1, and not a distribution. The function

$$g(x,y)=(|x|+|y|)|x-y|^{-1}-1$$

satisfies $g\geq 0$ and $g=0$ if and only if $y=-bx$ with $b\geq 0$. The set on which this occurs has six-dimensional Lebesgue measure zero. Thus, $g>0$ almost everywhere. Since $f>0$ on a set of positive measure, $\int fg>0$ and hence Eq. (A17) holds.

APPENDIX B: SOLUTION OF EQ. (5.20)

Let μ denote $(\mu_1,\ldots,\mu_K)$ and consider the function

$$F(\mu)=\sum_{s=1}^{K}[\delta_s(\mu)-\beta_s]^2 \tag{B1}$$

defined on the positive orthant D: $\mu_i\geq 0$, but excluding the origin $\mu=0$. The β_s are fixed, strictly *positive* constants satisfying $\sum_s \beta_s=1$. δ_s is of the form

$$\delta_s=\int \rho(x)[\mu_s|x-R_s|^{-1}/\phi(x)]d^3x , \tag{B2}$$

with $\int \rho=1$, and

$$\phi(x)=\sum_s \mu_s|x-R_s|^{-1} .$$

Equation (B2) implies $\sum \delta_s=1$.

Now $\delta_s(\mu)$ is continuous on D (in particular, $\delta_s(\mu)=0$ if $\mu_s=0$) and homogeneous of degree zero, i.e., $\delta_s(\lambda\mu)=\delta_s(\mu)$. Therefore, $F(\mu)$ has a minimum on D. We want to show that this minimum is zero, whence $\delta_s(\mu)=\beta_s$ for all s. Let μ be a minimum point and suppose (without loss of generality) that $\mu_1,\ldots,\mu_t>0$, $\mu_{t+1},\ldots,\mu_K=0$. (Not all the μ_s can vanish since 0 is not in D.) For $1\leq s\leq t$, $\delta_s(\mu)$ is differentiable in $\mu_1,\ldots,\mu_t$ and for $t+1\leq s\leq K$, $\delta_s(\mu)\equiv 0$ in a (t-dimensional) neighborhood of this point. Therefore, at the minimum,

$$0=\frac{\partial F}{\partial \mu_j}=2\sum_{s=1}^{t}\mu_s M_{sj}(\delta_j-\beta_j-\delta_s+\beta_s) \tag{B3}$$

for $1\leq j\leq t$, where M is the t-square matrix

$$M_{sj}=\int \rho(x)|x-R_s|^{-1}|x-R_j|^{-1}\phi(x)^{-2}d^3x . \tag{B4}$$

Clearly, M is symmetric and positive semidefinite and, most importantly, M has positive matrix elements.

Eq. (B3) can be rewritten in the following way (recalling that $\mu_s>0$ for $1\leq s\leq t$):

$$Nv=v , \tag{B5}$$

where N is a matrix and v is a vector given by

$$v_s=(\delta_s-\beta_s)\delta_s^{1/2} , \tag{B6}$$

$$N_{sj}=M_{sj}\mu_s\mu_j(\delta_s\delta_j)^{-1/2} . \tag{B7}$$

The fact that

$$\sum_s \mu_s\mu_j M_{sj}=\delta_j \tag{B8}$$

has been used which, in terms of N, reads

$$Nw=w , \tag{B9}$$

with

$$w_s=\delta_s^{1/2} . \tag{B10}$$

Now N is symmetric and has strictly positive matrix elements. By the Perron-Frobenius theorem, N has a *unique* eigenvalue of largest modulus λ. Moreover, this eigenvalue is positive and has only one eigenvector u, which (up to a phase) has *strictly positive* components. Equation (B9) implies that $\lambda=1$ and $u=w$ for, otherwise, taking the inner product of Eq. (B9) with u we would obtain $(\lambda-1)(u,w)=0$, which is impossible since $(u,w)>0$. Thus, the solution to Eq. (B5) is

$$v=cw , \tag{B11}$$

where c is a constant. This means that

$$\delta_s-\beta_s=c$$

for

$$1\leq s\leq t . \tag{B12}$$

Summing this on s we obtain (since $\delta_s=0$ for $s>t$)

$$1-\sum_{s=1}^{t}\beta_s=ct . \tag{B13}$$

If $t=K$, the left-hand side of Eq. (B13) is zero and we are finished. If $t<K$, then $c>0$. In the latter case, replace $\mu_s=0$ by $\mu_s=\epsilon$ for $t<s\leq K$ and with $\epsilon>0$. It is easy to see that δ_s decreases for $1\leq s\leq t$ and δ_s becomes strictly positive for $t<s\leq K$. If ϵ is small enough, $(\delta_s-\beta_s)^2$ will decrease for all $1\leq s\leq K$. Thus, $F(\mu)$ will decrease; as this is a contradiction, $t=K$ and the proof is complete.

[1]M. B. Ruskai, Commun. Math. Phys. 82, 457 (1982).
[2]I. M. Sigal, Commun. Math. Phys. 85, 309 (1982). See also *Mathematical Problems in Theoretical Physics,* Vol. 153 of *Lecture Notes in Theoretical Physics* (Springer, Berlin, 1982), pp. 149–156.
[3]I. M. Sigal, Institute Mittag Leffler Report No. 12 (1982, revised in 1983), Ann. Phys. (to be published).
[4]M. B. Ruskai, Commun. Math. Phys. 85, 325 (1982).
[5]R. Benguria and E. H. Lieb, Phys. Rev. Lett. 50, 1771 (1983).
[6]B. Baumgartner, J. Phys. A (to be published).
[7]R. Benguria and E. H. Lieb, (unpublished).
[8]E. H. Lieb, I. M. Sigal, B. Simon, and W. E. Thirring (unpublished). See also Phys. Rev. Lett. 52, 994 (1984).
[9]R. N. Hill, *Mathematical Problems in Theoretical Physics. Proceedings of the International Conference on Mathematical Physics, Lausanne, 1979,* Vol. 116 of *Lecture Notes in Physics* edited by K. Osterwalder (Springer, New York, 1979), pp. 52–56.
[10]G. Zhislin, Trudy, Mosk. Mat. Obšč. 9, 81 (1960).
[11]B. Baumgartner, Lett. Math. Phys. 7, 439 (1983).
[12]F. H. Stillinger and D. K. Stillinger, Phys. Rev. A 10, 1109 (1974); F. H. Stillinger, J. Chem. Phys. 45, 3623 (1966).
[13]E. H. Lieb, Phys. Rev. Lett. 52, 315 (1984).
[14]E. H. Lieb, Rev. Mod. Phys. 53, 603 (1981). Erratum: 54, 311(E) (1982).
[15]I. Daubechies and E. H. Lieb, Commun. Math. Phys. 90, 511 (1983).
[16]E. H. Lieb and B. Simon, J. Chem. Phys. 61, 735 (1974); Commun. Math. Phys. 53, 185 (1977). Theorem 2.4 states, inter alia, that $\epsilon_1, \ldots, \epsilon_N$ are the N lowest points of the spectrum of h. This is correct, but the proof is wrong. A correct proof is given in the earlier summary: E. H. Lieb, in *Proceedings of the International Congress of Mathematicians, Vancouver, 1974,* Vol. 2, pp. 383–386.