

One-Electron Relativistic Molecules with Coulomb Interaction

Ingrid Daubechies[†]* and Elliott H. Lieb[†]

Departments of Mathematics and Physics, Princeton University, Princeton, NJ 08544, USA

Abstract. As an approximation to a relativistic one-electron molecule, we study the operator $H = (-\Delta + m^2)^{1/2} - e^2 \sum_{j=1}^K Z_j |x - R_j|^{-1}$ with $Z_j \geq 0$, $e^{-2} = 137.04$. H is bounded below if and only if $e^2 Z_j \leq 2/\pi$, all j . Assuming this condition, the system is unstable when $e^2 \sum Z_j > 2/\pi$ in the sense that $E_0 = \inf \text{spec}(H) \rightarrow -\infty$ as the $R_j \rightarrow 0$, all j . We prove that the nuclear Coulomb repulsion more than restores stability; namely $E_0 + 0.069e^2 \sum_{i < j} Z_i Z_j |R_i - R_j|^{-1} \geq 0$. We also show that E_0 is an increasing function of the internuclear distances $|R_i - R_j|$.

Introduction

The problem of “stability of matter” consists in proving that a system of charged particles (electrons and nuclei), interacting electromagnetically, does not collapse. In the framework of nonrelativistic Schrödinger quantum mechanics, with Coulomb interactions between the particles, a first proof of this was given by F. Dyson and A. Lenard [1]. A shorter proof, leading to a much better lower bound on the binding energy per electron, was later given by E. Lieb and W. Thirring [2]. The strategy these proofs followed was first to consider the nuclei fixed (i.e. with infinite mass); the general case (nuclei with finite mass) then follows easily. With the K nuclei at fixed positions R_1, \dots, R_K , the problem then consists in proving that 1) the Hamiltonian describing N electrons and K nuclei (including also the repulsion terms between the electrons, and between the nuclei) is bounded below by a constant independent of the R_j ; 2) the energy per particle, i.e. the ground state energy of the system of N

[†] Work partially supported by U.S. National Science Foundation grant PHY-8116101-A01

* On leave from Vrije Universiteit Brussel, Belgium, and from Interuniversitair Instituut voor Kernwetenschappen, Belgium

electrons and K nuclei, divided by $(N + K)$, is bounded below by some constant.

If electrons were bosons, it is known [3] that matter would not be stable: statement 1) above still holds for bosons, but 2) does not.

As a first approximation to a relativistic approach to the problem of stability of matter, we shall study stability for the Hamiltonian

$$\sum_{j=1}^N (-\hbar^2 c^2 \Delta_j + m^2 c^4)^{1/2} - \sum_{j=1}^N \sum_{k=1}^K Z_k e^2 |x_j - R_k|^{-1} + \sum_{\substack{i,j=1 \\ i < j}}^N e^2 |x_i - x_j|^{-1} \\ + \sum_{\substack{k,l=1 \\ k < l}}^K Z_k Z_l e^2 |R_k - R_l|^{-1} = H_{N,K}(m, \underline{Z}, \underline{R}), \quad (1.1)$$

where m is the electron mass, and where again the K nuclei are fixed at the distinct positions R_1, \dots, R_K . We shall use the shorthand notation $\underline{Z}, \underline{R}$ for the sets $\{Z_j\}_{j=1}^K, \{R_j\}_{j=1}^K$.

To simplify our expressions, we rescale the x_j and R_k in units of the Compton wave length of the electron \hbar/mc . We also rescale the Hamiltonian by mc^2 . In terms of the variables $\tilde{x}_j = (mc/\hbar)x_j$, and with $\tilde{R}_k = (mc/\hbar)R_k$, $\tilde{H}_{N,K} = (mc^2)^{-1}H_{N,K}$ becomes

$$\tilde{H}_{N,K}(\underline{Z}, \underline{R}) = \sum_{j=1}^N (-\tilde{\Delta}_j + 1)^{1/2} - \sum_{j=1}^N \sum_{k=1}^K Z_k \alpha |\tilde{x}_j - \tilde{R}_k|^{-1} \\ + \sum_{\substack{i,j=1 \\ i < j}}^N \alpha |\tilde{x}_i - \tilde{x}_j|^{-1} + \sum_{\substack{k,l=1 \\ k < l}}^K Z_k Z_l \alpha |\tilde{R}_k - \tilde{R}_l|^{-1},$$

where $\alpha = e^2/\hbar c \simeq (137.04)^{-1}$.

It is obvious from the definition (1.1) that $H_{N,K}(m, \underline{Z}, \underline{R})$ is bounded below by $H_{N,K}(0, \underline{Z}, \underline{R})$. This will turn out to be very useful in some of our proofs below. We therefore introduce also a rescaled version of $H_{N,K}(0, \underline{Z}, \underline{R})$, with a different scaling, since the Compton wave length is infinity if $m = 0$. We rescale $H(0, \underline{Z}, \underline{R})$ in units $\hbar c/R_0$, where R_0 is an arbitrary length, and scale x and the R_j by R_0 . This yields

$$h_{N,K}(\underline{Z}, \underline{R}) = \sum_{j=1}^N (-\tilde{\Delta}_j)^{1/2} - \sum_{j=1}^N \sum_{k=1}^K Z_k \alpha |\tilde{x}_j - \tilde{R}_k|^{-1} \\ + \sum_{\substack{i,j=1 \\ i < j}}^N \alpha |\tilde{x}_i - \tilde{x}_j|^{-1} + \sum_{\substack{k,l=1 \\ k < l}}^K Z_k Z_l \alpha |\tilde{R}_k - \tilde{R}_l|^{-1},$$

where again the tildes indicate the scaled variables. We shall henceforth always work with these two scaled Hamiltonians, and omit the tildes.

It turns out that, unlike the Schrödinger case, even statement 1) above (i.e. the existence of an R -independent lower bound for the ground state energy of $H_{N,K}$) is not straightforward for $H_{N,K}$ even if we restrict ourselves to the case $N = 1$ (1 electron). First of all, there exists only a limited range of the Z_j for which $H_{N,K}$ is bounded below at all. This can already be seen for the simplest case

$$K = 1, N = 1,$$

$$H_{1,1}(Z) = (-\Delta + 1)^{1/2} - Z\alpha|x|^{-1}. \quad (1.2)$$

This operator was studied independently by R. Weder [4] and I. Herbst [5]. A first important fact (which can be found in Kato [6]) is the existence of a critical charge Z_c for the operator $H_{1,1}$, exactly as for the Klein–Gordon or Dirac theories. To be explicit, let

$$H_0 = (-\Delta + 1)^{1/2}, \quad Z_c = 2/(\alpha\pi) \simeq 87.2.$$

Then a) for $Z \leq Z_c$: $H_{1,1} \geq 0$ as a form on $Q(H_0)$.

b) for $Z > Z_c$: $H_{1,1}$ is unbounded below as a form on $Q(H_0)$.

Here $Q(H_0)$ is the form domain of H_0 , which consists of all the functions f in $L^2(\mathbb{R}^3)$ for which $|p|^{1/2}\hat{f}(p)$ is square integrable, where \hat{f} denotes the Fourier transform of f .

If Z is strictly subcritical (i.e. $Z < Z_c$), more information concerning the spectrum of $H_{1,1}(Z)$ is known [4, 5]. The eigenvalues of $H_{1,1}(Z)$ all lie between 0 and 1; as a matter of fact, they are even separated from 0 by a gap which increases with $Z_c - Z$.

$$\sigma_{\text{disc}}(H_{1,1}) \subset \left[\left[1 - (Z/Z_c)^2 \right]^{1/2}, 1 \right). \quad (1.3)$$

Since $(1 - \Delta)^{1/2} \leq 1 - \Delta/2$, and since $-\Delta/2 - Ze^2|x|^{-1}$ has infinitely many negative eigenvalues, it follows that $H_{1,1}(Z)$ has infinitely many eigenvalues smaller than 1. It turns out, from an argument we give in Sect. 3, that the lowest eigenvalue is nondegenerate, and that the associated ground state is strictly positive.

On the other hand the essential spectrum always starts at 1, and it consists of only an absolutely continuous spectrum: $\sigma_{\text{ess}}(H_{1,1}) = \sigma_{\text{abscont}}(H_{1,1}) = [1, \infty)$.

The operator $h_{1,1}(Z) = (-\Delta)^{1/2} - Z\alpha|x|^{-1}$ has the same critical Z as $H_{1,1}$: it is bounded below by 0 on $Q(|p|)$ if $Z \leq Z_c$, and unbounded below if $Z > Z_c$. Its spectrum is much simpler however: $\sigma(h_{1,1}) = \sigma_{\text{abscont}}(h_{1,1}) = [0, \infty)$.

One can easily extend the proofs in [4, 5] to show that $H_{N,K}$ is bounded below if and only if $Z_j\alpha \leq 2/\pi$ for all j (we assume that all the R_j are distinct). From now on we shall only consider this case. Even under this condition, however, it is not obvious that an \underline{R} -independent lower bound for $H_{N,K}$ exists.

Let us consider the case $N = 1$:

$$H_{1,K} = H_{1,K}^0 + \sum_{\substack{k,l=1 \\ k < l}}^K Z_k Z_l \alpha |R_k - R_l|^{-1}, \quad (1.4)$$

where

$$H_{1,K}^0 = (-\Delta + 1)^{1/2} - \sum_{k=1}^K Z_k \alpha |x - R_k|^{-1} \quad (1.5)$$

is the 1-electron Hamiltonian without nuclear repulsion. Suppose that $\sum_{k=1}^K Z_k \alpha > 2/\pi$ (this is possible when $K \geq 2$). It is a simple consequence of the unboundedness below of $H_{1,1}(Z)$ for $Z\alpha > 2/\pi$ that the ground state energy for $H_{1,K}^0$ will tend to $-\infty$ if the internuclear distances shrink to zero:

$$\liminf_{\lambda \rightarrow 0} \text{spec } H_{1,K}^0(\underline{Z}, \lambda \underline{R}) = -\infty, \quad \sum_{k=1}^K Z_k \alpha > 2/\pi.$$

Under the same conditions however, the nuclear repulsion term in (1.4) tends to $+\infty$, and this may “cancel” the behavior of $E_0 = \inf \text{spec } H_{1,K}^0$. The existence of an R -independent lower bound for $H_{1,K}(Z, R)$ is therefore only possible if the nuclear repulsion is strong enough to overcome the “collapsing tendency” present in $H_{1,K}^0$. The fact that this is the case is the main content of the following theorem:

Theorem 1. *Let $h_0 = (-\Delta)^{1/2}$. Suppose $z_k \leq 2/\pi$ for all k , and $R_k \neq R_l$ for all $k \neq l$.*

Then $h_0 - \sum_{k=1}^K z_k |x - R_k|^{-1}$, considered as a form on $Q(h_0)$, is bounded below by

$$-3\pi \sum_{\substack{k,l=1 \\ k < l}}^K z_k z_l |R_k - R_l|^{-1}.$$

Since $(-\Delta + 1)^{1/2} \geq (-\Delta)^{1/2}$, the theorem obviously also holds if we replace h_0 by $H_0 = (-\Delta + 1)^{1/2}$.

In terms of the fine-structure constant α and the nuclear charges $Z_k = z_k/\alpha$, Theorem 1 can be rewritten as

$$E_0 + \sum_{\substack{k,l=1 \\ k < l}}^K Z_k Z_l \alpha |R_k - R_l|^{-1} \geq (1 - b) \sum_{\substack{k,l=1 \\ k < l}}^K Z_k Z_l \alpha |R_k - R_l|^{-1}, \tag{1.6}$$

where $b = 3\pi\alpha \leq 6.88 \times 10^{-2}$. The fact that b is strictly smaller than 1 implies that the total energy (including the nuclear repulsion) increases to $+\infty$ if any two nuclei coincide. Thus the nuclear repulsion is not only strong enough to prevent collapse, it even pushes the nuclei apart!

For

$$h_{1,K}(Z, R) = h_{1,K}^0(Z, R) + \sum_{\substack{k,l=1 \\ k < l}}^K Z_k Z_l \alpha |R_k - R_l|^{-1}, \tag{1.7}$$

with

$$h_{1,K}^0(Z, R) = (-\Delta)^{1/2} - \sum_{k=1}^K Z_k \alpha |x - R_k|^{-1}, \tag{1.8}$$

we can say even more. By scaling one sees that $h_{1,K}(Z, R)$ and $\lambda h_{1,K}(Z, \lambda R)$ are unitarily equivalent. Theorem 1 implies therefore that in this case the total energy is minimal if the nuclei are infinitely far apart.

Incidentally, note that the exact numerical value of α plays a decisive role. This can be understood by the scaling behavior of $H_{1,K}$. By simple scaling one sees that $H_{1,K}(Z, R)$ is unitarily equivalent to $\lambda H_{1,K}(\lambda^{-1}Z, \lambda R)$, where we define

$$H_{1,K}(\mu; Z, R) = (-\Delta + \mu^2)^{1/2} - \sum_{k=1}^K Z_k \alpha |x - R_k|^{-1} + \sum_{\substack{k,l=1 \\ k < l}}^K Z_k Z_l \alpha |R_k - R_l|^{-1}.$$

(The same is true for $H_{N,K}$.) This means essentially that we can scale the mass away: the existence of any R -independent bound for $H_{1,K}(1; Z, R)$ implies that the zero-mass Hamiltonian $H_{1,K}(0; Z, R) = h_{1,K}(Z, R)$ is bounded below by zero, independently of R . (Since $H_{1,K} \geq h_{1,K}$, the converse statement is trivial.)

This, in turn, implies that the value of α is decisive. Let us illustrate this for the special case $K = 2$, $Z_1 \alpha = Z_2 \alpha = 2/\pi$. On the one hand, if $H_{1,2}(Z, R)$ is bounded

below by an R -independent constant, then $h_{1,2}(Z, R)$ is bounded below by zero. Thus, for all ψ in $Q[(-\Delta)^{1/2}]$,

$$\langle \psi, h_{1,2}^0(Z, R)\psi \rangle + (4/\pi^2\alpha)|R_1 - R_2|^{-1} \|\psi\|^2 \geq 0$$

($h_{1,2}^0$ is $h_{1,2}$ without the nuclear repulsion; see (1.8)). On the other hand, by a scaling argument,

$$\inf \text{spec} [h_{1,2}^0(Z, R)] = -C|R_1 - R_2|^{-1}$$

for some positive constant C , which is completely independent of α . It is here that the numerical value of α comes into play: if α is such that $4/(\alpha\pi^2) \geq C$, then the required bound will hold (for $K = 2$); if $4/(\alpha\pi^2) < C$, the one-electron-two-nucleus system is definitely unstable in this model. By the argument above, this conclusion will hold as well for $H_{1,K}$, i.e. for any positive mass, as for $h_{1,K}$, which has zero mass.

We prove Theorem 1 in Sect. 2.

In Sect. 3 we show that, for one electron, $E_{1,K}^0 = \inf \text{spec} H_{1,K}^0(Z, R)$, and $e_{1,K}^0 = \inf \text{spec} h_{1,K}^0(Z, R)$ are monotone increasing functions of the internuclear distances:

Theorem 2. *Keep the Z_k fixed, where $Z_k\alpha \leq 2/\pi$ for all k . If for all k, l : $|R_k - R_l| \leq |R'_k - R'_l|$, then $E_{1,K}^0(Z, R) \leq E_{1,K}^0(Z, R')$ and $e_{1,K}^0(Z, R) \leq e_{1,K}^0(Z, R')$.*

Our proof borrows its basic ideas from [7] and [8], where the analogous statement for the Schrödinger operator $-\Delta - \sum_{k=1}^K Z_k e^2 |x - R_k|^{-1}$ was proved (first for dilations, i.e. $R'_k = \lambda R_k$ with $\lambda \geq 1$, in [7]; this proof was extended to the general situation $|R'_k - R'_l| \geq |R_k - R_l|$ in [8]).

If all the Z_k are strictly subcritical, i.e. $Z_k\alpha < 2/\pi$ for all k , one can extend a result of [4, 5] and show that $\sigma_{\text{ess}}(H_{1,K}^0) = [1, \infty)$ and analogously $\sigma_{\text{ess}}(h_{1,K}^0) = [0, \infty)$. For any Z (provided all the Z_k are not zero) one can find a variational function (e.g. a suitable exponential) such that $\langle \psi, H_{1,K}^0\psi \rangle < 1$. This shows that for $Z_k < Z_c$, all k , $E_{1,K}^0$ is an isolated eigenvalue of $H_{1,K}^0$. (Actually, by the same argument as for $H_{1,1}$, one sees that $H_{1,K}^0$ has infinitely many eigenvalues smaller than 1.) For $h_{1,K}^0$, matters are slightly different. If $\sum_{k=1}^K Z_k \leq Z_c$, one can easily extend arguments of [4, 5] to show that $h_{1,K}^0 \geq 0$, in which case $e_{1,K}^0 = 0$, and $h_{1,K}^0$ has no eigenvalues. If $\sum_{k=1}^K Z_k > Z_c$, one can again find a variational function ψ for which $\langle \psi, h_{1,K}^0\psi \rangle < 0$, which implies that in that case $e_{1,K}^0$ is also an isolated eigenvalue.

We prove in Sect. 3 that, if $Z_k < Z_c$ for all k , $E_{1,K}^0$ is a nondegenerate eigenvalue of $H_{1,K}^0$, and that the corresponding ground state is strictly positive. The same is true for $e_{1,K}^0$ under the additional assumption $\sum_{k=1}^K Z_k > Z_c$.

We have no results in this paper concerning $H_{N,K}$ for $N > 1$. The same scaling argument as for $H_{1,K}$ applies, and the existence of an R -independent lower

bound for $H_{N,K}$ would therefore imply that $h_{N,K}$ (i.e. the same Hamiltonian with $(-\Delta)^{1/2}$ instead of $(-\Delta + 1)^{1/2}$) is bounded below by zero. We do not know, however, whether such a lower bound exists for general N, K .

Regardless of the statistics, and neglecting the electron repulsion, we know that if $K = 1$ (one nucleus only) $H_{N,1} \geq 0$ (because $H_{1,1} \geq 0$). On the other hand, the lower bound (1.6) shows that for $N \leq 14$ and arbitrary K , we also have $H_{N,K} \geq 0$ (because $1 - 14b \geq 0$, where $b = 3\pi\alpha$). For general values of N, K it is obvious that the Fermi statistics and the electron repulsion will have to play a role.

It is clear that the strategy of [2], which used a density functional lower bound on the kinetic energy, and then applied results of Thomas Fermi theory, will not work here. The lower bound on the kinetic energy for the operator $(-\Delta + 1)^{1/2}$ has the form [9] $\langle \psi, (-\Delta + 1)^{1/2} \psi \rangle \geq K \int d^3x F(|\psi(x)|^2)$, where $F(a) = \int_0^a du (u^{2/3} + 1)^{1/2}$. One sees that for small a , $F(a)$ behaves like $ca^{5/3} + a$, which is similar to the $\rho_\psi^{5/3}$ bound one obtains for $\langle \psi, (-\Delta)\psi \rangle$ [2], and which is caused by the fact that for small p , $(p^2 + 1)^{1/2}$ behaves like $1 + p^2/2$. For large p , however, $(p^2 + 1)^{1/2}$ behaves linearly in p , which is reflected in the $a^{4/3}$ behavior of $F(a)$ for large a . Hence the lower bound on the kinetic energy is of the form $\int d^3x \rho_\psi(x)^{4/3}$ in the region where ρ_ψ is large; the corresponding Thomas–Fermi functional (including the other contributions to the energy) is not bounded below, and therefore does not lead to a useful lower bound on the ground state energy of $H_{N,K}$.

Heuristically, one can argue that Thomas–Fermi theories are “large Z theories” (i.e. they give the correct asymptotic behavior for $Z \rightarrow \infty$; this is rigorously true for the Schrödinger case [10].) Hence it is only natural that the Thomas–Fermi theory corresponding to $H_{N,K}$ is unbounded below, since $H_{N,K}$ itself is unbounded below whenever one Z_j becomes supercritical.

Throughout this paper we consider only the three-dimensional case, i.e. the Hilbert space we use is $L^2(\mathbb{R}^3)$. For functions on \mathbb{R}^3 , the symbol $\| \cdot \|$ will be used to denote the $L^2(\mathbb{R}^3)$ -norm only; for operators, this symbol will always mean the norm as a bounded operator from $L^2(\mathbb{R}^3)$ to itself.

2. The Nuclear Repulsion Restores Stability

We shall prove Theorem 1 in two steps: first we shall prove it in the case $Z_k = Z_c$, all k . Then we shall apply a concavity argument to conclude the desired result for all Z_k . In the proof of the first step we shall need the following lemma, which can be considered as a refinement of the result $\| |x|^{-1/2} |p|^{-1/2} \| = (\pi/2)^{1/2}$. This norm was evaluated in [5], and by a different method also in [11]; the critical value $Z_c e^2 = 2/\pi$ actually has its origin in this number.

Notation. $B(a, R) = \{x \mid |x - a| \leq R\}$.

Lemma 2.1. *Let ψ be an L^2 -function with $\text{supp } \psi \subset B(0, R)$. We denote by K the bounded operator $K = 2\pi^{-1} |x|^{-1/2} |p|^{-1} |x|^{-1/2}$. Then*

$$\langle \psi, K\psi \rangle \leq \langle \psi, \psi \rangle - \pi^{-3} R^{-2} \left(\int d^3x |x|^{-1/2} |\psi|^2 \right). \tag{2.1}$$

Proof. Note first that since ψ has compact support, the integral $\int d^3x|x|^{-1/2}|\psi(x)|$ converges. By simple scaling, we can obviously assume that $R = 1$, which we shall do henceforth in this proof.

We denote the symmetric decreasing rearrangement of ψ by ψ^* . Since $|x|^{-1/2}$ is symmetric and decreasing, we have $\int \psi^*|x|^{-1/2} \geq \int |\psi||x|^{-1/2}$. Analogously, using the fact that $|p|^{-1}$ is the same as convolution with $(2\pi^2)^{-1}|x|^{-2}$, one sees that $\langle \psi^*, K\psi^* \rangle \geq \langle \psi, K\psi \rangle$ (this follows from the generalized rearrangement inequalities proved in [12].) Hence we can restrict our attention to ψ which are symmetric decreasing, i.e. $\psi = \psi^*$.

For any such $\psi(r)$, defined for $0 \leq r \leq 1$, we define, for $r \geq 1$, $\tilde{\psi}(r) = r^{-3}\psi(r^{-1})$. Consider the function $f(r) = \psi(r)$, $r \leq 1$, $f(r) = \tilde{\psi}(r)$, $r \geq 1$. One easily checks that f is in $L^2(\mathbb{R}^3)$, with $\|f\|^2 = 2\|\psi\|^2$. Straightforward calculation also leads to $\langle f, Kf \rangle = 2\langle \psi, K\psi \rangle + 2\langle \psi, L\psi \rangle$, where L has the integral kernel $L(x, y) = \pi^{-3}|x|^{-1/2}(1 + x^2y^2 - 2x \cdot y)^{-1}|y|^{-1/2}$. Since $\||x|^{-1/2}|p|^{-1/2}\| = (\pi/2)^{1/2}$, we have $\|K\| = 1$, hence $\langle f, Kf \rangle \leq \langle f, f \rangle$, which implies

$$\langle \psi, K\psi \rangle \leq \|\psi\|^2 - \langle \psi, L\psi \rangle. \tag{2.2}$$

By computing the spherical average of the kernel $L(x, y)$, we obtain

$$\begin{aligned} \langle \psi, L\psi \rangle &= (8/\pi) \int_0^1 \int_0^1 dr \int ds r^{1/2}\psi(r)s^{1/2}\psi(s)\ln[(1 + rs)/(1 - rs)] \\ &\geq (\pi^{-3}/2) [\int d^3x|x|^{-1/2}\psi(x)]^2 \min_{u \in [0, 1]} f(u), \end{aligned} \tag{2.3}$$

where $f(u) = u^{-1}\ln[(1 + u)/(1 - u)]$.

One can easily check that f attains its minimum at $u = 0$, with $f(0) = 2$. Expression (2.1) now follows immediately from (2.2) and (2.3). ■

Remark. The constant π^{-3} is probably not optimal in (2.1). We suspect that the optimal constant is given by functions behaving like $r^{-3/2}$. This is also the typical behavior of functions optimising the expectation value of $|x|^{-1/2}|p|^{-1}|x|^{-1/2}$ (see [11], [13]).

If we define $\psi_n(r) = 0$ for $r < n^{-1}$, $\psi_n(r) = r^{-3/2}$ for $n^{-1} < r \leq 1$, it turns out that

$$\lim_{n \rightarrow \infty} [\langle \psi_n, \psi_n \rangle - \langle \psi_n, K\psi_n \rangle] / [\int d^3x|x|^{-1/2}\psi_n(x)]^2 = a\pi^{-3},$$

with $a = \sum_{k=0}^{\infty} (2k + 1)^{-3} = (7/8)\zeta(3) \simeq 1.05$. We think $a\pi^{-3}$ is probably the best constant for (2.1); in any event the sharp constant lies in the interval $[\pi^{-3}, 1.05\pi^{-3}]$.

With the help of Lemma 2.1, we can now prove Theorem 1 for the case where all the Z_j are critical.

Proposition 2.2. *For all $\psi \in Q(|p|)$, $\|\psi\| = 1$:*

$$\langle \psi, |p|\psi \rangle - (2/\pi) \sum_{j=1}^K \langle \psi, |x - R_j|^{-1}\psi \rangle \geq - (12/\pi) \sum_{\substack{j,k=1 \\ j < k}}^K |R_j - R_k|^{-1}. \tag{2.4}$$

Proof. Note first that since $|x|^{-1}$ is relatively form-bounded with respect to $|p|$ (see [4–6]), the left hand side of (2.4) makes sense for all ψ in $Q(|p|)$.

Let us rewrite the desired result (2.4) as

$$\forall \psi \in Q(|p|): \langle \psi, |p|\psi \rangle \geq (2/\pi) \sum_{j=1}^K \left\langle \psi, \left[|x - R_j|^{-1} - A \sum_{k \neq j} |R_k - R_j|^{-1} \right] \psi \right\rangle. \tag{2.5}$$

For the time being, we shall not fix A . We shall determine below a value A_0 such that (2.5) holds for all $A \geq A_0$. It will turn out that $A_0 = 3$, which then implies (2.4).

For r real, we shall use the notation $r_+ = \max(0, r)$. For K arbitrary real numbers r_1, \dots, r_K , we have

$$\sum_{j=1}^K r_j \leq \left(\sum_j r_j \right)_+ \leq \sum_j (r_j)_+ \leq \left(\sum (r_j)_+^2 \right)^{1/2}.$$

Hence (2.5) will certainly be satisfied if

$$\forall \psi \in Q(|p|): \langle \psi, |p|\psi \rangle \geq \left\langle \psi, \left(\sum_{j=1}^K W_j(x) \right)^2 \psi \right\rangle, \tag{2.6}$$

where $W_j(x) = [(2/\pi)(|x - R_j|^{-1} - A \sum_{k \neq j} |R_k - R_j|^{-1})_+]^{1/2}$.

For any A , W_j has support in a ball with center R_j and radius t_j , where

$$t_j^{-1} = A \sum_{k \neq j} |R_k - R_j|^{-1}. \tag{2.7}$$

Provided A is large enough, all the balls $B(R_j, t_j)$ are disjoint: by (2.7), we have $t_j \leq A^{-1}|R_k - R_j|$ for all $k \neq j$. Therefore

$$|R_k - R_j| - (t_j + t_k) \leq (1 - 2/A)|R_k - R_j|. \tag{2.8}$$

which shows that $B(R_j, t_j) \cap B(R_k, t_k) = \emptyset$, for all $j \neq k$, if $A > 2$. From now on we shall assume $A > 2$.

For any j , let f_j be the characteristic function of $B(R_j, t_j)$. Since $W_j = W_j f_j$, and $W_j(x) \leq (2/\pi)^{1/2} |x - R_j|^{-1/2}$, we have

$$\sum_j W_j(x) = \sum_j W_j(x) f_j(x) \leq (2/\pi)^{1/2} \sum_j |x - R_j|^{-1/2} f_j(x) \equiv W(x). \tag{2.9}$$

Since $|x - R_j|^{-1/2} f_j \leq |x - R_j|^{-1/2}$, and $|x|^{-1/2} |p|^{-1/2}$ is bounded, one easily sees that $W|p|^{-1/2}$ is bounded. We shall prove now that as an operator from L^2 to L^2 , $\|W|p|^{-1/2}\| \leq 1$, provided A is larger than some constant which we shall evaluate below.

Let ψ be any function in C_0^∞ , so that $W\psi \in L^2$. Then

$$\begin{aligned} \| |p|^{-1/2} W\psi \|^2 &= (W\psi, |p|^{-1} W\psi) \\ &= (2/\pi) \sum_{j=1}^K (|x - R_j|^{-1/2} f_j \psi, |p|^{-1} |x - R_j|^{-1/2} f_j \psi) \\ &\quad + (2/\pi) \sum_{\substack{j,k=1 \\ j \neq k}}^K (f_j |x - R_j|^{-1/2} \psi, |p|^{-1} f_k |x - R_k|^{-1/2} \psi). \end{aligned} \tag{2.10}$$

The “diagonal” terms in (2.10) can be bounded above by applying Lemma 2.1:

$$\begin{aligned}
 & (2/\pi) \sum_{j=1}^K (f_j \psi, |x - R_j|^{-1/2} |p|^{-1} |x - R_j|^{-1/2} f_j \psi) \\
 & \leq \sum_{j=1}^K \|f_j \psi\|^2 - \pi^{-3} \sum_{j=1}^K t_j^{-2} m_j^2,
 \end{aligned} \tag{2.11}$$

where

$$m_j = \int d^3x |x - R_j|^{-1/2} f_j |\psi|. \tag{2.12}$$

To find an upper bound for the “non-diagonal” terms in (2.10) we use (2.8) and the fact that $|p|^{-1}$ is the same as convolution with $(2\pi^2)^{-1} |x|^{-2}$. We obtain

$$\begin{aligned}
 & (2/\pi) (f_j |x - R_j|^{-1/2} \psi, |p|^{-1} f_k |x - R_k|^{-1/2} \psi) \\
 & \leq \pi^{-3} [1 - 2/A]^{-2} |R_k - R_j|^{-2} m_j m_k,
 \end{aligned} \tag{2.13}$$

where m_j is defined by (2.12).

Combining (2.11) and (2.13) we obtain

$$\begin{aligned}
 & \| |p|^{-1/2} W \psi \|^2 \leq \| \psi \|^2 - \sum_{j=1}^K (\pi^{-3} t_j^{-2} m_j^2) \\
 & + \pi^{-3} [1 - 2/A]^{-2} \sum_{\substack{j,k=1 \\ j \neq k}}^K m_j m_k |R_j - R_k|^{-2}.
 \end{aligned}$$

Using $m_j m_k \leq (m_j^2 + m_k^2)/2$ we can rewrite this as

$$\| |p|^{-1/2} W \psi \|^2 \leq \| \psi \|^2 - \sum_{j=1}^K m_j^2 b_j,$$

where $b_j = \pi^{-3} t_j^{-2} - \pi^{-3} [1 - 2/A]^{-2} \sum_{k \neq j} |R_k - R_j|^{-2}$. By the definition (2.7) of t_j , we have

$$b_j \geq \pi^{-3} [A^2 - (1 - 2/A)^{-2}] \sum_{k \neq j} |R_k - R_j|^{-2}.$$

This shows that all the b_j will be positive, and hence $\| |p|^{-1/2} W \psi \| \leq \| \psi \|$, if $A^2 \geq (1 - 2/A)^{-2}$, or

$$A \geq A_0 = 3. \tag{2.14}$$

Thus, provided (2.14) is satisfied, we have $\| |p|^{-1/2} W \psi \| \leq \| \psi \|$ for all ψ in C_0^∞ , hence $\| W |p|^{-1/2} \| = \| |p|^{-1/2} W \| = 1$. This implies $|p| - W^2 = |p|^{1/2} [1 - (W |p|^{-1/2})^* (W |p|^{-1/2})] |p|^{1/2} \geq 0$.

Since $\left(\sum_j W_j \right)^2 \leq W^2$ by (2.9), we have therefore proved (2.6) and *a fortiori* (2.5) for all $A \geq A_0 = 3$. ■

Remarks

1. In terms of Z_c and e^2 , we can rewrite (2.4) as $\forall \psi \in Q(|p|), \|\psi\| = 1$:

$$\left\langle \psi, \left[|p| - Z_c \alpha \sum_{j=1}^K |x - R_j|^{-1} \right] \psi \right\rangle \geq -b Z_c^2 \alpha \sum_{\substack{j,k=1 \\ j < k}}^K |R_j - R_k|^{-1}, \quad (2.15)$$

where $b = \pi \alpha A_0 \leq 6.88 \times 10^{-2}$.

2. Note that (2.15) implies that $\sum_{j=1}^K Z_j \alpha |x - R_j|^{-1}$ is relatively form-bounded with respect to $|p|$, with relative bound $(\max Z_j)/Z_c$. This means that $h_{1,K}$ and $H_{1,K}$ can be defined as form sums, with form domain $Q(|p|)$ if $Z_j < Z_c$ for all j ; if $Z_j = Z_c$ for some j , we define $H_{1,K}, h_{1,K}$ to be the Friedrichs extensions of the quadratic forms on $Q(|p|)$.

With Proposition 2.2 ends our first step in the proof of Theorem 1 which is now established if $Z_j = Z_c$ for all j . From this result we shall derive Theorem 1 for all possible values of Z_j by a concavity argument. First we prove another technical lemma:

Lemma 2.3. *Let f be a real function on the n -dimensional unit cube $[0, 1]^n$, convex in each variable separately. Let h be the smallest concave function on $[0, 1]^n$ (jointly concave in all the variables) agreeing with f on all the vertices of the cube, i.e.*

$$\forall P_k \text{ vertex of } [0, 1]^n \quad (k = 1, \dots, M, \text{ with } M = 2^n): \quad h(P_k) = f(P_k).$$

Then $\forall (x_1, \dots, x_n) \in [0, 1]^n: h(x_1, \dots, x_n) \geq f(x_1, \dots, x_n)$.

Remarks

1. We shall apply this lemma to the function $f(x_1, \dots, x_n) = \sum_{i < j} \alpha_{ij} x_i x_j$. If we keep all the variables but one fixed, f is affine in the remaining variable. The requirement that f be convex in each variable separately is therefore obviously satisfied.

2. By changing signs, one immediately sees that the lemma is still true if we exchange “concave” and “convex,” if we make h the “largest convex function...” instead of the “smallest concave function...” and if we reverse the inequality sign in the conclusion.

3. Since h is the smallest concave function agreeing with f on the vertices, the conclusion $h \geq f$ automatically implies that h is the concave envelope of f . The lemma tells us therefore that to construct the concave envelope on a cube of a function which is convex in each variable separately, one only has to consider the values of f at the vertices of the cube.

4. At any point (x_1, \dots, x_n) in the cube, h can be explicitly constructed as follows: for any decomposition of (x_1, \dots, x_n) into a convex combination of the vertices: $x_m = \sum_{j=1}^M \lambda_j (P_j)_m, m = 1, \dots, n$, with $\lambda_j \geq 0$ and $\sum_{j=1}^M \lambda_j = 1$, we define $F(\underline{\lambda}) = \sum_{j=1}^M \lambda_j f(P_j)$; h is given by the maximum of $F(\underline{\lambda})$, taken over all possible convex decompositions of (x_1, \dots, x_n) .

Proof of Lemma 2.3. Fix y_2, \dots, y_n so that each is either 0 or 1. Considered as a function of x_1 only, $(h-f)(x_1, y_2, \dots, y_n) = h(x_1, y_2, \dots, y_n) - f(x_1, y_2, \dots, y_n)$ is concave (h is jointly concave, and $-f$ is concave in each variable separately.)

Since h and f agree on the vertices, we have $(h-f)(0, y_2, \dots, y_n) \geq 0$, $(h-f)(1, y_2, \dots, y_n) \geq 0$. By the concavity of $h-f$ in x_1 , this implies $(h-f)(x_1, y_2, \dots, y_n) \geq 0$ for all $x_1 \in [0, 1]$. Now fix x_1 , and note that h and f satisfy the stated hypotheses in the variables y_2, \dots, y_n . The complete result follows by iterating the argument n times. ■

With the help of this lemma and Proposition 2.2, we can now prove Theorem 1:

Proof of Theorem 1. For $(z_1, \dots, z_K) \in [0, 1]^K$, we define

$$F(z_1, \dots, z_K) = \inf_{\psi \in Q(|p|), \|\psi\|=1} [\langle \psi, |p|\psi \rangle - (2/\pi) \sum_{k=1}^K z_k \langle \psi, |x - R_k|^{-1} \psi \rangle].$$

As the infimum of linear functions, F is concave. Define

$$f(z_1, \dots, z_K) = -(12/\pi) \sum_{\substack{k,l=1 \\ k < l}}^K |R_k - R_l|^{-1} z_k z_l. \tag{2.16}$$

Proposition 2.2 tells us that

$$F(P_k) \geq f(P_k) \quad k = 1, \dots, 2^K,$$

where the P_k , $k = 1, \dots, 2^K$ are the vertices of the cube $[0, 1]^K$. (When all the components of the vertices are 1, i.e. $P_k = (1, \dots, 1)$, (2.16) follows from Proposition 2.2. When some of the components of P_k are zero, (2.16) still follows from Proposition 2.2, but now with a smaller value for K .) Since F is concave, (2.16) implies that F is larger than h , the smallest concave function agreeing with f at the cornerpoints P_k of the cube. But h is larger than f by Lemma 2.3; by combining these two inequalities we obtain $F \geq f$, or for all $z_k \in [0, 1]$, $k = 1, \dots, K$:

$$\begin{aligned} & \inf_{\psi \in Q(|p|), \|\psi\|=1} [\langle \psi, |p|\psi \rangle - (2/\pi) \sum_{k=1}^K z_k \langle \psi, |x - R_k|^{-1} \psi \rangle] \\ & \geq -(12/\pi) \sum_{\substack{k,l=1 \\ k < l}}^K z_k z_l |R_k - R_l|^{-1}. \quad \blacksquare \end{aligned} \tag{2.17}$$

Remark. In terms of the fine-structure constant α and of the nuclear charges Z_k , with $Z_k \leq Z_c = 2/(\alpha\pi)$ for all k , (2.17) can be rewritten as

$$|p| - \sum_{k=1}^K Z_k \alpha |x - R_k|^{-1} \geq -b \sum_{\substack{k,l=1 \\ k < l}}^K Z_k Z_l \alpha |R_k - R_l|^{-1},$$

where $b \leq 6.88 \times 10^{-2}$.

3. Monotonicity of the Ground State Energy in the Nuclear Coordinates

We prove Theorem 2 by essentially the same method as used by one of us [8] in the proof of the analogous statement for the Schrödinger Hamiltonian

$-\Delta - \sum_{k=1}^K Z_k |x - R_k|^{-1}$. The strategy followed there was the following (modulo some epsilons):

$$1) E_0(\underline{R}) = \inf \text{spec } H(\underline{R}) = - \lim_{t \rightarrow \infty} t^{-1} \ln G(x, x, t, \underline{R}) \tag{3.1}$$

for all x , where $G(x, y, t, \underline{R})$ is the kernel $\exp[-tH(\underline{R})](x, y)$.

2) By the Trotter product formula, $G(x, y, t, \underline{R})$ can be approximated by multiple integrals involving $\exp[-(t/n)(-\Delta)](x, y)$ and $\exp[(t/n)Z_k|x - R_k|^{-1}]$. (Here one of the epsilons mentioned above comes into play: in order to ensure convergence of the integrals, one writes $G(x, y, t, \underline{R})$ as $\lim_{\varepsilon \rightarrow 0} G_\varepsilon(x, y, t, \underline{R})$,

where G_ε is the kernel of the semigroup corresponding to $-\Delta - \sum_{k=1}^K Z_k(|x - R_k|^2 + \varepsilon^2)^{-1/2}$. The whole argument is then carried out for G_ε rather than for G and the limit $\varepsilon \rightarrow 0$ is taken at the end.)

3) Since $\exp[as^{-1/2}]$ is completely monotone, it can be written as the Laplace transform of a positive measure; this can be used to write $\exp[(t/n)Z|x - R|^{-1}]$ as an integral, over a positive measure, of $\exp[-s|x - R|^2]$.

4) Since all the kernels in the Trotter product approximation for G have been reduced to positive superpositions of Gaussian kernels (the kernel of $\exp[(t/n)\Delta]$ was already Gaussian in the Schrödinger case [8]), a lemma involving only Gaussian functions (Lemma 4 in [8]) can be applied, leading to the conclusion:

5) If $\underline{R}, \underline{R}'$ are given with $|R_i - R_j| \leq |R'_i - R'_j| \forall i, j$, and if x, y, x', y' are such that $|x - R_i| \leq |x' - R'_i|, |y - R_i| \leq |y' - R'_i|$ for all i , and $|x - y| \leq |x' - y'|$, then

$$G(x, y, t, \underline{R}) \geq G(x', y', t, \underline{R}')$$

Combined with 1) this proves the theorem. One uses the fact that given $\underline{R}, \underline{R}'$ and x , one can always find x' satisfying $|x' - R'_i| \geq |x - R_i|$ by choosing x' far enough away from the R'_i .

The same argument can be used to prove Theorem 2 in our case. Some modifications are needed, however. We shall only describe these in detail, and not repeat the whole proof as given in [8].

The first modification concerns the kernels $\exp[(t/n)(-\Delta + 1)^{1/2}]$ and $\exp[-(t/n)(-\Delta)^{1/2}]$. These are not Gaussian kernels, but they can again be written as positive superpositions of Gaussian kernels. Indeed, $\exp[-bs^{1/2}]$ is again a completely monotone function, and can therefore again, by Bernstein's theorem, be written as a Laplace transform of a positive measure:

$$\exp[-bs^{1/2}] = \int e^{-sv} d\mu(v, b);$$

an explicit formula for μ can be found in any table of Laplace transforms: $d\mu(v, b) = (b\pi^{-1/2}/2)v^{-3/2} \exp[-b^2/4v]dv$. It follows now that

$$\begin{aligned} \exp[-(t/n)(-\Delta + 1)^{1/2}](x, y) &= \int (2v)^{-3/2} \exp[-|x - y|^2/4v] e^{-v} d\mu(v, t/n), \\ \exp[-(t/n)(-\Delta)^{1/2}](x, y) &= \int (2v)^{-3/2} \exp[-|x - y|^2/4v] d\mu(v, t/n), \end{aligned} \tag{3.2}$$

which are indeed positive superpositions of Gaussians.

The second modification is more a set of comments on step 1) than a real modification.

First of all, note that it is enough to prove Theorem 2 for the case $Z_k < Z_c$ for all k . As the infima of decreasing linear functions of the Z_k , $E_{1,K}^0(Z)$ and $e_{1,K}^0(Z)$ are decreasing, upper semicontinuous functions of the Z_k , which ensures that

$$E_{1,K}^0(Z) = \lim_{\varepsilon \rightarrow 0} E_{1,K}^0(\underline{Z} - \varepsilon), e_{1,K}^0(Z) = \lim_{\varepsilon \rightarrow 0} e_{1,K}^0(\underline{Z} - \varepsilon),$$

where we use the notation $\underline{Z} - \varepsilon$ for the set $\{Z_j - \varepsilon\}$. This implies that if Theorem 2 holds for the case where $Z_k < Z_c$ for all k , it also holds in the case where some of the Z_k are equal to Z_c .

We can therefore restrict ourselves to the case $Z_k < Z_c$, all k . In this case $E_{1,K}^0$ is an isolated eigenvalue (one can easily show $E_{1,K}^0 < 1$ by a variational argument, and since $Z_k < Z_c$ for all k , we know that $\sigma_{\text{ess}}(H_{1,K}^0) = [1, \infty)$). For

$e_{1,K}^0$, matters are slightly different. If $\sum_{k=1}^K Z_k \leq Z_c$, we know that $e_{1,K}^0(Z, \underline{R}) = 0$, independently of \underline{R} , in which case the monotonicity in \underline{R} is trivial. We shall

therefore not consider this case in the discussion below. Whenever $\sum_{k=1}^K Z_k > Z_c$,

however, one has again $e_{1,K}^0 < 0$ by a variational argument, and since $\sigma_{\text{ess}}(h_{1,K}^0) = [0, \infty)$, this shows that then $e_{1,K}^0$ too is an isolated eigenvalue. We shall therefore assume in the following that $E_{1,K}^0, e_{1,K}^0$ are isolated eigenvalues. We know then that (3.1) will be true, for all x , if there exist ground states for $H_{1,K}^0, h_{1,K}^0$ which are strictly positive. The argument in the next paragraph shows that this is indeed the case: we show that the ground states for $H_{1,K}^0, h_{1,K}^0$

(where $\sum_{k=1}^K Z_k > Z_c$ in the last case, and $Z_k < Z_c$ for all k for both Hamiltonians) are nondegenerate and strictly positive, which is an interesting result in its own right.

Since $H_{1,K}^0 = H_0 - V(R, x), h_{1,K}^0 = h_0 - V(R, x)$, where $H_0 = (-\Delta + 1)^{1/2}$, $h_0 = (-\Delta)^{1/2}$, and where V is positive, the kernels $\exp[-tH_{1,K}^0](x, y)$, $\exp[-th_{1,K}^0](x, y)$ will be pointwise larger than $\exp[-tH_0](x, y)$ and $\exp[-th_0](x, y)$, respectively. By (3.2) one sees immediately that $\exp[-tH_0](x, y)$, $\exp[-th_0](x, y)$ are strictly positive for all x, y , which implies therefore $\exp[-tH_{1,K}^0](x, y) > 0$, $\exp[-th_{1,K}^0](x, y) > 0$ for all x, y . By the extension of the Perron-Frobenius theorem to operators (see e.g. [14], Theorem XIII.43), this implies that the ground states of $H_{1,K}^0, h_{1,K}^0$ (which correspond to the largest eigenvalues of $\exp[-tH_{1,K}^0], \exp[-th_{1,K}^0]$) are nondegenerate, and the corresponding eigenvectors strictly positive.

Taking into account all the remarks made above, the proof given in [8] can now be completely transcribed to our case; this completes our proof of Theorem 2.

References

1. Dyson, F., Lenard, A.: Stability of matter. I. J. Math. Phys. **8**, 423-434 (1967)
- Lenard, A., Dyson, F.: Stability of matter. II. J. Math. Phys. **9**, 698-711 (1968)
2. Lieb, E., Thirring, W.: A bound for the kinetic energy of fermions which proves the stability of matter.

Phys. Rev. Lett. **35**, 687–689 (1975); Errata: Phys. Rev. Lett. **35**, 1116 (1975). See also Lieb, E.: The stability of matter. Rev. Mod. Phys. **48**, 553–569 (1976)

3. Dyson, F.: Ground-state energy of a finite system of charged particles. J. Math. Phys. **8**, 1538–1545 (1967)
Lieb, E.: The $N^{5/3}$ law for bosons. Phys. Lett. **A70**, 71–73 (1979)
4. Weder, R.: Spectral analysis of pseudodifferential operators. J. Funct. Anal. **20**, 319–337 (1975)
5. Herbst, I.: Spectral theory of the operator $(p^2 + m^2)^{1/2} - Ze^2/r$. Commun. Math. Phys. **53**, 285–294 (1977); Errata: Commun. Math. Phys. **55**, 316 (1977)
6. Kato, T.: Perturbation theory for linear operators. Berlin, New York: Springer 1966 (2nd edn. 1976)
7. Lieb, E., Simon, B.: Monotonicity of the electronic contribution to the Born–Oppenheimer energy. J. Phys. **B11**, L537–542 (1978)
8. Lieb, E.: Monotonicity of the molecular electronic energy in the nuclear coordinates. J. Phys. **B15**, L63–L66 (1982)
9. Daubechies, I.: An uncertainty principle for fermions with generalized kinetic energy. Commun. Math. Phys. (1983)
10. Lieb, E., Simon, B.: The Thomas–Fermi theory of atoms, molecules and solids. Adv. Math. **23**, 22–116 (1977)
11. Kovalenko, V., Perelmuter, M., Semenov, Ya.: Schrödinger operators with $L_w^{1/2}(\mathbb{R}^d)$ potentials. J. Math. Phys. **22**, 1033–1044 (1981)
12. Brascamp, H., Lieb, E., Luttinger, M.: A general rearrangement inequality for multiple integrals. J. Funct. Anal. **17**, 227–237 (1974)
13. Lieb, E.: Sharp constants in the Hardy–Littlewood–Sobolev and related inequalities (submitted)
14. Reed, M., Simon, B.: Methods of modern mathematical physics Vol. IV: Analysis of operators. New York: Academic Press 1978

Communicated by J. Fröhlich

Received March 24, 1983