Monotonicity of Born-Oppenheimer electronic energies for excited molecular states

1993 J. Phys. A: Math. Gen. 26 159

(http://iopscience.iop.org/0305-4470/26/1/017)

J. Phys. A: Math. Gen. 26 (1993) 159–170. Printed in the UK

Monotonicity of Born–Oppenheimer electronic energies for excited molecular states

H Hogreve

Centre de Physique Théorique, Luminy Case 907, F-13288 Marseille Cedex 9, France
and PHYMAT, Université de Toulon et du Var, F-83130 La Garde, France

Received 19 June 1992

Abstract. We consider adiabatic molecular systems consisting of one electron and $N \geqslant 2$ nuclei with arbitrary positive charges. If all the nuclei are arranged on a line, the angular momentum around the internuclear axis is conserved; using the corresponding separability in cylindrical coordinates it is shown that the lowest electronic energies in each symmetry sector of fixed angular momentum increase if the internuclear distances become larger. This extends previous monotonicity results for the ground state by Lieb and Simon. For the physically most important case $N = 2$ (or $N = 3$) we give a separate proof that emphasizes the basic role of *reflection positivity* in this context.

1. Introduction

In the Born–Oppenheimer approximation a non-relativistic one-electron molecule is modelled by the quantum Hamiltonian

$$H = h + \sum_{i<j}^{N} \frac{Z_i Z_j}{|\boldsymbol{X}_i - \boldsymbol{X}_j|}. \tag{1.1}$$

Here h stands for the electronic part of H,

$$h = h^{\text{kin}} - V \tag{1.2a}$$

where

$$h^{\text{kin}} = -\Delta$$
$$V(r) = \sum_{i=1}^{N} v(r; \boldsymbol{X}_i, Z_i) = \sum_{i=1}^{N} \frac{Z_i}{|r - \boldsymbol{X}_i|} \tag{1.2b}$$

and h depends parametrically on the nuclear positions $\boldsymbol{X}_i \in \mathbb{R}^3$ and charges $Z_i > 0$, $i = 1, \ldots, N$. The second term in (1.1) describes the nuclear repulsion. Clearly it always gives a positive contribution to the total energy and decreases for increasing nuclear separations $R_{ij} = |\boldsymbol{R}_{ij}| = |\boldsymbol{X}_i - \boldsymbol{X}_j|$. In contrast, the behaviour of the electronic energies resulting from h is *a priori* not obvious and may depend

0305-4470/93/010159+12$07.50

crucially on the respective state. In fact, all existing rigorous results pertain exclusively to the ground-state energy

$$e_0(\{R_{ij}\}_{1\leqslant i<j\leqslant N}) = \inf \sigma\{h(\{R_{ij}\}_{1\leqslant i<j\leqslant N})\}$$

(since the dependence on the charges Z_i is irrelevant for our considerations, we do not indicate it explicitly). The first proof of the monotonic increase of $e_0(R_{12})$ in R_{12} is due to Lieb and Simon (1978) and was given for the case of two nuclei. The same result, for $N = 2$, was rederived by Hoffmann–Ostenhoff (1980) employing a somewhat different approach based on differential inequalities (cf also Thirring 1981). Later, Lieb (1982) generalized this monotonicity to systems with $N \geqslant 2$ nuclei and demonstrated that $e_0(\{R'_{ij}\}_{1\leqslant i<j\leqslant N}) \geqslant e_0(\{R_{ij}\}_{1\leqslant i<j\leqslant N})$ if $R'_{ij} \geqslant R_{ij}$ for all $1 \leqslant i < j \leqslant N$.

Molecular binding requires the existence of a (global) minimum at a finite configuration $R_{ij} < \infty$, $1 \leqslant i < j \leqslant N$, of the potential energy curve associated with the total Hamiltonian H. A sufficiently rapid increase in the electronic energies is necessary for such a minimum to overcome (locally) the decay of the nuclear repulsion. This competition depends sensibly on the actual values of the Z_i, and for large Z_i one expects that the repulsion effect will dominate and exclude a binding minimum for the total energy. Indeed, for the situation $N = 2$ and $Z_1 = Z_2 = Z$ it is known (Duclos and Hogreve 1991, 1992) that a charge $Z > 4$ renders the total energy monotonically decreasing not only for the *ground* but also for *all excited* states of the system.

On the other hand, apparently there are no rigorous results for the behaviour of the electronic curves for excited states. Our aim here is to show that—in analogy to the ground state—a monotonic increase continues to hold for a certain class of excited electronic energies. To this end, we assume that all nuclei are located on a line which we will take as the ζ-axis of cylindrical coordinates for $\mathbb{R}^3$. Then the system is invariant under the symmetry group $C_{\infty v}$ (Tinkham 1964). In particular, the operator $L_\varphi = \mathrm{i}\partial/\partial\varphi$ of angular momentum around the internuclear axis commutes (at least formally) with H and h, implying that its eigenvalues $m \in \mathbb{Z}$ label distinct symmetry subspaces of eigenvectors of H and h (and irreducible representations of $C_{\infty v}$). Denoting by e_m the lowest eigenvalue of h restricted to the subspace labelled by m (cf (2.17) for the precise definition), our main result is the following monotonicity property:

Theorem 1.1. Let $\{R_{ij}\}_{1\leqslant i<j\leqslant N}$ specify a nuclear configuration by $\boldsymbol{R}_{ij} = \boldsymbol{X}_i - \boldsymbol{X}_j = R_{ij}\hat{\boldsymbol{R}}$, $R_{ij} > 0$, with $\hat{\boldsymbol{R}}$ a fixed unit vector in $\mathbb{R}^3$, and similarly for $\{R'_{ij}\}_{1\leqslant i<j\leqslant N}$. Then $R_{ij} \leqslant R'_{ij}$ for all $1 \leqslant i < j \leqslant N$ implies

$$e_m(\{R'_{ij}\}_{1\leqslant i<j\leqslant N}) \geqslant e_m(\{R_{ij}\}_{1\leqslant i<j\leqslant N}) \tag{1.3}$$

for all $m \in \mathbb{Z}$ and $Z_i > 0$, $i = 1, \ldots, N$.

Molecules containing only two nuclei, e.g. H_2^+ or HeH^{++}, are always invariant under $C_{\infty v}$ and their energies depend only on one distance parameter $R = R_{12}$. In view of the physical importance of these systems, we reformulate theorem 1.1 for this specific situation:

Corollary 1.2. For all $Z_1, Z_2 > 0$ the lowest electronic energy $e_m(R)$ for each given angular momentum $m \in \mathbb{Z}$ is monotonically increasing in R.

The basic ingredients of our proof for theorem 1.1 are the (partial) separability of the Schrödinger equation due to the conservation of the azimuthal angular momentum component together with the observation that in each m-sector of the separated problem the wavefunction for the lowest state can be chosen to be non-negative. A rigorous discussion of the separation procedure and a derivation of some important properties of the separated operators is given subsequently in section 2. Armed with these results, the strategy of Lieb (1982) can be easily adapted to complete the proof of theorem 1.1. Since the required changes are minimal, in section 2 we shall only sketch the modifications to Lieb's proof. On the other hand, section 3 is devoted to an independent and complete proof of corollary 1.2 based on the fact that the potentials involved obey a 'reflection positivity' property. We conclude in section 4 with various remarks about possible generalizations and limitations of the monotonicity properties of electronic curves.

2. Separation of variables

Although the separation procedure discussed later is quite standard in physics if applied formally, its rigorous justification including an analysis of the operators arising after the separation seems never to have been accomplished in the literature. In fact, as pointed out by Miller (1977), separating variables 'proves suprisingly subtle and difficult to describe in general'. Thus, the corresponding results of this section may be of interest in their own right.

For the configurations considered, the $\mathcal{C}_{\infty v}$ symmetry suggests a separation in cylindrical coordinates (ρ, ζ, ϕ) with $\rho \in \mathbb{R}_0^+ = \mathbb{R}^+ \cup \{0\}$, $\zeta \in \mathbb{R}$, $\varphi \in S^1 = [0, 2\pi]$. Setting $\Omega = \mathbb{R}^+ \times \mathbb{R}$ and $\mathrm{d}\omega = \rho \, \mathrm{d}\rho \, \mathrm{d}\zeta$, the relevant Hilbert space is decomposed as

$$L^2(\mathbb{R}^3, \mathrm{d}^3 r) = L^2(\Omega, \mathrm{d}\omega) \otimes L^2(S^1, \mathrm{d}\varphi) = \bigoplus_{m \in \mathbb{Z}} \mathcal{L}_m \tag{2.1}$$

where

$$\mathcal{L}_m = L^2(\Omega, \mathrm{d}\omega) \otimes \Lambda_m$$

and $\Lambda_m = \{c g_m \mid c \in \mathbb{C}\}$, $g_m(\varphi) = \exp(\mathrm{i}m\varphi)$. The Hamiltonian h is reduced by the $\mathcal{L}_m$ in the sense that

$$h = \bigoplus_{m \in \mathbb{Z}} h_m \tag{2.2}$$

where h_m acts in $\mathcal{L}_m$ by

$$h_m = h_m^{\mathrm{kin}} - V \qquad h_m^{\mathrm{kin}} = h_m^{\perp} \otimes \mathbf{1} + \mathbf{1} \otimes h^{\parallel} \tag{2.3a}$$

with

$$h_m^{\perp} = -\frac{1}{\rho} \frac{\partial}{\partial \rho} \rho \frac{\partial}{\partial \rho} + \frac{m^2}{\rho^2} \qquad h^{\parallel} = -\frac{\partial^2}{\partial \zeta^2} \tag{2.3b}$$

and we continue to denote the multiplication operator for the potential (now involving the variables (ρ, ζ)) by V.

We recall that the original operator h is self-adjoint on the Sobolev space $W_{2,2}(\mathbb{R}^3)$. The analogous properties of the separated operators are provided in the following theorem:

Theorem 2.1. Let $h_m, h_m^{\mathrm{kin}}, h_m^{\perp}$ and $h^{\parallel}$ be given as in (2.3).

(i) The kinetic parts h_m^{kin} of the Hamiltonians h_m are essentially self-adjoint on $\mathcal{D}_m \otimes C_0^\infty(\mathbb{R})$ for all $m \in \mathbb{Z}$, where for $m \neq 0$ the domain $\mathcal{D}_m$ is always $\mathcal{D}_m = C_0^\infty(\mathbb{R}^+)$, whereas

$$\mathcal{D}_0 = \{f \in C^\infty(\mathbb{R}^+) \mid \exists c_f > 0 \text{ such that } f(\rho) = 0 \text{ if } \rho \geqslant c_f, \text{ and } f'(\rho) = o(1) \text{ as } \rho \to 0\}. \tag{2.4}$$

(ii) For all $m \in \mathbb{Z}$ the operator h_m^{kin} is semibounded $h_m^{\mathrm{kin}} \geqslant 0$, and the kernel of the generated semigroup $\exp(-t h_m^{\mathrm{kin}}) = \exp(-t h_m^{\perp}) \otimes \exp(-t h^{\parallel})$ is pointwise non-negative.

(iii) For all $m \in \mathbb{Z}$ the potential V is relative compact with respect to h_m^{kin}, and therefore h_m essentially self-adjoint on the same domains as specified in (i).

(iv) The spectra of the original and separated Hamiltonians are related by

$$\bigcup_{m \in \mathbb{Z}} \sigma(h_m) \subset \sigma(h). \tag{2.5}$$

Proof. (i) Invoking well known facts about tensor products (cf theorem VIII.33 of Reed and Simon (1980)) we need only show the essential self-adjointness of $h_m^{\perp}$ and $h^{\parallel}$ regarded as ordinary differential operators on $\mathcal{D}_m$ and $C_0^\infty(\mathbb{R})$, respectively. For $h^{\parallel}$, the statement is proved in textbooks (see e.g. Reed and Simon 1975). In case of $h_m^{\perp}$, first note that this situation corresponds to a singular Sturm–Liouville problem (Jörgens and Rellich 1976). The associated differential equation $-y''(\rho) - \rho^{-1}y'(\rho) + m^2\rho^{-2}y(\rho) = 0$ has the general solutions $y(\rho) = c_1 y_1(\rho) + c_2 y_2(\rho)$ where

$$y_1(\rho) = \rho^m \qquad y_2(\rho) = \rho^{-m} \qquad \text{if } m \neq 0$$
$$y_1(\rho) = 1 \qquad y_2(\rho) = \ln(\rho) \qquad \text{if } m = 0.$$

For $m \neq 0$, choosing $c_1, c_2 \neq 0$ leads to y that are neither at $\rho \to 0$ nor at $\rho \to \infty$ in $L^2(\mathbb{R}^+, \rho\,\mathrm{d}\rho)$. Hence, at both endpoints we are in the limit-point case, implying the essential self-adjointness of $h_m^{\perp}$ on $C_0^\infty(\mathbb{R}^+)$ by standard methods and results (Jörgens and Rellich 1976, Reed and Simon 1975). If $m = 0$, then y is not square-integrable as $\rho \to \infty$ but all solutions y are in $L^2(\mathbb{R}^+, \rho\,\mathrm{d}\rho)$ as $\rho \to 0$. Thus, again we have the limit-point case at infinity, but now a limit-circle situation at the other endpoint $\rho = 0$, the latter excluding $C_0^\infty(\mathbb{R}^+)$ as domain for essential self-adjointness of $h_m^{\perp}$. Applying the usual procedure to construct boundary conditions (at $\rho = 0$) that produce self-adjoint extensions (cf chapter III.4 of Jörgens and Rellich (1975), or chapter 8.4 of Weidmann (1980)), one such possible boundary condition is of the form $\lim_{\rho \to 0} \rho f'(\rho) = 0$, i.e. as imposed on functions in $\mathcal{D}_0$. That this particular choice from the one-parameter family is indeed the 'good' boundary condition for the separated operator will become clear in (iv).

(ii) To prove $h_m^{\mathrm{kin}} \geqslant 0$, it is sufficient to verify $\langle \phi, h_m^{\perp}\phi\rangle_{L_2(\mathbb{R}^+, \rho\,\mathrm{d}\rho)} \geqslant 0$ and $\langle \psi, h^{\parallel}\psi\rangle_{L_2(\mathbb{R})} \geqslant 0$ for $\phi \in \mathcal{D}_m$, $\psi \in C_0^\infty(\mathbb{R})$. Whereas the latter inequality is trivial, the first one follows from

$$\langle \phi, h_m^{\perp}\phi\rangle_{L_2(\mathbb{R}^+, \rho\,\mathrm{d}\rho)} = \int_0^\infty \mathrm{d}\rho\, \{\rho|\phi'(\rho)|^2 + m^2\rho^{-1}|\phi(\rho)|^2\}.$$

As a consequence, all $z < 0$ are in the resolvent sets of $h_m^{\perp}$, $h^{\parallel}$ and h_m^{kin}.

Looking now at the generated semigroups, the kernel of $\exp(-th^{\parallel})$ is explicit

$$e^{-th^{\parallel}}(\zeta_1, \zeta_1) = (4\pi t)^{-1/2} e^{-(\zeta_1 - \zeta_2)/4t}. \tag{2.6}$$

The pointwise non-negativity of the kernel of $\exp(-th_m^{\perp})$ results from

$$(h_m^{\perp} + z)^{-1}(\rho_1, \rho_2) \geqslant 0 \tag{2.7}$$

for $z > 0$, and the relation (Reed and Simon 1978)

$$e^{-th_m^{\perp}} = \lim_{n \to \infty} (1 + (t/n)h_m^{\perp})^{-n}.$$

To show the inequality (2.7), let us construct an explicit expression for the resolvent in terms of appropriate solutions u of

$$\left(\frac{d^2}{d\rho^2} + \frac{1}{\rho} \frac{d}{d\rho} - \frac{m^2}{\rho^2} - z \right) u = 0. \tag{2.8}$$

After a scaling, (2.8) becomes identical to the modified Bessel equation; linear independent solutions are, for instance, the modified Bessel functions I_m, K_m. From the known asymptotics of these functions and their derivatives as $\rho \to 0$ and $\rho \to \infty$ (Olver 1974) in combination with theorem III.3.2 of Jörgens and Rellich (1976) (or a direct verification) we conclude that for $z > 0$ the kernel of $(h_m^{\perp} + z)^{-1}$ is given by

$$(h_m^{\perp} + z)^{-1}(\rho_1, \rho_2) = \begin{cases} I_m(z^{1/2}\rho_1) K_m(z^{1/2}\rho_2) & \text{if } \rho_1 \leqslant \rho_2 \\ K_m(z^{1/2}\rho_1) I_m(z^{1/2}\rho_2) & \text{if } \rho_1 \geqslant \rho_2. \end{cases} \tag{2.9}$$

Finally, in view of (2.9), the non-negativity (2.7) is just a consequence of $I_m(\rho) \geqslant 0$ and $K_m(\rho) \geqslant 0$ for $\rho \geqslant 0$ and all $m \geqslant 0$ (Olver 1974).

We remark that the explicit estimates on the kernels imply that $\exp(-th_m^{\mathrm{kin}})$ is positivity preserving; alternatively, this could be proved in a more abstract way by a verification of the first Beurling–Deny criterion (Reed and Simon 1978) for h_m^{kin}.

(iii) We first decompose V into a part V_s with compact support that includes all points where V becomes singular, and a non-singular part $V_c = V - V_s$. By standard arguments, the existence of a $n \geqslant 1$ with a finite trace

$$\mathrm{Tr}\{(V_{s,c}^{1/2}(h_m^{\mathrm{kin}} + s)^{-1} V_{s,c}^{1/2})^n\} < \infty \tag{2.10}$$

(for $s > 0$) entrains the relative compactness of V_s, V_c and thus of $\pm(V_s + V_c)$ with respect to h_m^{kin}. For V_s and $n = 1$, the trace in (2.10) becomes

$$\int_0^{\infty} \rho \, d\rho \int_{-\infty}^{\infty} d\zeta \, \{V_s(\rho, \zeta)(h_m^{\mathrm{kin}} + s)^{-1}(\rho, \rho; \zeta, \zeta)\}. \tag{2.11}$$

The integral (2.11) is finite because at the singularities of V the variable ρ has, of necessity, to vanish, yielding $\lim_{\rho \to 0} \rho V_s(\rho, \cdot) < \infty$, and because integrability problems as $\rho, |\zeta| \to \infty$ are absent due to the compact support of V_s.

For V_c, we use the inequality $\mathrm{Tr}\{(B^{1/2}AB^{1/2})^n\} \leqslant \mathrm{Tr}\{B^{n/2}A^nB^{n/2}\}$ (which holds for $A, B > 0$, $n \geqslant 1$, cf Lieb and Thirring (1976), Araki (1990)) and estimate $V_c(\rho,\zeta) \leqslant V_c(\rho \equiv 0, \zeta)$, $V_c(\rho,\zeta) \leqslant V_c(\rho, \zeta \equiv 0)$, to bound the trace in (2.10) from above by

$$\int_0^\infty \rho \,\mathrm{d}\rho \int_{-\infty}^{\infty} \mathrm{d}\zeta \,\{V_c(\rho \equiv 0,\zeta)^{n/2}(h_m^{\mathrm{kin}} + s)^{-n}(\rho,\rho;\zeta,\zeta)V_c(\rho,\zeta \equiv 0)^{n/2}\}. \quad (2.12)$$

From the asymptotic behaviour $V_c(\rho \equiv 0,\zeta)^{n/2} \propto |\zeta|^{-n/2}$, $V_c(\rho,\zeta \equiv 0)^{n/2} \propto \rho^{-n/2}$ as $\rho, |\zeta| \to \infty$, and the uniform boundedness of the kernel of $(h_m^{\mathrm{kin}} + s)^{-1}$, we deduce that $\{\ldots\}$ in (2.12) is certainly integrable if $n > 4$.

(iv) For the essential spectrum, relation (2.5) is easy. Since for all $m \in \mathbb{Z}$ we have $\sigma_{\mathrm{ess}}(h_m^{\perp}), \sigma_{\mathrm{ess}}(h^{\parallel}) \subset [0,\infty) = \sigma_{\mathrm{ess}}(h)$ and therefore also $\sigma_{\mathrm{ess}}(h_m) = \sigma_{\mathrm{ess}}(h_m^{\mathrm{kin}}) \subset \sigma_{\mathrm{ess}}(h)$, clearly $\bigcup_{m\in\mathbb{Z}} \sigma_{\mathrm{ess}}(h_m) \subset \sigma_{\mathrm{ess}}(h)$. Next, to show $\sigma_{\mathrm{disc}}(h_m) \subset \sigma_{\mathrm{disc}}(h)$, let ψ be an eigenfunction of h_m and—as before—$g_m(\varphi) = \exp(im\varphi)$. Since (after change of coordinates) the differential equations for the eigenproblems are identical, it suffices to verify $\psi \otimes g_m \in D(h)$, or, because $D(h) = W_{2,2}(\mathbb{R}^3)$ equals the domain of the maximal operator constructed from the differential expression h on $C_0^\infty(\mathbb{R}^3)$, to prove that $h(\psi \otimes g_m) \in L^2(\mathbb{R}^3)$. Starting by regarding $h(\psi \otimes g_m)$ as a distribution, for all $\phi \in C_0^\infty(\mathbb{R}^3)$

$$\int \mathrm{d}^3r\, \phi(r) h(\psi \otimes g_m)(r) = \int \mathrm{d}^3r\, (h\phi)(r)(\psi \otimes g_m)(r)$$

$$= \lim_{n\to\infty} \int \mathrm{d}^3r\, (-\Delta\phi(r) - V(r)\phi(r))(\psi_n \otimes g_m)(r). \quad (2.13)$$

Here ψ_n is an approximating sequence for ψ (with respect to the graph norm of h_m) of the form $\psi_n = \psi_n^{\perp} \otimes \psi_n^{\parallel}$ with $\psi_n^{\perp} \in \mathcal{D}_m$, $\psi_n^{\parallel} \in C_0^\infty(\mathbb{R})$. Changing to cylindrical coordinates in (2.13), via partial integrations all derivatives $\partial^2/\partial\varphi^2$, $\partial^2/\partial\zeta^2$, $\rho^{-1}(\partial/\partial\rho)\rho(\partial/\partial\rho)$ can be moved to $\psi \otimes g_m$; the properties of ψ_m and g_m ensure that all arising boundary terms do vanish inclusive those for the ρ-variable. In consequence,

$$\lim_{n\to\infty} \int \mathrm{d}^3r(-\Delta\phi(r) - V(r)\,\phi(r))(\psi_n \otimes g_m)(r)$$

$$= \lim_{n\to\infty} \int_0^\infty \rho\,\mathrm{d}\rho \int_{-\infty}^{\infty} \mathrm{d}\zeta \int_0^{2\pi} \mathrm{d}\varphi\, \phi(r(\rho,\zeta,\varphi))$$

$$\times \left\{ -\frac{1}{\rho}\frac{\partial}{\partial\rho}\rho\frac{\partial}{\partial\rho} + \frac{m^2}{\rho^2} - \frac{\partial^2}{\partial\zeta^2} - V(\rho,\zeta) \right\} \psi_n^{\perp}(\rho)\psi_n^{\parallel}(\zeta)\mathrm{e}^{\mathrm{i}m\varphi} \quad (2.14)$$

and since in (2.14) $\{\ldots\}(\psi_n^{\perp} \otimes \psi_n^{\parallel} \otimes g_m) = h_m\psi_n \otimes g_m$ converges to a L^2-function as $n \to \infty$, it follows that $h(\psi \otimes g_m) \in L_2(\mathbb{R}^3)$, thus concluding the proof of (2.5) (the inverse inclusion, although true, is not needed for our purposes).

Note that at this point it turns out that the additional boundary condition (at $\rho = 0$) specified for $\mathcal{D}_0$ is indeed the 'correct' one. Namely, all other possible self-adjoint extensions for $h_0^{\perp}$ have domains containing functions $\psi_n^{\perp}$ with $\psi_n^{\perp}(\rho) \propto \ln(\rho)$ as $\rho \to 0$. However, for such functions the boundary term in (2.14)

$$\rho\phi(r(\rho,\zeta,\phi)) \left\{ \frac{\mathrm{d}\psi_n^{\perp}}{\mathrm{d}\rho}(\rho) \right\} \Big|_{\rho=0}^{\infty} = -\phi(r(\rho = 0,\zeta,\phi))$$

does not vanish in general; instead, it leads to a $\delta(\rho)$-type contribution for $h(\psi_n \otimes g_m)$ showing that in this case $h(\psi \otimes g_m) \notin L^2(\mathbb{R}^3)$. □

With theorem 2.1 at our disposal, we can proceed with the proof of theorem 1.1 along the same steps as employed by Lieb (1982). Therefore, here a brief sketch of the simple modifications necessary to adapt Lieb's proof to the present situation is sufficient. First we observe that, also when working in cylindrical coordinates, the representation of the exponential of the Coulomb potential (or its regularized version) as stated in equation (7) of Lieb (1982) remains valid, i.e. for $b > 0$

$$e^{b|r-X_i|^{-1}} = \int e^{-t\{\rho^2-(\zeta-|X_i|)^2\}}\, d\mu(t;b) \tag{2.15}$$

with a positive measure $d\mu(t;b)$ (note: we have chosen $\hat{\boldsymbol{R}}$ in the ζ-direction). Furthermore, introducing, in analogy to equation (8) of Lieb (1982),

$$Z(\{R_{ij}\}_{1\leqslant i<j\leqslant N};\rho) = \int d\zeta_1 \ldots d\zeta_K \times \exp\left\{ -\sum_{i,j=1}^{K} \zeta_i \zeta_j A_{ij} - \sum_{i=1}^{K}\sum_{j=1}^{M} \lambda_{ij}(\zeta_i - X_{ij})^2 \right\} F(\rho_1, \ldots, \rho_K) \tag{2.16}$$

a one-dimensional version of lemma 4 of Lieb (1982) shows that now pointwise in ρ the inequality $Z(\{R'_{ij}\}_{1\leqslant i<j\leqslant N};\rho) \leqslant Z(\{R_{ij}\}_{1\leqslant i<j\leqslant N};\rho)$ holds if $R_{ij} \leqslant R'_{ij}$ for all $1 \leqslant i < j \leqslant N$. In (2.16) the quantities A_{ij}, λ_{ij} satisfy the same requirements as those in lemma 4 of Lieb, and the X_{ij} correspond to the parameters called R_{ij} by Lieb. More precisely, the X_{ij} arise from a multiple application of (2.15) within a Trotter product formula, i.e. $X_{ij} \in \{|\boldsymbol{X}_k|,\ k = 1, \ldots, N\}$ where we place the 'leftmost' nucleus, e.g. $k = 1$, at the origin (hence $|\boldsymbol{X}_1| = 0$) and the remaining ones on the positive ζ-axis. Enlarging the internuclear distances $R'_{ij} \geqslant R_{ij}$, $1 \leqslant i < j \leqslant N$, evidently entails $|X'_{kl} - X'_{mn}| \geqslant |X_{kl} - X_{mn}|$ for all k, l, m, n. The function $F(\rho_1, \ldots, \rho_K)$ in (2.16) can be assumed to be non-negative due to (2.15) and assertion (ii) of theorem 2.1. This form of Z as written in (2.16) reflects, in particular, the fact that the kernel

$$e^{-th_m^{\mathrm{ks}}}(\rho_i, \rho_j; \zeta_i, \zeta_j) = e^{-th_m^{\perp}}(\rho_i, \rho_j)(4\pi t)^{-1/2} e^{-(\zeta_i-\zeta_j)^2/4t}$$

is still Gaussian in the relevant ζ-variables and $\exp(-th_m^{\perp})(\rho_i, \rho_j) \geqslant 0$. Hence, pointwise in ρ, an analogue of theorem 5 (in the ζ and R variables) of Lieb (1982) can be easily established for the kernel $G_m(\rho, \zeta, t, \{R_{ij}\}_{1\leqslant i<j\leqslant N}) = \exp(-th_m(\{R_{ij}\}_{1\leqslant i<j\leqslant N}))(\rho, \zeta)$. Then, the desired inequality (1.3) for the energies

$$\begin{aligned} e_m(\{R_{ij}\}_{1\leqslant i<j\leqslant N}) &\stackrel{\text{def}}{=} \inf \sigma\{h_m(\{R_{ij}\}_{1\leqslant i<j\leqslant N})\} \\ &= -\lim_{t\to\infty} t^{-1} \ln G_m(\rho, \zeta, t, \{R_{ij}\}_{1\leqslant i<j\leqslant N}) \end{aligned} \tag{2.17}$$

follows by the same arguments as in Lieb (1982).

3. Reflection positivity and monotonicity of electronic energies

In this section, we bring in the concept of reflection positivity and establish the constituents that will eventually be brought together for a proof of corollary 1.2. Including such an independent proof of the physically significant corollary 1.2 makes not only the present paper self-contained in this respect, but it is our belief that the role of reflection positivity in monotonicity problems is of fundamental importance. In fact, the reflection principle is already implicit in the first monotonicity proof of Lieb and Simon (1978) and our line of reasoning is just a generalization of those previous arguments.

Going back to Nelson (1973), the reflection principle serves as an important tool in constructive quantum field theory; reflection positivity is one of the so-called Osterwalder–Schrader axioms. Schrader (1977) employed the reflection principle to derive correlation inequalities for the Ising model (for the closely related method of 'duplicated variables', cf Messager and Miracle-Sole (1977), and some generalizations cf Hegerfeldt (1977)). For our purposes, we can restrict the discussion to reflections in $\mathbb{R}^3$ with respect to planes orthogonal to the ζ-axis (of cylindrical coordinates). Thus let ϑ_a denote the reflection with respect to the plane passing through the point $\zeta = a$, i.e. $\vartheta_a(\rho,\zeta,\varphi) = (\rho, 2a - \zeta, \varphi)$, and θ_a the induced mapping on functions $f:\mathbb{R}^3 \to \mathbb{C}$, i.e. $\theta_a f(r) = f(\vartheta_a r)$. Furthermore, setting $\Theta_a = \{r \in \mathbb{R}^3 \mid r = \rho\hat{e}_\rho + \zeta\hat{e}_\zeta + \varphi\hat{e}_\varphi \text{with} \zeta \leqslant a\}$, we define the set $\mathcal{RP}_a$ of functions obeying the reflection principle on Θ_a by

$$\mathcal{RP}_a = \{f:\mathbb{R}^3 \to \mathbb{R} \mid \forall r \in \Theta_a : \theta_a f(r) \leqslant f(r)\}.$$

The set of 'reflection positive functions' is defined by

$$\mathcal{RP}_a^+ = \{f \in \mathcal{RP}_a \mid \forall r \in \Theta_a : f(r) \geqslant 0\}.$$

Note that $\mathcal{RP}_a^+$ is a positive cone (i.e. $\sum c_i f_i \in \mathcal{RP}_a^+$ if all $f_i \in \mathcal{RP}_a^+$ and $c_i \geqslant 0$) closed under multiplication. In addition, if $g:\mathbb{R} \to \mathbb{R}$, $g \geqslant 0$ is monotonically increasing and $f \in \mathcal{RP}_a^+$, then also $g \circ f \in \mathcal{RP}_a^+$. Any function $f:\mathbb{R}^3 \to \mathbb{R}$ can be decomposed into its even $(+)$ and odd $(-)$ components with respect to θ_a,

$$f^\pm = \tfrac{1}{2}(f \pm \theta_a f)$$

and clearly $f \in \mathcal{RP}_a^+$ implies $f^\pm \in \mathcal{RP}_a^+$.

The following 'correlation inequality' is more or less the first Griffiths inequality (cf also lemma 1 of Lieb and Simon (1978)) rewritten into a form that will be applied later.

Lemma 3.1. Consider $f_i \in \mathcal{RP}_a^+$, $i = 1,\ldots,n$ such that for the given set of coefficients $\alpha_{ij} < 0$, $i,j = 1,\ldots,n$ the integral in (3.1) below does exist. Then with $\mathrm{d}^n\zeta = \prod_{i=1}^n \mathrm{d}\zeta_i$, for all $r_k = \rho_k\hat{e}_\rho + \zeta_k\hat{e}_\zeta + \varphi_k\hat{e}_\varphi \in \mathbb{R}^3$

$$\int_{\mathbb{R}^n} \mathrm{d}^n\zeta \exp\Big(\sum_{i,j=1}^n \alpha_{ij}(\zeta_i - \zeta_j)^2\Big) \prod_{k=1}^n f_k(r_k) \geqslant 0. \tag{3.1}$$

Proof. Due to translation invariance $(\zeta_i - a - (\zeta_j - a))^2 = (\zeta_i - \zeta_j)^2$ we can assume that ϑ_a is a reflection with respect to the plane passing through the origin

$\zeta = 0$. Expanding the exponential and decomposing the f_k into even f_k^+ and odd f_k^- components with respect to θ_0, (3.1) becomes a sum of integrals of the form

$$\int_{\mathbb{R}^n} \mathrm{d}^n\zeta \prod_{\substack{l,m \\ l\neq m}} g_l^+(\zeta_l) g_m^-(\zeta_m)$$

that either vanish if there is a factor g^- in the product, or due to symmetry can be reduced to integrations over half-lines where the integrands are non-negative since all $g_l^+ \in \mathcal{RP}_0^+$. □

The next lemma shows that the potentials and their R derivatives are reflection positive if the reflection is taken with respect to the 'outer' nucleus (regarded from Θ_R). Since in the following also regularized potentials will be used, we formulate the statement in an appropriate generality:

Lemma 3.2. Let $\boldsymbol{R} \in \mathbb{R}^3$ be given by $\boldsymbol{R} = R\hat{e}_\zeta$ with $0 \leqslant R$. Then for all $\varepsilon \in \mathbb{R}$ the potential

$$V_\varepsilon(\boldsymbol{r}) = \frac{Z_1}{|\boldsymbol{r}| + \varepsilon^2} + \frac{Z_2}{|\boldsymbol{r} - \boldsymbol{R}| + \varepsilon^2} \tag{3.2}$$

and its derivative $-\mathrm{d}V_\varepsilon/\mathrm{d}R$ are reflection positive with respect to θ_R:

$$V_\varepsilon, -\frac{\mathrm{d}V_\varepsilon}{\mathrm{d}R} \in \mathcal{RP}_R^+.$$

Proof. Straightforward computation produces

$$V_\varepsilon(\boldsymbol{r}) - \theta_R V_\varepsilon(\boldsymbol{r}) = 4Z_1 R(R-\zeta)/\left[\left(\sqrt{\rho^2+\zeta^2}+\varepsilon^2\right)\left(\sqrt{\rho^2+(2R-\zeta)^2}+\varepsilon^2\right)\right.$$
$$\left.\times\left(\sqrt{\rho^2+\zeta^2}+\sqrt{\rho^2+(2R-\zeta)^2}\right)\right] \tag{3.3}$$

which is obviously non-negative if $\zeta \leqslant R$. Under the same condition, non-negativity holds for the difference

$$\theta_R \frac{\mathrm{d}V_\varepsilon}{\mathrm{d}R}(\boldsymbol{r}) - \frac{\mathrm{d}V_\varepsilon}{\mathrm{d}R}(\boldsymbol{r}) = \frac{2Z_2(R-\zeta)}{\left(\sqrt{\rho^2+(\zeta-R)^2}+\varepsilon^2\right)^2}$$

of the derivatives. □

Proof of corollary 1.2. To carry out the proof of corollary 1.2, as Lieb and Simon (cf also (2.17)) we express the energies e_m as $e_m = -\lim_{t\to\infty} t^{-1}\ln\langle\phi, \exp(-th_m)\phi\rangle$ where ϕ is any non-negative square-integrable function $\phi \not\equiv 0$. Taking $\phi \in C_0^\infty(\mathbb{R}^3)$ and using the differentiability of e_m (cf Thirring (1981) and references therein) we get

$$\frac{\mathrm{d}e_m}{\mathrm{d}R}(R) = \lim_{t\to\infty} t^{-1}\langle\phi, \mathrm{e}^{-th_m(R)}\phi\rangle^{-1}\int_0^t \mathrm{d}s\,\langle\phi, \mathrm{e}^{-sh_m(R)}\left(-\frac{\mathrm{d}V}{\mathrm{d}R}\right)\mathrm{e}^{-(t-s)h_m(R)}\phi\rangle \tag{3.4}$$

(the right-hand side of (3.4) converges uniformly in R thus justifying the exchange of $\lim_{t\to\infty}$ and $\mathrm{d}/\mathrm{d}R$). Furthermore, we choose ϕ of the form $\phi(\rho,\zeta,\varphi) = \phi^{\perp}(\rho)\phi^{\parallel}(\zeta)$ with $\phi^{\perp}(\rho) \geqslant 0$ and $\phi^{\parallel} \in \mathcal{RP}_R^+$. Approximating the potential in h_m by its regularized version (3.2) then allows the application of the Trotter product formula for the semigroups, yielding

$$\left\langle \phi, \mathrm{e}^{-sh_m(R)} \left(-\frac{\mathrm{d}V}{\mathrm{d}R}\right) \mathrm{e}^{-(t-s)h_m(R)}\phi \right\rangle = \lim_{\varepsilon\to 0} \lim_{\substack{n\to\infty \\ k\to\infty}} \int_0^{\infty} \prod_{i=-k}^{n} \rho_i \,\mathrm{d}\rho_i \; K^{\perp}_{t,s;k,n}(\rho^{(k,n)})$$

$$\times \left\{ \int_{-\infty}^{\infty} \prod_{j=-k}^{n} \mathrm{d}\zeta_j \; K^{\parallel}_{t,s;k,n}(\zeta^{(k,n)}) W^{\varepsilon}_{t,s;k,n}(\rho^{(k-1,n)};\zeta^{(k-1,n)}) \right\} \quad (3.5)$$

where we used the abbreviations $\rho^{(k,n)} = (\rho_{-k}, \rho_{-k+1}, \ldots, \rho_{n-1}, \rho_n)$ and similarly for $\zeta^{(k,n)}$, and where

$$K^{\perp}_{t,s;k,n}(\rho^{(k,n)}) = \phi^{\perp}(\rho_{-k})\mathrm{e}^{-(s/n)h^{\perp}_m(\rho_{-k},\rho_{-k+1})} \times \cdots$$

$$\times \mathrm{e}^{-(s/n)h^{\perp}_m(\rho_{-1},\rho_0)}\mathrm{e}^{-((t-s)/k)h^{\perp}_m(\rho_0,\rho_1)} \times \cdots$$

$$\times \mathrm{e}^{-((t-s)/k)h^{\perp}_m(\rho_{n-1},\rho_n)}\phi^{\perp}(\rho_n)$$

is pointwise non-negative by (ii) of theorem 2.1. The analogous quantity $K^{\parallel}$ contains explicit Gaussians

$$K^{\parallel}_{t,s;k,n}(\zeta^{(k,n)}) = \left(\frac{4\pi s}{n}\right)^{-n/2} \left(\frac{4\pi(t-s)}{k}\right)^{-k/2} \phi^{\parallel}(\zeta_{-k})$$

$$\times \exp\left(-\frac{(\zeta_{-k}-\zeta_{-k+1})^2}{4s/k}\right) \times \cdots \times \exp\left(-\frac{(\zeta_{-1}-\zeta_0)^2}{4s/k}\right)$$

$$\times \exp\left(-\frac{(\zeta_0-\zeta_{-1})^2}{4(t-s)/n}\right) \times \cdots \times \exp\left(-\frac{(\zeta_{n-1}-\zeta_n)^2}{4(t-s)/n}\right) \phi^{\parallel}(\zeta_n)$$

and $W^{\varepsilon}_{t,s;k,n}$ is given by

$$W^{\varepsilon}_{t,s;k,n}(\rho^{(k-1,n)};\zeta^{(k-1,n)}) = \mathrm{e}^{(s/k)V_{\varepsilon}(\rho_{-k+1},\zeta_{-k+1})} \times \cdots \times \mathrm{e}^{(s/k)V_{\varepsilon}(\rho_0,\zeta_0)}$$

$$\times \left(-\frac{\mathrm{d}V}{\mathrm{d}R}(\rho_0,\zeta_0)\right) \mathrm{e}^{((t-s)/n)V_{\varepsilon}(\rho_1,\zeta_1)} \times \cdots \times \mathrm{e}^{((t-s)/n)V_{\varepsilon}(\rho_n,\zeta_n)}.$$

The non-negativity of $\{\ldots\}$ in (3.5) and thus also $(\mathrm{d}e_m/\mathrm{d}R)(R) \geqslant 0$ follows now from the observation that all terms in $\{\ldots\}$ satisfy the prerequisites of lemma 3.1, completing the proof of corollary 1.2. □

4. Monotonicity for other states and systems

For $N > 2$ and arbitrary configurations $\{R_{ij}\}_{1 \leqslant i < j \leqslant N}$, a generalization of lemma 3.2 (concerning $\mathrm{d}V/\mathrm{d}R$) is unfortunately not straightforward. Nevertheless, if only an outermost nucleus is moved such that, for example, $\boldsymbol{X}_i' = \boldsymbol{X}_i$ and $R_{iN}' \geqslant R_{iN}$ for $i = 1, \ldots, N-1$, then in the corresponding derivative of the potential only the term linked with the Nth nucleus appears, and lemma 3.2 can be simply extended to such a situation if the reflection is with respect to this outer nucleus. As a consequence, for these configurations (which include the general $N = 3$ case) theorem 1.1 also follows from a proof based on the reflection principle. Moreover, all statements can be generalized to potentials other than Coulomb potentials if these potentials allow the partial separation induced by the cylinder symmetry and obey the prerequisites assumed in Lieb's proof or the reflection positivity of section 3. The monotonicity proofs of Lieb and Simon, Hoffmann–Ostenhoff, and that of section 3 require differentiability (in R) of the electronic curves; to avoid this condition, one might try to modify the 'infinitesimal' nature of these proofs by estimating the difference $e_m(R') - e_m(R)$ instead of the derivative $\mathrm{d}e_m/\mathrm{d}R$. However, due to the behaviour of $V(R) - V(R')$ on the interval $\zeta \in [R, R']$, the desired extension of lemma 3.2 is not obvious.

An attempt to establish monotonicity for more electronic curves has to face the fact that there are states whose electronic energies definitely do not behave monotonically (see also the remark (4.6.26;3) in Thirring (1981)). On the other hand, for $N = 2$ numerical experience (Hogreve 1992) suggests the existence of a larger class of states for which electronic monotonicity holds. To characterize these states, recall that one-electron $N = 2$ systems are completely separable in prolate spheroidal coordinates, and in the united atoms limit $R = 0$ all states become hydrogenic-like (with $Z = Z_1 + Z_2$). Therefore we can label the electronic energies e_{nlm} by the usual hydrogen quantum numbers n, l, m. Then, motivated by the numerical results, we put forward the following conjecture.

Conjecture 4.1. For all $n \in \mathbb{N}$, $l \in \mathbb{N} \cup \{0\}$, $m \in \mathbb{Z}$ with $n \geqslant |m| + 1$, $l = |m|$ the electronic curves $e_{nlm}(R)$ are monotonically increasing in R.

In corollary 1.2 we demonstrated the monotonicity for states with $n = |m| + 1$, $l = |m|$. The reason preventing a direct extension of the proofs of the preceding sections to the more general assertion of conjecture 4.1 is related to the following. Despite complete separability, the energies $e_{nlm}(R)$ are only determined as implicit functions because a (non-explicit) separation constant $A_{nlm}(R)$ is also involved. The prolate spheroidals (ξ, η, φ) are, in a certain sense, not 'good' coordinates for a monotonicity proof, since neither ξ nor η is parallel to the ζ-variable (of cylindrical coordinates) whereas our arguments in sections 2 and 3 showed that just the behaviour in the ζ-variable is decisive for the monotonicity.

For systems with more than one electron—as already mentioned by Lieb and Simon—the electron–electron repulsion may also cause additional problems for the validity of the electronic monotonicity. Nevertheless, at least if the system is neutral or has a positive total charge, not only the ground-state energy can be expected to behave monotonically, but a monotonicity of the more general kind as in theorem 1.1 should hold.

Conjecture 4.2. Theorem 1.1 remains valid for the electronic energies of an n-electron molecule, i.e. if the kinetic and potential operators in (1.2) are replaced by

$$h^{\text{kin}} = -\sum_{i=1}^{n} \Delta_i$$

$$V(r_1, \ldots, r_n) = \sum_{i=1}^{N} \sum_{j=1}^{n} v(r_j; X_i, Z_i) - \sum_{i<j}^{n} |r_i - r_j|^{-1}$$

and if $n \leqslant \sum_{i=1}^{N} Z_i$.

For $n = 2$ and electronic singlet states, it might be interesting to prove conjecture 4.2 using the methods of Lieb (1982) and the present study. More intricate situations where the state is no longer represented by a nodeless wavefunction (at least away from a symmetry axis), however, certainly call for other and probably more elaborate strategies.

Acknowledgment

I am grateful to Dr P Duclos for his hospitality at the CPT.

References

Araki H 1990 *Lett. Math. Phys.* **19** 167

Duclos P and Hogreve H 1991 On the stability of positive diatomic molecular ions *Proc. Conf. Rigorous Results in Quantum Dynamics (Liblice, 1990)* ed J Dittrich and P Exner (Singapore: World Scientific) pp 63–70

—— 1992 On the monotonicity of abiabatic curves of the one-electron diatomic molecular ion, to be submitted, and work in progress

Hegerfeldt G C 1977 *Commun. Math. Phys.* **57** 259

Hoffmann-Ostenhoff T 1980 *J. Phys. A: Math. Gen.* **13** 417

Hogreve H 1992 On the stability of the one-electron bond *J. Chem. Phys.* (to appear); 1992 On the binding and binding-metamorphosis of the $Z_1 Z_2 e^-$ system, to be submitted

Jörgens K and Rellich F 1976 *Eigenwerttheorie gewöhnlicher Differentialgleichungen* (Berlin: Springer)

Lieb E H 1982 *J. Phys. B: At. Mol. Phys.* **15** L63

Lieb E H and Simon B 1978 *J. Phys. B: At. Mol. Phys.* **11** L537

Lieb E H and Thirring W 1976 Inequalities for the moments of the eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities *Studies in Mathematical Physics, Essays in Honor of Valentine Bargmann* ed E H Lieb, B Simon and A Wightman (Princeton: Princeton University Press)

Messager A and Miracle-Sole S 1977 *J. Stat. Phys.* **17** 245

Miller W Jr 1977 *Symmetry and Separation of Variables (Encyclopedia of Mathematics and its Applications 4)* ed G C Rota (Reading: Addison-Wesley)

Nelson E 1973 *Constructive Quantum Field Theory* ed G Velo and A Wightman (Berlin: Springer)

Olver F W J 1974 *Asymptotics and Special Functions* (New York: Academic)

Reed M and Simon B 1975 *Methods of Modern Mathematical Physics II, Fourier Analysis, Self-Adjointness* (New York: Academic)

—— 1978 *Methods of Modern Mathematical Physics IV, Analysis of Operators* (New York: Academic)

—— 1980 *Methods of Modern Mathematical Physics I, Functional Analysis* 2nd edn (New York: Academic)

Schrader R 1977 *Phys. Rev.* B **15** 2798

Thirring W 1981 *A Course in Mathematical Physics vol 3, Quantum Mechanics of Atoms and Molecules* (Wien: Springer)

Tinkham M 1964 *Group Theory and Quantum Mechanics* (New York: McGraw-Hill)

Weidmann J 1980 *Linear Operators in Hilbert Spaces* (New York: Springer)