

Nonlinear Schrodinger operators and molecular structure

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1995 J. Phys. A: Math. Gen. 28 4025

(<http://iopscience.iop.org/0305-4470/28/14/021>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.68

The article was downloaded on 02/06/2010 at 00:12

Please note that [terms and conditions apply](#).

Nonlinear Schrödinger operators and molecular structure

E B Davies†

Department of Mathematics, King's College, Strand, London WC2R 2LS, UK

Received 29 November 1994

Abstract. We minimize the energy of a quantum Hamiltonian on the tensor product of two Hilbert spaces within the class of product states. This yields a nonlinear Schrödinger equation, whose ground state may bifurcate, producing symmetry breaking. We describe a procedure for computing the ground states numerically, and prove that it converges. We argue that the nonlinear Schrödinger equation has relevance to the issue of molecular structure in quantum chemistry, and study an exactly soluble example in detail to support this claim. The paper concludes with a brief discussion of other approaches to molecular structure.

1. Introduction

It is well known that the Hilbert space associated with the composite system C obtained by combining two spatially separated quantum systems A and B is

$$\mathcal{H}_C := \mathcal{H}_A \otimes \mathcal{H}_B$$

in an obvious notation. In principle, any vector of norm one in \mathcal{H}_C describes a pure state of C . We shall call those of the form $\phi_C := \phi_A \otimes \phi_B$ product states, and refer to other states as entangled, to a greater or lesser extent depending on how close they are to product states.

Simplicity suggests that one should assume that the state of a composite system is a product state, in the absence of any contrary information. However, this is not always an appropriate choice. In particular, the EPR paradox rests upon the possibility of creating highly entangled states, whose properties under measurement are very different from corresponding results for product states. If A represents some collection of atoms or molecules interacting with each other, and B represents the external environment, then one is almost forced into the use of a product state ansatz, even though it is often implicit, in that the environment is not actually mentioned. Putting it the other way around, if entanglements between the system studied and the environment are important in some context, then one has to make the division between system and environment in a different manner than that originally considered.

Now suppose that \mathcal{H}_A and \mathcal{H}_B are the Hilbert spaces describing two interacting molecules. In this paper we shall examine the thesis that it is appropriate to use product states to investigate the combined system, at least as a first approximation. The obvious defence of this proposition is that if such an approximation is completely invalid, then the collection of atoms and electrons in C must be regarded as a single molecule rather than two separate ones. We shall see that the use of product states provides a bridge

† E-mail address: udah210@bay.cc.kcl.ac.uk

between treating molecules as completely isolated non-interacting systems and handling the technically difficult problems associated with an exact quantum mechanical treatment.

Our approach can also be described as imposing coarse graining or decoherence [O, Z] between the two Hilbert spaces, or as localizing the particles associated with each of the two spaces. This is followed by the detailed analysis of the eigenstates of a nonlinear Schrödinger operator associated with this procedure. These eigenstates are of interest in their own right, but can also be used as a basis for expansion of the true eigenfunctions. We pursue these ideas partly theoretically, and partly through the numerical investigation of a simple model. A more philosophical discussion of our results is given in the final section.

The product state ansatz leads to the study of a certain nonlinear Schrödinger equation. This equation is associated with a non-convex energy functional which has been studied by a number of people since it was first introduced [D1, D2, D3, GM, P1, P2, P3]. The number of different derivations confirms the significance of the energy functional as a way of taking some account of external influences on an open quantum system. We make some further progress in the general theory of the functional, and illustrate it by solving a model problem describing two interacting molecules. We show, at least in the model, that use of the nonlinear equation yields remarkably good numerical results, considering the simplicity of the method.

It has been pointed out in earlier papers on this energy functional that its properties throw some light on the problems of molecular structure and optical activity. The first discussion of molecular structure was given by Hund, who explained why optically active states of certain molecules might be very long lived [H]. The problem was revived by Woolley [W1, W2], and the current issue is to explain why optically active states of certain molecules seem to be positively preferred to the true quantum mechanical ground state. In the author's view the best current explanation uses the instability of the true ground states of these molecules under small external perturbations. Various types of external perturbations have been considered in this context and all lead to rather similar conclusions [CJ, JMS, P1, P2, P3].

Wightman and Glance [WG] have summarized one group of papers on molecular structure by concentrating on the idea of superselection rules which put different optical isomers into distinct superselection sectors. These superselection sectors only exist for systems with an infinite number of degrees of freedom, but statistical mechanics and quantum field theory provide standard tools for formulating such problems in precise terms. A difficulty with this approach is that it depends upon the solution of substantial technical problems [A1, S] even for highly simplified models. It is not clear whether more realistic models exhibit the phase transitions which the approach requires [WG].

The behaviour of a large assembly of molecules can be studied at four levels of approximation. The simplest is to consider a single molecule, completely ignoring the influence of its neighbours. A better procedure is to model the environment of one of the molecules by a classical polarizing medium or by a random external electric field, and to compute the effects of the additional terms. The third approximation considers all of the molecules quantum mechanically, but involves minimizing the energy within the class of product states. The final approach is the very difficult problem of finding the true ground-state energy for the full interaction Hamiltonian of the multi-particle system.

We shall study the third of these models. Our results indicate the existence of symmetry-breaking metastable states at this level of approximation. Although these results are artifacts of the method, it is interesting that they are so closely related to one's intuition about the behaviour of molecules. We believe that they are likely to be indications of the existence of phase transitions and superselection rules for the full statistical mechanical models.

In section 4 of the paper we consider a highly idealized model of a molecule consisting of a rotating covalent bond joining two groups, each of which has four possible orientations relative to the bond. We obtain expressions for the exact ground state, the lowest energy product states, a symmetrized approximate ground state and their associated energies. The striking numerical accuracy of our method provides support for the value of the method in other more realistic problems in which the exact solution is not readily computed.

2. The nonlinear energy functional

Let X be a periodic lattice in \mathbf{R}^N , and let \mathcal{H} denote the Hilbert space describing a collection of one or more atoms or molecules with Hamiltonian H . We put a copy \mathcal{H}_x of \mathcal{H} at each site $x \in X$, and let $K := K_0 + K_I$ denote the total self-energy of the system of molecules in the Hilbert space $\mathcal{K}_X := \otimes_{x \in X} \mathcal{H}_x$. Here $K_0 := \sum_{x \in X} H_x$, and K_I is the interaction energy of the molecules. We assume that the interaction is of dipole type, that is

$$K_I := - \sum_{(x,y) \in \Lambda} \sum_{r=1}^n A_{x,r} \otimes A_{y,r}$$

where $A_{x,r}$ are copies in \mathcal{H}_x of bounded self-adjoint operators A_r on \mathcal{H} , and Λ is the set of nearest neighbour edges in the lattice X . More complicated edge dependence of the interactions can be accommodated within our theory.

We comment that the simplest case in which X consists of just two points, is still of interest. It describes the interaction of two atoms or molecules, and exhibits the symmetry-breaking effect just as the statistical mechanical model does.

The ground-state energy density of the system is by definition the quantity

$$E_X := \inf\{\langle Kf, f \rangle / |X| : f \in \mathcal{K} \text{ and } \|f\| = 1\}$$

where $|X|$ is the number of sites in X . If we minimize over product states of the form $f := \otimes_{x \in X} \phi_x$, where ϕ_x are all copies of a single unit vector $\phi \in \mathcal{H}$, then we obtain the inequality

$$E_X \leq \bar{E}_\lambda \leq \inf\{\langle H\phi, \phi \rangle : \phi \in \mathcal{H} \text{ and } \|\phi\| = 1\}$$

where

$$\bar{E}_\lambda := \inf\{\mathcal{E}_\lambda(\phi) : \phi \in \mathcal{H} \text{ and } \|\phi\| = 1\} \tag{1}$$

and

$$\mathcal{E}_\lambda(\phi) := \langle H\phi, \phi \rangle - \lambda \sum_{r=1}^n \langle A_r \phi, \phi \rangle^2 \tag{2}$$

where λ is the number of edges per lattice site.

We emphasize that sites in X should not always be identified with individual atoms or molecules. We argue in the last section of the paper that one should group atoms together into clusters whose diameter is a certain coherence length determined by the environment. The justification of the above product ansatz within the context of equilibrium statistical mechanics is discussed at length in the last section.

We have already mentioned that the functional (2) above has been studied by a number of authors with different justifications for its use. The above derivation is, however, new and has the feature of not introducing any undetermined parameters. The technical content of the present paper consists of an analysis of the critical points of the functional, and procedures for determining its local minima. We refer to [D3] for related results using group symmetries to simplify the analysis, and to [D2] for the study of the analogous formula at positive temperatures. We leave the proof of the following easy result to the reader.

Lemma 1. The infimum in (1) is actually attained for at least one unit vector $\phi \in \mathcal{H}$. The quantity \tilde{E}_λ is a concave decreasing function of the parameter λ .

3. Analysis of critical points

Throughout this section we consider the functional (2) under the assumptions that \mathcal{H} is finite-dimensional, that H and A_r are all self-adjoint operators and that $\lambda \geq 0$. We understand that \mathcal{E} is always applied to states, that is to $f \in \mathcal{H}$ with unit norm, where two such vectors are identified if they differ only by a phase factor.

Lemma 2. The critical points f of \mathcal{E} are precisely those states f which satisfy the eigenvalue equation

$$Hf - 2\lambda \sum_{r=1}^n \alpha_r A_r f = \mu f \quad (3)$$

for some $\mu \in \mathbf{R}$, where

$$\alpha_r := \langle A_r f, f \rangle. \quad (4)$$

Proof. A state f is a critical point if

$$\left. \frac{d\mathcal{E}(f(t))}{dt} \right|_{t=0} = 0$$

where

$$f(t) := \frac{f + tg}{\|f + tg\|} \quad (5)$$

and $g \in \mathcal{H}$ is any vector satisfying $\langle f, g \rangle = 0$. On performing the necessary differentiations, we find that this is equivalent to

$$\langle Hf - 2\lambda \sum_{r=1}^n \alpha_r A_r f, g \rangle = 0$$

for all such g . This, in turn, is equivalent to (3). □

Lemma 3. Let f be a critical point of \mathcal{E} satisfying (3) and (4). If we define

$$\tilde{H} := H - 2\lambda \sum_{r=1}^n \alpha_r A_r - \mu \tag{6}$$

$$\tilde{A}_r := A_r - \alpha_r 1 \tag{7}$$

for $1 \leq r \leq n$, then

$$\tilde{H}f = 0 \quad \langle \tilde{A}_r f, f \rangle = 0$$

for all $1 \leq r \leq n$, and

$$\mathcal{E}(g) = \langle \tilde{H}g, g \rangle - \lambda \sum_{r=1}^n \langle \tilde{A}_r g, g \rangle^2 + c$$

for all states g , where

$$c := \lambda \sum_{r=1}^n \alpha_r^2 + \mu.$$

We omit the proof, which involves only straightforward algebra. We call the above expressions the normalized representation of \mathcal{E} relative to the critical point f . We say that a critical point f of \mathcal{E} is a non-degenerate local minimum if

$$\left. \frac{d^2 \mathcal{E}(f(t))}{dt^2} \right|_{t=0} > 0$$

for all $f(t)$ of the form (4) and all $g \in \mathcal{H}$ such that $\langle f, g \rangle = 0$.

Theorem 4. Suppose that the state f satisfies

$$Hf = 0 \quad \langle A_r f, f \rangle = 0$$

for all $1 \leq r \leq n$. Then f is a non-degenerate local minimum of

$$\mathcal{E}(\phi) := \langle H\phi, \phi \rangle - \lambda \sum_{r=1}^n \langle A_r \phi, \phi \rangle^2$$

if the following two conditions hold:

(i) H is strictly positive on $\mathcal{L} := \{f\}^\perp$.

(ii) The $n \times n$ matrix

$$K_{r,s} := \langle H^{-1} A_r f, A_s f \rangle$$

satisfies $4\lambda \|K\| < 1$.

Note. These conditions are related to those of Grecchi and Martinez [GM], but are perhaps more transparent.

Proof. In the notation of (5) we have

$$\begin{aligned} \mathcal{E}(f(t)) &= \|f + tg\|^{-2} \langle H(f + tg), f + tg \rangle - \lambda \|f + tg\|^{-4} \sum_{r=1}^n \langle A_r(f + tg), f + tg \rangle^2 \\ &= (1 - t^2 \|g\|^2) t^2 \langle Hg, g \rangle \\ &\quad - \lambda (1 - 2t^2 \|g\|^2) t^2 \sum_{r=1}^n \{ \langle A_r f, g \rangle + \langle A_r g, f \rangle \}^2 + O(t^3) \\ &= ct^2 + O(t^3) \end{aligned}$$

where

$$\begin{aligned} c &:= \langle Hg, g \rangle - \lambda \sum_{r=1}^n \{ \langle A_r f, g \rangle + \langle A_r g, f \rangle \}^2 \\ &\geq \langle Hg, g \rangle - 4\lambda \sum_{r=1}^n | \langle A_r f, g \rangle |^2 \\ &= \langle (H - 4\lambda C)g, g \rangle \end{aligned} \tag{8}$$

where

$$C := \sum_{r=1}^n |A_r f\rangle \langle A_r f|.$$

If we put $\mathcal{L} := \{g \in \mathcal{H} : \langle g, f \rangle = 0\}$, then it is easy to see that $(H - 4\lambda C)(\mathcal{L}) \subseteq \mathcal{L}$. Moreover $c > 0$ for all non-zero $g \in \mathcal{L}$ if the restriction of $H - 4\lambda C$ to \mathcal{L} is a strictly positive operator.

The remainder of the proof involves showing that

$$\langle (H - 4\lambda C)g, g \rangle > 0 \tag{9}$$

for all non-zero $g \in \mathcal{L}$ if and only if the conditions (i) and (ii) of the theorem hold. The condition (i) is an easy consequence of (9) and the inequality $C \geq 0$, and will be assumed below. The condition (9) may be rewritten in the form

$$A^*A = H^{-1/2}CH^{-1/2} \leq \theta I \tag{10}$$

on \mathcal{L} , for some $\theta < 1/(4\lambda)$, where $A : \mathcal{L} \rightarrow \mathbb{C}^n$ is defined by

$$(Ag)_r := \langle H^{-1/2}g, A_r f \rangle.$$

A well known result states that (10) is equivalent to the inequality $A^*A \leq \theta I$ on \mathbb{C}^n . Since this operator has the matrix K , we see that the condition is equivalent to (ii).

Theorem 5. Suppose that the real-valued state f satisfies

$$Hf = 0 \quad \langle A_r f, f \rangle = 0$$

for all $1 \leq r \leq n$, and that H, A_r are all real-symmetric matrices. Then f is a non-degenerate local minimum of

$$\mathcal{E}(\phi) := \langle H\phi, \phi \rangle - \lambda \sum_{r=1}^n \langle A_r \phi, \phi \rangle^2$$

if and only if the conditions (i) and (ii) of theorem 4 are both satisfied.

Proof. We follow the notation of the last proof. If f is a non-degenerate local minimum of \mathcal{E} and $g \in \mathcal{L}$ is real then the proof of (8) yields

$$c = \langle Hg, g \rangle - 4\lambda \sum_{r=1}^n \langle A_r f, g \rangle^2$$

$$= \langle (H - 4\lambda C)g, g \rangle.$$

Since $H - 4\lambda C$ is a real-symmetric matrix, it follows that if $c > 0$ for all non-zero real $g \in \mathcal{L}$ then $H - 4\lambda C$ is strictly positive; conversely if $H - 4\lambda C$ is strictly positive then $c > 0$ for all non-zero complex $g \in \mathcal{L}$ by the argument of theorem 4. The remainder of the proof follows theorem 4.

If this type of nonlinear Schrödinger equation is to be computationally useful, it is essential to develop procedures for finding the local minima of \mathcal{E} . We next investigate one iterative scheme for doing this. A short discussion of this scheme is given after corollary 8. Starting from a sequence $\alpha^{(m)} := (\alpha_1^{(m)}, \dots, \alpha_n^{(m)})$, let f_m be the ground state of the operator

$$H_m := H - \sum_{r=1}^n \alpha_r^{(m)} A_r$$

with corresponding eigenvalue μ_m , and put

$$\alpha_r^{(m+1)} := \langle A_r f_m, f_m \rangle$$

for $1 \leq r \leq n$. The precise meaning of the following lemma will emerge during its proof. It is of theoretical rather than computational value, since it depends upon the knowledge of the limit state f .

Lemma 6. If f is a critical point of \mathcal{E} , then the issue of convergence of $\alpha^{(m)}$ to α may be solved in the normalized representation of \mathcal{E} relative to f .

Proof. Suppose that H, λ, A_r and the limit state f are known and that $f_m, \alpha^{(m)}$ and μ_m satisfy

$$Hf_m - 2\lambda \sum_{r=1}^n \alpha_r^{(m)} A_r f_m = \mu_m f_m \quad \langle A_r f_m, f_m \rangle = \alpha_r^{(m+1)}.$$

A direct substitution of (6) and (7) establishes that

$$\tilde{H} f_m - 2\lambda \sum_{r=1}^n \alpha_r^{(m)} \tilde{A}_r f_m = \tilde{\mu}_m f_m$$

where $\tilde{\alpha}_r^{(m)} := \alpha_r^{(m)} - \alpha_r$ satisfy

$$\tilde{\alpha}_r^{(m+1)} = \langle \tilde{A}_r f_m, f_m \rangle.$$

Thus $f_m, \tilde{\alpha}^{(m)}$ and $\tilde{\mu}_m$ satisfy the 'same' recurrence relations as before but with respect to the renormalized operators. The convergence of $\alpha^{(m)}$ to α is equivalent to the convergence of $\tilde{\alpha}^{(m)}$ to 0 as $m \rightarrow \infty$.

Theorem 7. Suppose that the state f satisfies

$$Hf = 0 \quad \langle A_r f, f \rangle = 0$$

for all $1 \leq r \leq n$, and that the following two conditions hold:

(i) H is strictly positive on $\mathcal{L} := \{f\}^\perp$.

(ii) The $n \times n$ matrix

$$M_{r,s} := \operatorname{Re}\{ \langle H^{-1} A_r f, A_s f \rangle \} \tag{11}$$

satisfies $4\lambda \|M\| < 1$. Suppose also that $f_m, \alpha^{(m)}$ and μ_m satisfy the above recurrence relations. Then $\lim_{m \rightarrow \infty} \alpha^{(m)} = 0$ if $\alpha^{(1)}$ is sufficiently small.

Proof. Condition (i) implies that 0 is an eigenvalue of multiplicity one of H with eigenvector f . Analytic perturbation theory implies that if $\alpha \in \mathbb{R}^n$ is sufficiently small then the smallest eigenvalue $\mu(\alpha)$ of $H - 2\lambda\alpha \cdot A$ is of multiplicity one. Moreover, $\mu(\alpha)$ and the corresponding eigenvector $f(\alpha)$ satisfy

$$\mu(\alpha) = \alpha \cdot \gamma + O(|\alpha|^2) \quad f(\alpha) = f + \alpha \cdot g + O(|\alpha|^2)$$

where $\gamma \in \mathbb{R}^n$ and $\langle g, f \rangle = 0$. Therefore

$$(H - 2\lambda\alpha \cdot A)(f + \alpha \cdot g) = \alpha \cdot \gamma(f + \alpha \cdot g) + O(|\alpha|^2)$$

and

$$H g_r - 2\lambda A_r f = \gamma_r f$$

for $1 \leq r \leq n$. Taking inner products with f yields $\gamma_r = 0$, and hence

$$g_r = 2\lambda H^{-1} A_r f.$$

If we now define the nonlinear map $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ by

$$(T\alpha)_r := \langle A_r f(\alpha), f(\alpha) \rangle$$

then

$$\begin{aligned} (T\alpha)_r &= \langle A_r(f + \alpha \cdot g), (f + \alpha \cdot g) \rangle + O(|\alpha|^2) \\ &= 2\operatorname{Re}\langle A_r f, \alpha \cdot g \rangle + O(|\alpha|^2) \\ &= 4\lambda M\alpha + O(|\alpha|^2) \end{aligned}$$

where M is defined by (11). If $4\lambda \|M\| < c < 1$ then $|T\alpha| \leq c|\alpha|$ provided $|\alpha|$ is small enough. Therefore $\alpha^{(n)} := T^{n-1}\alpha^{(1)}$ converges to zero provided $\alpha^{(1)}$ is small enough.

Corollary 8. Let f be a real local minimum of a functional \mathcal{E} defined using real-symmetric matrices H and A_r and real constants $\lambda > 0$ and α_r . If $f_m, \alpha^{(m)}$ and μ_m satisfy the above recurrence relations, then $\lim_{m \rightarrow \infty} \alpha^{(m)} = \alpha$ if $|\alpha^{(1)} - \alpha|$ is sufficiently small.

Proof. We use lemma 6 to transfer to the normalized case, noting that this preserves the reality of the various matrices. We then use theorem 5 to prove the hypotheses of theorem 7.

The above corollary shows that each local minimum $\alpha \in \mathbf{R}^n$ has a domain of attraction under the above recurrence relation. This is an open set which contains a neighbourhood of the point α itself. On the basis of limited numerical experiments, we conjecture that the union of the domains of attraction is dense in \mathbf{R}^n . Numerically this means that starting from essentially any initial point $\alpha \in \mathbf{R}^n$, repetition of the iterative scheme always leads to a local minimum. Some general results comparing the rate of convergence of this method with that of the steepest descent method applied to \mathcal{E} would be interesting.

We next suppose that G is a finite group of symmetries of the nonlinear functional \mathcal{E} in the following sense. We assume that G has a unitary representation U on the Hilbert space \mathcal{H} , and that H commutes with $U(g)$ for all $g \in G$. We relabel the subscript r of the operators A_r by points from a finite set X and suppose that G acts as a permutation group on X . We assume that

$$A_{gx} = U(g)A_xU(g)^*$$

for all $x \in X$ and all $g \in G$. The operator subspace $\text{lin}\{A_x : x \in X\}$ is then invariant under the action of G , and we finally assume that the decomposition of the representation of G on this space into irreducibles does not contain the identity representation. This is weaker than the assumption of [D3] that the representation of G on $\text{lin}\{A_x : x \in X\}$ is irreducible and different from the identity representation.

The simplest case of the above assumptions happens when there is a single spatial symmetry S of order two acting on \mathcal{H} , for which $SH = HS$ and $A_rS = -SA_r$ for all r .

It is known from a variety of examples that the minimum value of $\mathcal{E}(\phi)$ need not be achieved at a unique point, and that the individual minima may break the group symmetries described above. This phenomenon is relevant to optical and structural isomerism. The following theorem is a slight improvement upon Theorem 6 of [D3], and is included for completeness.

Theorem 9. Under the above hypotheses, every G -invariant eigenfunction of H is a critical point of the functional \mathcal{E} . Suppose that $H \geq 0$ and that there exists $\phi \in \mathcal{H}$ with $\|\phi\| = 1$ and $\mathcal{E}(\phi) < 0$. Then the minimum value of \mathcal{E} is achieved at more than one point and G acts non-trivially on the set of all minima. In other words the global minima of \mathcal{E} break the group symmetry.

Proof. Let $S := \{\psi \in \mathcal{H} : \|\psi\| = 1\}$, and let $\phi \in S$ be G -invariant in the weak sense that for all $g \in G$ we have

$$U(g)\phi = e^{i\theta(g)}\phi \tag{12}$$

where $\theta(g) \in \mathbf{R}$. Then the linear map $T : \mathcal{L} = \text{lin}\{A_x : x \in X\} \rightarrow \mathbf{C}$ defined by

$$T(B) := \langle B\phi, \phi \rangle$$

satisfies

$$T(U(g)BU(g)^*) = \langle BU(g)^*\phi, U(g)^*\phi \rangle = T(B)$$

for all $B \in \mathcal{L}$. Therefore $\mathcal{L}_0 := \{B \in \mathcal{L} : T(B) = 0\}$ is G -invariant, and the action of G on $\mathcal{L}_1 := \mathcal{L} \cap \mathcal{L}_0^\perp$ is trivial. Our assumption on this action implies that $\mathcal{L}_1 = 0$, and hence that $\mathcal{L}_0 = \mathcal{L}$. In other words $\langle A_x \phi, \phi \rangle = 0$ for all $x \in X$.

If ϕ is also an eigenvector of H then it follows from lemma 2 that it is a critical point of \mathcal{E} . If ϕ is a critical point of \mathcal{E} in addition to (12) then it follows from lemma 2 that it is an eigenfunction of H . Since $H \geq 0$, we deduce that $\mathcal{E}(\phi) \geq 0$, so \mathcal{E} does not take its minimum value at ϕ .

4. A soluble model

Application of the above computational scheme to problems in quantum chemistry will inevitably involve a great amount of numerical work. The purpose of this section is to describe a simple model exhibiting the main phenomena which can occur. We shall see that the numerical results obtained by our method are remarkably good considering its simplicity.

Although the derivation of the non-convex energy functional in section 2 was phrased in terms of a lattice of sites in \mathbb{R}^N , we commented that it could also be applied to a pair of identical systems. We make this choice in our model, in order that the exact eigenvalues and their nonlinear approximations can be calculated in closed form for comparison.

If we put $\lambda = \mu$ below, then the model can be regarded as a simplified description of a single molecule consisting of two parts joined by a single rotating covalent bond. The simplest molecule of this type, namely ethane, could be modelled very crudely by the Hilbert space $\mathbb{C}^3 \otimes \mathbb{C}^3$, one copy of \mathbb{C}^3 for each end of the molecule.

The molecule methyl hydrazine has more features in common with the model we consider, although the two ends are not identical. Lattimer and Harmony [LH] have computed the potential energy function associated with the different configurations of the molecule obtained by rotating around the N-N bond. This potential has two maxima and two minima as a function of the angle, and these configurations are at angles of approximately 90° to each other. A possible Hilbert space for a very approximate study of this molecule is therefore $\mathbb{C}^4 \otimes \mathbb{C}^4$.

In our highly simplified model we assume that $\mathcal{H} := \mathbb{C}^4$ and construct a Hamiltonian using the matrices

$$H := \begin{pmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{pmatrix}$$

$$A := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad B := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

We consider the anisotropic Schrödinger operator on $\mathcal{H} \otimes \mathcal{H}$ defined by

$$K := \frac{1}{2}(H \otimes 1 + 1 \otimes H) - \lambda A \otimes A - \mu B \otimes B$$

where $\lambda \geq 0$ and $\mu \geq 0$. The first term on the RHS is a discrete kinetic energy operator, and the other two terms define a potential with four local minima.

If we put

$$\mathcal{E}(\phi) := \langle K\phi \otimes \phi, \phi \otimes \phi \rangle$$

where $\phi \in \mathcal{H}$ and $\|\phi\| = 1$ then

$$\mathcal{E}(\phi) = \langle H\phi, \phi \rangle - \lambda \langle A\phi, \phi \rangle^2 - \mu \langle B\phi, \phi \rangle^2.$$

The operator K is obviously invariant with respect to the interchange of the two copies of \mathcal{H} . It is also invariant with respect to the rotational symmetry $R \otimes R$ where

$$R := \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

provided $\lambda = \mu$. This last property is inherited by the functional \mathcal{E} . However, even in the case $\lambda \neq \mu$ there are two commuting reflection symmetries, each of order two, induced by

$$S := \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad T := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

The influence of these symmetries on the calculations will be kept in the background, but has been essential in finding the various eigenvectors. We will constantly refer below to expansions in terms of the orthonormal basis

$$e_1 := \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad e_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} \quad e_3 := \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \quad e_4 := \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

consisting of simultaneous eigenvectors of S and T .

Lemma 10. The unit vector e_1 is a non-degenerate local minimum of the functional \mathcal{E} , with zero energy, if and only if $\lambda < 1$ and $\mu < 1$.

Proof. Direct computations establish the identities $He_1 = 0$, $He_2 = 2e_2$, $He_3 = 2e_3$ and $He_4 = 4e_4$. If we put $A_1 := \lambda^{1/2}A$ and $A_2 := \mu^{1/2}B$ then $A_1e_1 = \sqrt{(\lambda/2)}e_2$ and $A_2e_1 = \sqrt{(\mu/2)}e_3$. We may now apply the results of theorems 4 and 5. Since

$$K := \begin{pmatrix} \lambda/4 & 0 \\ 0 & \mu/4 \end{pmatrix}$$

it follows that $K \leq \theta 1$ for some $\theta < 1/4$ if and only if $\lambda < 1$ and $\mu < 1$.

Lemma 11. If $\lambda \leq 1$ and $\mu \leq 1$ then e_1 is a global minimum of the functional \mathcal{E} .

Proof. Since $\mathcal{E}(e_1) = 0$, we need to prove that $\mathcal{E}(f) \geq 0$ for all $f \in \mathbb{C}^4$ such that $\|f\| = 1$, provided $\lambda \leq 1$ and $\mu \leq 1$. This is equivalent to the inequality

$$\langle Af, f \rangle^2 + \langle Bf, f \rangle^2 \leq \|f\|^2 \langle Hf, f \rangle$$

being valid for all $f \in \mathbb{C}^4$. Since the map $f \rightarrow |f|$ reduces the RHS while leaving the LHS unchanged, it is sufficient to treat the case where $0 \leq f_r \in \mathbb{R}$ for all $1 \leq r \leq 4$. We expand f in the form $f := \alpha e_1 + \beta e_2 + \gamma e_3 + \delta e_4$, where $\alpha, \beta, \gamma, \delta \in \mathbb{R}$. Direct computations establish that

$$\langle Hf, f \rangle = 2\beta^2 + 2\gamma^2 + 4\delta^2$$

$$\langle Af, f \rangle = \sqrt{2}\beta(\alpha + \delta)$$

$$\langle Bf, f \rangle = \sqrt{2}\gamma(\alpha - \delta).$$

Our task is therefore to establish that

$$2\beta^2(\alpha + \delta)^2 + 2\gamma^2(\alpha - \delta)^2 \leq (2\beta^2 + 2\gamma^2 + 4\delta^2)(\alpha^2 + \beta^2 + \gamma^2 + \delta^2)$$

for all real variables α, β, γ and δ . Now assume that $\alpha\delta \geq 0$. (The case $\alpha\delta < 0$ may be handled in a similar way.) By expanding and simplifying both sides we see that it is sufficient to prove that

$$\beta^4 + 2(\delta^2 - \alpha\delta)\beta^2 + 2\delta^2(\alpha^2 + \delta^2) \geq 0.$$

for all $\alpha, \beta, \delta \in \mathbb{R}$. Regarding this as a quadratic inequality in β^2 , it is certainly valid if

$$(\delta^2 - \alpha\delta)^2 \leq 2\delta^2(\alpha^2 + \delta^2)$$

and the proof of this is elementary.

Lemma 12. Let $\lambda > 1$ and let τ be the positive solution of $\lambda^2 = 1 + \tau^2$. Then the unit vectors

$$g_{\pm} := \left(\frac{\lambda \pm \tau}{2\lambda}, \frac{1}{2\lambda}, \frac{\lambda \mp \tau}{2\lambda}, \frac{1}{2\lambda} \right)$$

are critical points of the functional \mathcal{E} , with energy

$$\mathcal{E}(g_{\pm}) = 2 - \lambda - \frac{1}{\lambda}. \quad (13)$$

Proof. An application of the symmetry S establishes that it is sufficient to treat the vector $g := g_+$. It is easy to check from their definitions that $\|g\| = 1$, $\langle Ag, g \rangle = \tau/\lambda$ and $\langle Bg, g \rangle = 0$. Putting

$$\begin{aligned} L &:= H - 2\lambda \langle Ag, g \rangle A - 2\mu \langle Bg, g \rangle B \\ &= H - 2\tau A \end{aligned}$$

it follows by a direct calculation that

$$Lg = 2(1 - \lambda)g.$$

Lemma 2 now implies that g is a critical point of \mathcal{E} . Moreover,

$$\begin{aligned} \mathcal{E}(g) &= \langle Lg, g \rangle + \lambda \langle Ag, g \rangle^2 \\ &= 2(1 - \lambda) + \lambda(\tau/\lambda)^2 \\ &= 2 - \lambda - \frac{1}{\lambda}. \end{aligned}$$

If $\lambda \geq \mu > 1$ then there is another pair of critical points, given by

$$h_{\pm} := \left(\frac{1}{2\mu}, \frac{\mu \pm \sigma}{2\mu}, \frac{1}{2\mu}, \frac{\mu \mp \sigma}{2\mu} \right)$$

where $\mu^2 := 1 + \sigma^2$. These have energies

$$\mathcal{E}(h_{\pm}) = 2 - \mu - \frac{1}{\mu}. \tag{14}$$

Numerical explorations give strong grounds for believing that these four points are local minima of \mathcal{E} , that no other local minima exist, and that g_{\pm} are the global minima, but we have not done the calculations needed to prove this.

The two pairs of energies obtained above can be compared with the four smallest eigenvalues of the operator K . We label these $\{E_r\}_{r=1}^4$ in increasing order. Because of the group symmetries present, they can probably all be evaluated in closed form, although they are the eigenvalues of a 16×16 matrix depending upon two parameters. We have however only determined E_2 and E_3 . Representing the eigenvectors $\phi_r \in \mathbb{C}^4 \otimes \mathbb{C}^4$ by 4×4 matrices, the statements of the following lemma may be checked by direct computations.

Lemma 13. If $0 < \mu < \lambda$ and $\alpha := (\sqrt{\lambda^2 + 4} - \lambda)/4$ then

$$\phi_2 := \begin{pmatrix} 1 & \alpha & 0 & \alpha \\ \alpha & 0 & -\alpha & 0 \\ 0 & -\alpha & -1 & -\alpha \\ \alpha & 0 & -\alpha & 0 \end{pmatrix}$$

is an eigenvector of K with eigenvalue

$$E_2 := 2 - \frac{1}{2}\lambda - \frac{1}{2}\sqrt{\lambda^2 + 4}.$$

Moreover, if $\beta := (\sqrt{\mu^2 + 4} - \mu)/4$ then

$$\phi_3 := \begin{pmatrix} 0 & \beta & 0 & -\beta \\ \beta & 1 & \beta & 0 \\ 0 & \beta & 0 & -\beta \\ -\beta & 0 & -\beta & -1 \end{pmatrix}$$

is an eigenvector of K with eigenvalue

$$E_3 := 2 - \frac{1}{2}\mu - \frac{1}{2}\sqrt{\mu^2 + 4}.$$

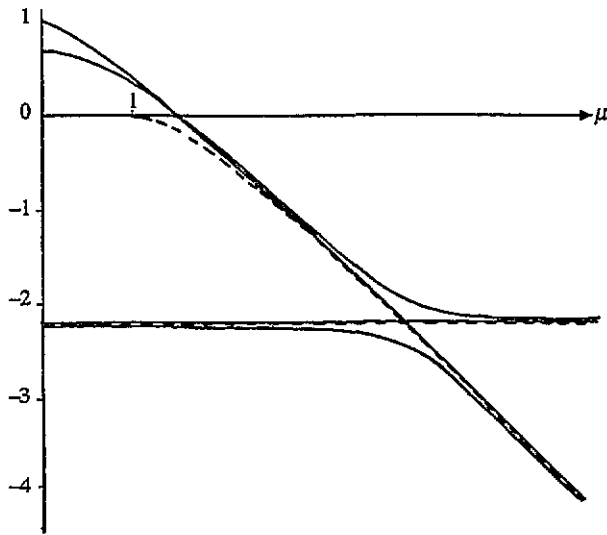


Figure 1.

The following diagram displays the bottom four eigenvalues of K for $\lambda := 4$ and various values of μ by full curves, and the approximate expressions (13) and (14) by broken curves.

We see that the agreement between the true eigenvalues and those computed by the product ansatz is very good, except when λ and μ are very close. Here the quantum tunnelling between the four different wells of the potential $V := -\lambda A \otimes A - \mu B \otimes B$ becomes important. We can, however, carry out a second stage of our procedure, described as use of the generator coordinate method in [W3], in order to improve its accuracy. One of the main uses of the generator coordinate method is to restore symmetries which have been removed by some other procedure, and we shall see that taking an appropriate nonlinear combination of the states g_{\pm} and h_{\pm} enables us to approximate the true ground state even better. We mention that Amann [A2] has recently made similar use of the generator coordinate method for a different model.

To illustrate this, we suppose that $\lambda = \mu > 1$ and consider

$$\bar{f} := g_+ \otimes g_+ + g_- \otimes g_- + h_+ \otimes h_+ + h_- \otimes h_-$$

as an approximation to the true ground state $f := \phi_1$ of K . In our approach this coherent superposition of the product states is unstable with respect to small external perturbations, just as the true ground state is. The simple forms of \bar{f} and f below can be explained by the large symmetry group of the model, but we first observed it in numerical calculations.

Lemma 14. The unnormalized vector \bar{f} is of the form

$$\bar{f} := \begin{pmatrix} \lambda^2 & \lambda & 1 & \lambda \\ \lambda & \lambda^2 & \lambda & 1 \\ 1 & \lambda & \lambda^2 & \lambda \\ \lambda & 1 & \lambda & \lambda^2 \end{pmatrix}$$

and has energy

$$\bar{E} := \frac{\langle K \bar{f}, \bar{f} \rangle}{\langle \bar{f}, \bar{f} \rangle^2} = \frac{2(\lambda - 1)^2 - \lambda(\lambda^2 - 1)}{\lambda^2 + 1}.$$

The true unnormalized ground state f is of the form

$$f := \begin{pmatrix} u^2 & u & 1 & u \\ u & u^2 & u & 1 \\ 1 & u & u^2 & u \\ u & 1 & u & u^2 \end{pmatrix} \quad (15)$$

where $u := \frac{\lambda}{2} + \frac{1}{2}\sqrt{\lambda^2 + 4}$. It has energy

$$E := 2 - \sqrt{\lambda^2 + 4}.$$

Proof. The statements about \tilde{f} are obtained by direct computations. Another direct computation shows that a vector $f \in \mathbf{C}^4 \otimes \mathbf{C}^4$ of the form (15) is an eigenvector of the full Hamiltonian K with the stated energy E if u has the stated value. Now the coefficients of f are all positive, and K is a discrete Schrödinger operator. Therefore f is the ground-state eigenvector of K , by the arguments of [D4, section 1.4].

We can compare the approximate and exact ground-state energies \tilde{E} and E by two methods. For large λ we have the two asymptotic expansions

$$\begin{aligned} \tilde{E} &= -\lambda + 2 - \frac{2}{\lambda} + \frac{3}{\lambda^3} + O\left(\frac{1}{\lambda^4}\right) \\ E &= -\lambda + 2 - \frac{2}{\lambda} + \frac{2}{\lambda^3} + O\left(\frac{1}{\lambda^4}\right) \end{aligned}$$

which correspond extremely closely. However, even for moderate values of λ the two expressions are remarkably similar, as the following table indicates.

5. Discussion

There are several phenomena which go under the general heading of 'the molecular structure problem' in quantum chemistry, and we have chosen to concentrate on one which involves the breaking of a rotational symmetry group. The task is to find a procedure for constructing localized states which break the symmetry group in a manner which corresponds to the intuitions of chemists, and which can be justified in some theoretical manner. We have achieved the first goal by imposing a product state ansatz, and have shown in a soluble

Table 1.

Ground-state energy for $\lambda = \mu$		
λ	Exact	Approx.
1.0	-0.2361	-0.0000
1.2	-0.3324	-0.1836
1.4	-0.4413	-0.3459
1.6	-0.5612	-0.4989
1.8	-0.6907	-0.6491
2.0	-0.8284	-0.8000
3.0	-1.6056	-1.6000
4.0	-2.4721	-2.4706
5.0	-3.3852	-3.3846

model that it has the expected effect. It is our belief that similar behaviour occurs in more realistic situations.

The more difficult task is to justify the product state ansatz, and here we mention three approaches. The first invokes the process of measurement and the associated collapse of the wavepacket. If one is primarily interested in the behaviour of observed molecules, then one can refer to the fact that the act of observation destroys phase correlations. It has been maintained in this context that the act of observation confers objective reality upon the molecules, and that an object/subsystem can only be regarded as having context-free individuality if it is in a pure state [W3].

The second idea is to refer to the decoherence effect of the environment [O, Z], or more specifically to random external influences such as electric or even gravitational fields. It can be shown in certain models for which the energy gap between the smallest two eigenvalues is extremely small [CJ, GM, JMS], that these introduce random phases into the symmetric zero-field eigenstates, and that localized symmetry-breaking states emerge after taking the external fields into account. The instability of the exact solutions of the Schrödinger equation under tiny perturbations has some similarities with well known effects in classical chaotic dynamics. A recent model of the effect of random fields using an evolution equation for mixed states [B] exhibits decoherence on a very short timescale, but also a further mixing on a much retarded timescale associated with the quantum Zeno effect.

The final proposal is that one can consider the tensor product ansatz as a mathematical device for constructing eigenstates in two stages, the second stage being the use of the generator coordinate method. In this view chemists choose to invoke only the first stage of the procedure because it leads to results which have simple classical interpretations, and which are computationally satisfactory for many purposes.

The author does not consider it his role as a mathematician to arbitrate between these justifications. It is sufficient that several different physically based arguments lead to the same conclusion, and that the study of tensor product states is mathematically sufficiently precise for detailed investigation.

The need for a justification of our approach is more acute in the context of equilibrium statistical mechanics, because the conventional approach via Gibbs states and the thermodynamic limit is very well established. The conventional attitude to the tensor product ansatz is presumably that it is an approximation to the correct conventional procedure, and may or may not be a good one, depending upon the model. We argue that this is an oversimplification of a rather complex situation. We will only discuss the zero-temperature problem, that is the properties of ground states.

A conventional infinite volume statistical mechanical model is an idealization of a physical problem, in which interactions with particles such as photons and effects of the environment including gravitation are disregarded. This is achieved by removing entanglements, that is by assuming a tensor product state for the combination of infinite volume model and the environment. This is normally done implicitly, by setting the interactions between the system and environment equal to zero, which ignores the effect of the interaction energy. The conventional procedure therefore has a restricted domain of validity. If the ground state is unique and there is a positive energy gap as in [AKLT, FNW], then the use of the conventional model may be justified. If, however, the conventionally calculated ground state has no energy gap, we would argue that the conventional model has no more justification than the use of a tensor product ansatz which explicitly recognizes the decoherence effects of the environment.

We would emphasize that the use of a tensor product ansatz involves estimating a coherence length associated with the environmental effects. Once this has been fixed, one

groups particles together into clusters, which become the sites of the lattice X of section 2. These clusters should have diameter at least equal to the coherence length, but the model does not become better by allowing the cluster size to diverge to infinity. The introduction of an extra parameter, the coherence length, may or may not be welcomed, but its mathematical effect can only be judged by detailed calculations in particular models.

Acknowledgments

I should like to thank R F Streater for many stimulating conversations both before and during the course of this work. Thanks are also due to R G Woolley for further comments.

References

- [AKLT] Affleck I, Kennedy T, Lieb E H and Tasaki H 1988 Valence bond ground states in isotropic quantum antiferromagnets *Commun. Math. Phys.* **115** 477–528
- [A1] Amann A 1991 Ground states of a spin-boson model *Ann. Phys.* **208** 414–48
- [A2] Amann A 1992 Applying the variational principle to a spin-boson Hamiltonian *J. Chem. Phys.* **96** 1317–24
- [B] Berry M 1994 Two-state quantum asymptotics *Preprint*
- [CJ] Claverie P and Jona-Lasinio G 1986 Instability of tunneling and the concept of molecular structure in quantum mechanics: the case of pyramidal molecules and the enantiomer problem *Phys. Rev. A* **33** 2245–53
- [D1] Davies E B 1979 Symmetry breaking for a non-linear Schrödinger operator *Commun. Math. Phys.* **64** 191–210
- [D2] Davies E B 1980 Metastable states of molecules. *Commun. Math. Phys.* **75** 263–83
- [D3] Davies E B 1981 Symmetry breaking for molecular open systems *Ann. Inst. H. Poincaré A* **35** 149–71
- [D4] Davies E B 1989 *Heat Kernels and Spectral Theory* (Cambridge: Cambridge University Press)
- [FNW] Fannes M, Nachtegale B and Werner R F 1992 Finitely correlated states on quantum spin chains. *Commun. Math. Phys.* **144** 443–90.
- [GM] Grecchi V and Martinez A 1993 Non-linear Stark effect and molecular localization *Preprint*
- [H] Hund F 1927 *Z. Phys.* **43** 805
- [JMS] Jona-Lasinio G, Martinelli F and Scoppola E 1981 New approach to the semiclassical limit of quantum mechanics *Commun. Math. Phys.* **80** 223–45
- [LM] Lattimer R P and Harmony M D 1972 The N–N torsional potential function in methylhydrazine *J. Chem. Soc.* **94** 351–6
- [O] Omnes R 1992 Consistent interpretations of quantum mechanics *Rev. Mod. Phys.* **64** 339
- [P1] Pfeifer P 1980 Chiral molecules, a superselection rule induced by the radiation field *Thesis* ETH, Zurich
- [P2] Pfeifer P 1980 A non-linear Schrödinger operator yielding the shape of molecules by spontaneous symmetry breaking *Classical, Semiclassical and Quantum Mechanical Problems in Mathematics, Chemistry and Physics* ed K Gustafson and W P Reinhardt (New York: Plenum Press)
- [P3] Pfeifer P 1983 Hartree theory for matter in a photon or phonon field *Physica* **6D** 393–403; erratum 1983 *Physica* **9D** 266
- [S] Spohn H 1989 Ground states of the spin-boson Hamiltonian *Commun. Math. Phys.* **123** 277–304.
- [WG] Wightman A S and Glance N 1989 Superselection rules in molecules *Nucl. Phys. B (Proc. Suppl.)* **6** 202–6
- [W1] Woolley R G 1976 Quantum theory and molecular structure *Adv. Phys.* **25** 27–52
- [W2] Woolley R G 1980 Quantum mechanical aspects of the molecular structure hypothesis *Israel J. Chem.* **19** 30–46
- [W3] Woolley R G 1987 The evolution of the molecular conception of matter The molecule in biology, chemistry and physics II *National Meeting Storia e Fondamenti della Chimica (Rome)* pp 35–51
- [Z] Zurek W H 1991 Decoherence and the transition from quantum to classical *Phys. Today* 36–44