BRIEF COMMUNICATION

On the variational principle for the non-linear Schrödinger equation

**Zsuzsanna É. Mihálka[1] · Ádám Margócsy[1] · Ágnes Szabados[2] ·
Péter R. Surján[2]**

Abstract

While variation of the energy functional yields the Schrödinger equation in the usual, linear case, no such statement can be formulated in the general nonlinear situation when the Hamiltonian depends on its eigenvector. In this latter case, as we illustrate by sample numerical calculations, the points of the energy expectation value hypersurface where the eigenvalue equation is satisfied separate from those where the energy is stationary. We show that the variation of the energy at the eigensolution is determined by a generalized Hellmann–Feynman theorem. Functionals, other than the energy, can, however be constructed, that result the nonlinear Schrödinger equation upon setting their variation zero. The second centralized moment of the Hamiltonian is one example.

Keywords Nonlinear Schrödinger equation · Variational principle · Generalized Hellmann–Feynman theorem · Second centralized moment

1 Introduction

Apart from the linear Schrödinger equation written in standard notation

$$(\hat{H} - E\hat{I})\, \Psi \;=\; 0 \tag{1}$$

nonlinear equations are also considered in various areas of physics and chemistry. These can be grouped into two categories:

(1) The eigenproblem is nonlinear in energy.

   (i) An example for such nonlinearity is

$$(\hat{A}_0 - E\hat{A}_1 - E^2\hat{A}_2 - E^3\hat{A}_3 \ldots)\Psi = 0$$

   where $\hat{A}_i$-s are linear operators. Results on problems of this type in the mathematical literature are extensive, including the search of appropriate variational principles [1–3].

   (ii) Another example is the Feshbach-type [4] effective Hamiltonian appearing in Löwdin's energy dependent partitioning technique [5,6]:

$$H^{eff}\,\Psi\,=\,\left[H_{AA} + H_{AB}\left(H_{BB} - E\right)^{-1} H_{BA}\right]\Psi\,=\,E\,\Psi,$$

   where AA, AB, ... indicate blocks (projections) of the Hamiltonian.

We shall not further consider this case in this paper.

(2) The equation is nonlinear in the eigenvector Ψ, which may occur when the Hamiltonian depends on the eigenvector:

$$\left(\hat{H}(\Psi) - E\right)\Psi\,=\,0. \tag{2}$$

This is the case we shall be concerned with here.

Examples for case (2) are

  (i) the nonlinear Schrödinger equation describing the interaction of a molecule with a polarizable medium (setting $\hat{A}$ and $\hat{B}$ appropriately) [7–12]:

$$\left(\hat{H}^0 - \hat{A}\,\frac{\langle\Psi|\hat{B}|\Psi\rangle}{\langle\Psi|\Psi\rangle}\right)\Psi\,=\,E\,\Psi. \tag{3}$$

  (ii) The Gross–Pitajevskij equation [13] (GPE), which is the bosonic analogue of the Hartree–Fock equations:

$$(\hat{T} + \hat{V} + \lambda||\psi||^2)\psi = E\psi.$$

Here $\hat{T}$ is the kinetic energy operator, $\hat{V}$ is an external potential, and λ is a coupling constant for inter-particle interactions. Effects of nonlinearity and variational principles for GPE and its generalizations have been widely studied [14–18].

(iii) The Hartree or the Hartree–Fock equations for many-electron systems. The latter in canonical form are written as

$$\hat{F}\psi_i = \varepsilon_i\psi_i \tag{4}$$

where the nonlinearity arises from the dependence of the Fockian $\hat{F}$ on the density matrix P constructed with occupied one-electron orbitals $\{\psi_k\}$ in basis set $\{\mu\}$ as:

$$P_{\mu v} = \sum_{k}^{occ} \langle \mu | \psi_k \rangle \langle \psi_k | v \rangle. \tag{5}$$

(iv) Certain nonlinear differential equations for Hamiltonian systems have the form of nonlinear eigenvalue problems. These can originate from special variational principles [19–21].

(v) The non-perturbative form of the Bloch equation [22–24]

$$\hat{H} \, \hat{\Omega} \; = \; \hat{\Omega} \, \hat{H} \, \hat{\Omega} \tag{6}$$

where the wave operator is either of form $\hat{\Omega} = |\Psi\rangle\langle\Psi|$ or $\hat{\Omega} = |\Psi\rangle\langle\varphi|$ where φ is a reference function, constitute another example for wave function nonlinearity. In fact, Eq. (6) was named as the nonlinear Schrödinger equation by Löwdin [23].

In this paper, we shall treat cases (i) and (iii) discussing variational principles.

When relating the ordinary linear Schrödinger equation (1) to the variation of the energy (preserving normalization of the eigenvector by a Lagrangian multiplier E), one proceeds as

$$\delta \, [\langle \Psi | H | \Psi \rangle + E(1 - \langle \Psi | \Psi \rangle)] \; = \; 0$$

resulting

$$\langle \delta \Psi | H - E | \Psi \rangle \; + \text{c.c.} \; = \; 0,$$

which for arbitrary variations $\langle \delta \Psi |$ is satisfied only when $H|\Psi\rangle = E|\Psi\rangle$.

In this procedure we have utilized that $\delta H = 0$.

2 The nonlinear case

2.1 Variation of the energy expectation value

When treating the nonlinear Schrödinger equation (2), one must not assume that $\delta H = 0$. One should rather write the variation of the energy expectation value $\mathcal{E}$ with the appropriate Lagrange multiplier E as

$$\delta\mathcal{E} = \delta \left(\langle \Psi | \hat{H} - E | \Psi \rangle \right),$$

where the constant term E whose variation is zero has been omitted. Carrying out the variation we have

$$\delta\mathcal{E} = \left(\langle \delta\Psi | \hat{H} - E | \Psi \rangle + c.c. \right) + \langle \Psi | \delta\hat{H} | \Psi \rangle. \tag{7}$$

This shows that satisfying the Schrödinger equation is insufficient to make the variation of the energy expectation value zero. On the contrary, imposing the variational condition $\delta\mathcal{E} = 0$, the Schrödinger equation will not be satisfied.

2.2 The generalized Hellmann–Feynman theorem

If Ψ is chosen to satisfy the Schrödinger equation, the first term in Eq. (7) as well as its complex conjugate vanish, and one is left with

$$\delta\mathcal{E} = \langle\Psi|\delta\hat{H}|\Psi\rangle. \tag{8}$$

This equation gives us the variation of the energy expectation value when the wave function is an eigenfunction of the nonlinear Hamiltonian $H(\Psi)$. The formula is in analogy with the Hellmann–Feynman equation [25], which gives the derivative of the energy with respect to an external parameter R of the potential:

$$\frac{\partial E}{\partial R} = \left\langle\Psi|\frac{\partial\hat{H}}{\partial R}|\Psi\right\rangle.$$

It seems that the similar equation (8) is valid for the variation of the energy expectation value in the nonlinear Schrödinger equation, too. One may call Eq. (8) the generalized Hellmann–Feynman theorem.

2.3 The Hartree–Fock case

The general formulation of the previous sections can be applied to the variation of Fockian expectation values in a straightforward manner. We write the orbital energies using standard notation as

$$\varepsilon_i = \langle\psi_i|\hat{F}|\psi_i\rangle.$$

The variation of ε_i with respect to orbital rotations is

$$\delta\varepsilon_i = \left(\langle\delta\psi_i|\hat{F}|\psi_i\rangle + c.c.\right) + \langle\psi_i|\delta\hat{F}|\psi_i\rangle. \tag{9}$$

The orbital variations can be described by a unitary matrix (this will take care of preserving the orthonormality of orbitals):

$$\psi_i + \delta\psi_i = \sum_k \left[e^{\epsilon}\right]_{ik}\psi_k = \sum_k \left[\delta_{ik} + \epsilon_{ik} + \mathcal{O}(2)\right]\psi_k,$$

where the exponential parametrization of a unitary matrix is used with antihermitean matrix ϵ. Linear variation of orbital ψ_i is, accordingly

$$\delta\psi_i = \sum_k \epsilon_{ik}\,\psi_k = \sum_k^{occ} \epsilon_{ik}\,\psi_k + \sum_k^{virt} \epsilon_{ik}\,\psi_k. \tag{10}$$

To keep the variation orthogonal to the varied orbital, one may want to include the restriction $i \neq k$ in the occupied sum. This is taken care by ϵ_{ii} being 0 due to the antihermiticity of ϵ. It follows that we consider variations with the property

$$\langle\psi_i|\delta\psi_i\rangle = 0.$$

This property ensures that quantities of type

$$\langle\delta\psi_i|\hat{F}|\psi_i\rangle = \langle\delta\psi_i|\psi_i\rangle\varepsilon_i = 0,$$

since we take the variation at the canonical orbitals ψ_i which are eigenvectors of $\hat{F}$. As a consequence, the first term and its complex conjugate in Eq. (9) disappear and one is left with

$$\delta\varepsilon_i = \langle\psi_i|\delta\hat{F}|\psi_i\rangle, \tag{11}$$

which is the manifestation of the generalized Hellmann–Feynman theorem discussed above in the Hartree–Fock case.

To evaluate $\delta\hat{F}$, we write the diagonal elements of $\hat{F}$ as

$$\varepsilon_i = h_{ii} + \sum_k^{occ} \big(2(\psi_i\psi_i|\psi_k\psi_k) - (\psi_i\psi_k|\psi_k\psi_i)\big) = h_{ii} + 2J_i - K_i$$

with standard notation. When taking the variation of this, we utilize Eq. (11). This means that the terms involving the variations of ψ_i drop out, and one should consider only the variations with respect to ψ_k in the summation of the Coulomb and exchange operators. This will be expressed by the variation symbol δ_i. One gets:

$$\delta\varepsilon_i = 2\,\delta_i J_i - \delta_i K_i = \sum_k^{occ} \big(2\,(\psi_i\,\psi_i|\delta\psi_k\,\psi_k) - (\psi_i\,\psi_k|\delta\psi_k\,\psi_i)\big) + c.c.$$

The next step is to expand the orbital variations according to Eq. (10). Writing this down for the Coulomb term J_i first:

$$\delta_i J_i = \sum_k^{occ}\sum_j^{occ} \big(\epsilon_{kj}(\psi_i\psi_i|\psi_j\psi_k) + \epsilon_{kj}^*(\psi_i\psi_i|\psi_k\psi_j)\big)$$

$$+ \sum_k^{occ} \sum_j^{virt} \left(\epsilon_{kj}(\psi_i\psi_i|\psi_j\psi_k) + \epsilon_{kj}^*(\psi_i\psi_i|\psi_k\psi_j) \right)$$

$$= \sum_k^{occ} \sum_j^{occ} \left(\epsilon_{kj}(\psi_i\psi_i|\psi_j\psi_k) + \underbrace{\epsilon_{jk}^*}_{-\epsilon_{kj}}(\psi_i\psi_i|\psi_j\psi_k) \right)$$

$$+ \sum_k^{occ} \sum_j^{virt} \left(\epsilon_{kj}(\psi_i\psi_i|\psi_j\psi_k) + \epsilon_{kj}^*(\psi_i\psi_i|\psi_k\psi_j) \right)$$

$$= \sum_k^{occ} \sum_j^{virt} \left(\epsilon_{kj}(\psi_i\psi_i|\psi_j\psi_k) + \epsilon_{kj}^*(\psi_i\psi_i|\psi_k\psi_j) \right)$$

with the asterisk indicating complex conjugation. In the last-but-one step we applied an interchange of labels j, k, and realized that the two terms involving summation over occupied orbitals ψ_k and ψ_j cancel due to the antihermiticity of ϵ. The two terms containing virtual summation survive, since there j and k are not equivalent indices.

The same manipulation can be repeated for the exchange terms to get

$$\delta_i K_i = \sum_k^{occ} \sum_j^{virt} \left(\epsilon_{kj}(\psi_i\psi_k|\psi_j\psi_i) + \epsilon_{kj}^*(\psi_i\psi_j|\psi_k\psi_i) \right)$$

The quantity $2\,\delta_i J_i - \delta_i K_i$ constitutes the internal variation of the Fockian, which determines orbital energy variations through (11).

The fact that orbital variations within the occupied space do not contribute to the variation of orbital energies is trivial, since $\delta\hat{F}$ comes from the variation of the density matrix P which is invariant to any mixing within the occupied space.

2.4 Numerical illustrations

We include two simple illustrations. The first one is a two-state model for equations of type (3). We set $A = B$ and aim to solve the nonlinear equation
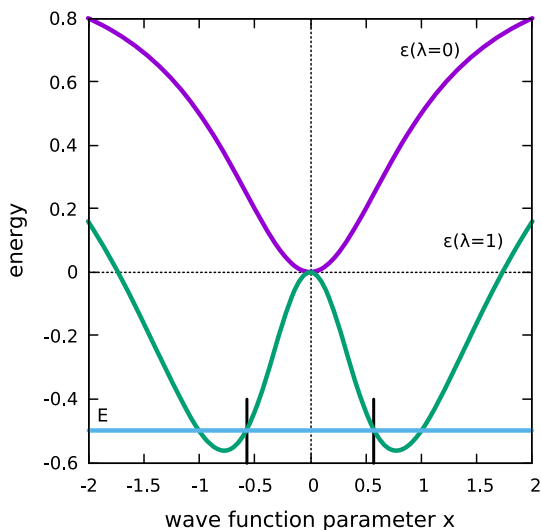
$$H^0\,\Psi - \lambda\,\frac{\langle\Psi|A|\Psi\rangle}{\langle\Psi|\Psi\rangle}\,A\,\Psi = E\,\Psi \qquad (12)$$

where the strength of nonlinearity is characterized by λ. Working in the basis set of the eigenvectors of H^0 in the two-state model, choosing the energy origin at H_{11}^0 and the energy unit at H_{22}^0, then

$$H^0 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Further, assuming only off-diagonal perturbation in A,

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Finally, writing the eigenvector as

$$\Psi = \begin{pmatrix} 1 \\ x \end{pmatrix}$$

the normalization $\langle \Psi | \Psi \rangle = 1 + x^2$ resulted.

This simple model can be solved analytically by the substitution of these choices into Eq. (12) yielding, apart from a trivial solution $x = 0$, the self-consistent solutions:

$$x^2 = \frac{2\lambda - 1}{2\lambda + 1}$$

as a function of the perturbation strength λ. Choosing $\lambda = 1$ we get the two solutions $x = \pm 1/\sqrt{3}$. The energy expectation value as a function of x is plotted in Fig. 1. The non-interacting case, when the Schrödinger equation is linear, is $\lambda = 0$. Then, trivially, the expectation value has a stationary point (minimum) at $x = 0$, the zero order solution. The non-linear case is illustrated by choosing $\lambda = 1$. Then the expectation value is not stationary for the self-consistent wave function $x = \pm 1/\sqrt{3}$, and when $\mathcal{E}(x)$ is stationary (at around $x = 0.775$), x does not correspond to any solution of the nonlinear problem.

The second example we consider are Fockian expectation value variations for a model system consisting of one helium atom and two hydrogens. Pairwise orbital rotations are performed with the unitary matrix

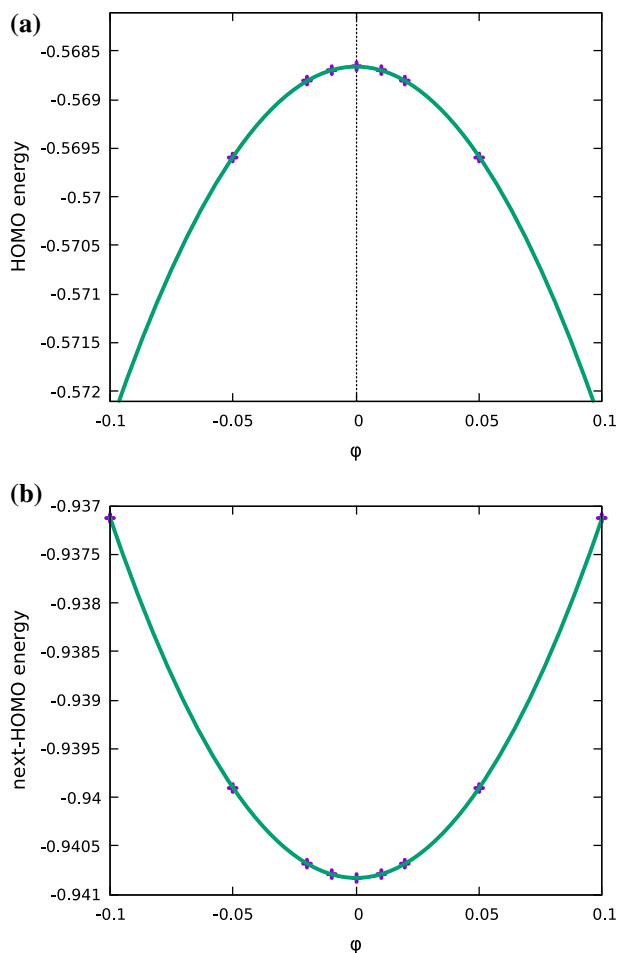$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

(a)

(b)

Fig. 2 Variation of Fockian expectation values for the model system HeH$_2$ in a DZ basis set with respect to occupied orbital mixing angle φ. Both the HOMO (Fig. 2a.) and the next-HOMO (Fig. 2b.) energies are stationary for canonical orbitals $\varphi = 0$

Figure 2. shows the trivial situation when only the two occupied orbitals are mixed. As a function of the mixing angle φ, the energies of the HOMO (Fig. 2a.) and the next-HOMO (Fig 2b.) vary as expected, showing extremal values at $\varphi = 0$ corresponding to the canonical case, the former showing a maximum, the latter a minimum.

Figure 3. shows that when occupied and virtual orbitals are allowed to mix, the stationary points of orbital energies deviate from the canonical $\varphi = 0$ value. The deviations are not too large, however.
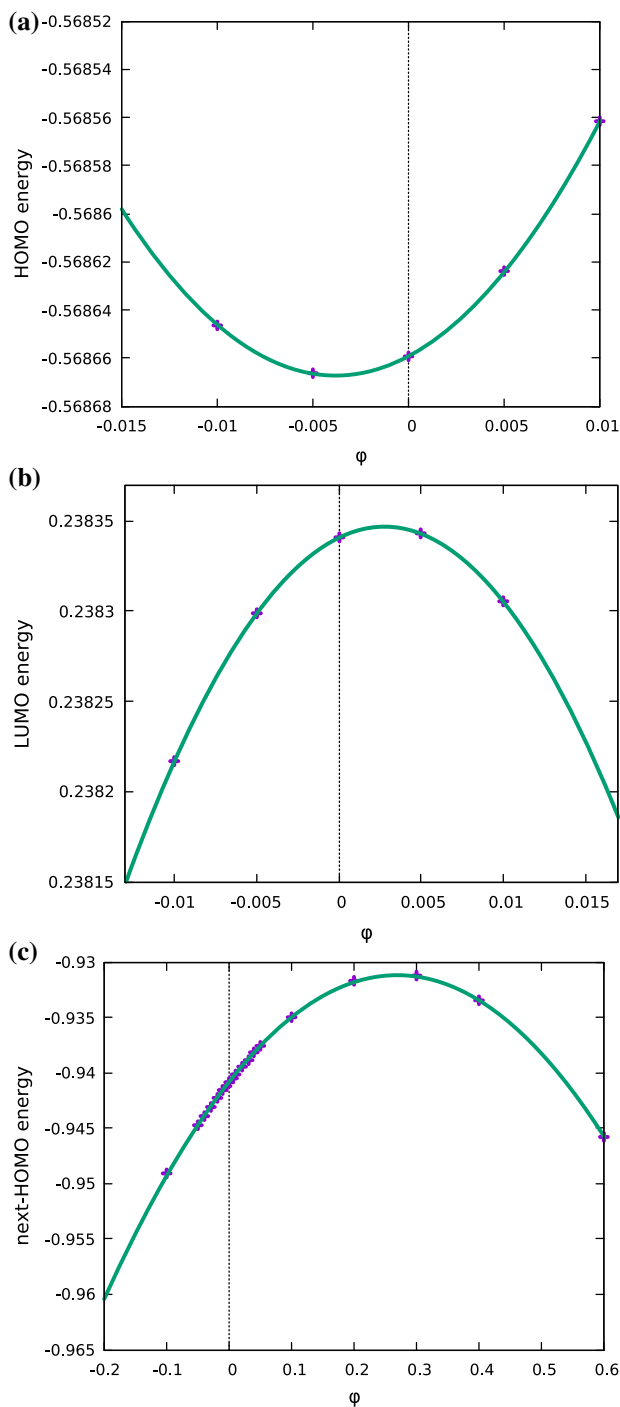
Fig. 3 Variation of Fockian expectation values for the model system HeH$_2$ with respect to HOMO-LUMO mixing described by rotation angle φ in a DZ basis set. The orbital energy curves for the HOMO (Fig. 3a.), the LUMO (Fig. 3b.) and the next-HOMO (Fig. 3c.) are drawn

3 A simple variational principle for nonlinear Schrödinger equations

We note first that for the linear, but time-dependent Schrödinger equation the original Dirac–Frenkel variational principle [26,27] has been reformulated by McLachlan and Ball [28], and discussed thoroughly by Löwdin and Mukherjee [29]. The essence of the reformulation was the utilization of the second centralized moment of the Hamiltonian

$$m_2 = \left\langle \Psi | \left(\hat{H} - \langle \hat{H} \rangle \right)^2 | \Psi \right\rangle \tag{13}$$

with $\langle \cdots \rangle$ indicating Rayleigh expectation value $\langle \Psi | \dots | \Psi \rangle / \langle \Psi | \Psi \rangle$.

In this section we emphasize that the second moment may also serve as a variational functional for the time-independent nonlinear Schrödinger equation.

An alternative form of m_2 can be written as the square norm

$$m_2 = ||r||^2 = \langle r | r \rangle$$

whith the residual vector

$$|r\rangle = \left(\hat{H} - \langle \hat{H} \rangle \right) | \Psi \rangle.$$

It is evident that if the nonlinear Schrödinger equation (2) is satisfied, $|r\rangle = 0$, thus $m_2 = 0$, too. A formal proof that vanishing m_2 is equivalent to the linear Schrödinger equation was given by Nakatsuji and Davidson [30]. We repeat their arguments here with the necessary modification requested by nonlinearity to show that the statement remains valid also in the nonlinear case.

All examples studied here assume that the nonlinear Hamiltonian entering Eq. (2) is Hermitean. Therefore at some wave function Ψ we may consider the spectral resolution of $H(\Psi)$ as

$$\hat{H}(\Psi) = \sum_k E_k(\Psi) |\Psi_k\rangle\langle\Psi_k|$$

leading to

$$\left(\hat{H} - \langle \hat{H} \rangle \right)^2 = \sum_k \left(E_k - \langle \hat{H} \rangle \right)^2 |\Psi_k\rangle\langle\Psi_k|.$$

One must emphasize that the expansion vectors Ψ_k also depend on Ψ, which is not indicated in the equations to avoid a cumbersome notation.

Wave function Ψ entering Eq. (13) can be expanded as

$$\Psi = \sum_k C_k |\Psi_k\rangle,$$

leading to

$$m_2 = \sum_k |C_k|^2 \left(E_k - \langle \hat{H} \rangle \right)^2.$$

Since every term is nonnegative, setting $m_2 = 0$ involves that each term in the summation over k must be zero individually. This involves (since $E_k \neq \langle \hat{H} \rangle$ in general) that $C_k = 0$ for most k. To avoid a trivial solution $\Psi = 0$, there must be at least one value of k, say i, for which

$$E_i = \langle \hat{H} \rangle$$

with the associated coefficient $C_i \neq 0$. This means that Ψ must be a pure state $\Psi = \Psi_i$, which is then a self-consistent solution of the non-linear Schrödinger equation. In case of $\langle \hat{H} \rangle$ being equal to a degenerate level, Ψ becomes a linear combination of the degenerate states.

Based on this conclusion, one may write down the variational principle

$$\delta \left\langle \Phi \left| \left(\hat{H} - \frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle} \right)^2 \right| \Phi \right\rangle = 0 \tag{14}$$

where Φ is a (class of) trial function(s). One may distinguish two cases:

(a) Φ is parametrized sufficiently well so that the family of functions it represents contains one or more true self-consistent eigenvector of the nonlinear Eq. (2). In this case the absolute minimum of the non-negative functional is 0, i.e., the second centralized moment vanishes and the optimal Φ becomes an exact eigenstate.
(b) Φ is less flexible than in case (a), so that solving the variational problem (14) does not yield $m_2 = 0$. In this case the variation of Φ provides the best possible approximation within the permitted function class, resulting the smallest possible positive value for m_2. The optimal Φ serves as an approximate eigenfunction.

It is worth mentioning that the above procedure may be, in principle applied to any self-consistent eigenstate of a nonlinear Hamiltonian, though optimization for excited states may not be easy. Numerically stable procedures for determining excited states variationally, that make use of second centralized moments, were proposed recently [31,32] for linear Hamiltonians.

References

1. R.E.L. Turner, J. Math. Anal. Appl **17**, 151 (1967)
2. R.E.L. Turner, J. Funct. Anal. **7**, 297 (1968)
3. H. Voss, Lecture Notes in Computational Science **103**, 305 (2014)
4. H. Feshbach, Ann. Phys. (N.Y.) **19**, 287 (1962)
5. P.-O. Löwdin, J. Mol. Spectrosc. **10**, 12 (1963)
6. P.-O. Löwdin, J. Math. Phys. **3**, 969 (1962)
7. G. Klopman, Chem. Phys. Lett. **1**, 200 (1967)
8. J.L. Rivail, D. Rinaldi, Chem. Phys. **18**, 233 (1976)
9. J.G. Ángyán, M. Allavena, M. Picard, A. Potier, O. Tapia, J. Chem. Phys. **77**, 4723 (1982)
10. P.R. Surján, J.G. Ángyán, Phys. Rev. A **28**, 45 (1983)
11. J.G. Ángyán, P.R. Surján, Phys. Rev. A **44**, 2188 (1991)
12. P.R. Surján, J. Math. Chem. **8**, 151 (1991)
13. F. Dalfovo, S. Giorgini, L. Pitaevskii, S. Stringari, Rev. Mod. Phys. **71**, 463 (1999)
14. J.C.D. Perez, C. Trallero-Giener, V. Lopez-Richard, C. Trallero-Herrero, J.L. Birman, Pys. Stat. Sol (c) **2**, 3665 (2005)
15. C. Trallero-Giener, J.C.D. Perez, V. Lopez-Richard, J.L. Birman, Physica D **237**, 2342 (2008)
16. J.-H. He, Chaos. Solitons Fractals **19**, 847 (2004)
17. L. Xu, Phys. Lett. A **359**, 627 (2006)
18. X.-W. Zhou, L. Wang, Comput. Math. Appl. **61**, 2035 (2011)
19. P. Antonelli, J. Arbunich, C. Sparber, SIAM J. Math. Anal. **51**, 110 (2019)
20. A. Eydeland, J. Spruck, B. Turkington, Math. Comput. **55**, 509 (1990)
21. R.D. Benguria, M.C. Depassier, Phys. Rev. Lett. **77**, 2847 (1996)
22. I. Lindgren, J. Morrison, *Atomic Many-Body Theory* (Springer, Berlin, 1986)
23. P.O. Löwdin, Int. J. Quant. Chem. **72**, 379 (1999)
24. P. Szakács, P.R. Surján, J. Math. Chem. **43**, 314 (2008)
25. I. Mayer, *Simple Theorems, Proofs, and Derivations in Quantum Chemistry* (Kluwer, New York, 2003)
26. P.A.M. Dirac, Proc. R. Soc. A **112**, 673 (1926)
27. J. Frenkel, *Wave Machanics, Andvanced General Theory* (Dover, New York, 1950)
28. A.D. McLahlan, M.A. Ball, Rev. Mod. Phys **36**, 844 (1964)
29. P.-O. Löwdin, P.K. Mukherjee, Chem. Phys. Lett. **14**, 1 (1972)
30. H. Hakatsuji, E.R. Davidson, J. Chem. Phys. **115**, 2000 (2001)
31. L. Zhao, E. Neuscamman, J. Chem. Theor. Comput. **12**, 3436 (2016)
32. H.-Z. Ye, T. Van-Voorhis, J. Chem. Theor. Comput. **15**, 2954 (2019)