

Generalized variational principle for excited states using nodes of trial functions

Dario Bressanini*

Dipartimento di Scienze Chimiche e Ambientali, Università dell'Insubria, Via Lucini 3, I-22100 Como, Italy

Peter J. Reynolds†

Physics Division, Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709, USA

(Received 14 June 2011; revised manuscript received 1 August 2011; published 18 October 2011)

The familiar variational principle provides an upper bound to the ground-state energy of a given Hamiltonian. This allows one to optimize a trial wave function by minimizing the expectation value of the energy. This approach is also trivially generalized to excited states, so that given a trial wave function of a certain symmetry, one can compute an upper bound to the lowest-energy level of that symmetry. In order to generalize further and build an upper bound of an arbitrary excited state of the desired symmetry, a linear combination of basis functions is generally used to generate an orthogonal set of trial functions, all bounding their respective states. However, sometimes a compact wave-function form is sought, and a basis-set expansion is not desirable or possible. Here we present an alternative generalization of the variational principle to excited states that does not require explicit orthogonalization to lower-energy states. It is valid for one-dimensional systems and, with additional information, to at least some n -dimensional systems. This generalized variational principle exploits information about the nodal structure of the trial wave function, giving an upper bound to the exact energy without the need to build a linear combination of basis functions. To illustrate the theorem we apply it to a nontrivial example: the $1s2s\ ^1S$ excited state of the helium atom.

DOI: [10.1103/PhysRevE.84.046705](https://doi.org/10.1103/PhysRevE.84.046705)

PACS number(s): 02.70.Ss, 03.65.Ge

I. INTRODUCTION

One of the most powerful methods for obtaining highly accurate approximate solutions to the Schrödinger equation is based on the variational principle

$$\langle \hat{H} \rangle = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0, \quad (1)$$

where \hat{H} is the Hamiltonian operator of the system we wish to study, E_0 is the lowest-energy eigenvalue, and Ψ is any trial wave function of the correct symmetry. In its simplest form, the variational method consists of choosing a physically motivated trial wave function that depends on one or more adjustable parameters and then minimizing the expectation value in Eq. (1) by varying the parameters in the trial function. Within the flexibility of the trial function, this minimum is the best upper bound to the unknown ground-state energy.

The above method can also be applied readily to optimize the lowest eigenstate of a given symmetry simply by constraining Ψ to be of the desired symmetry. However, if the state of interest is an excited state having the same symmetry as a lower state, there is no longer a guarantee that the expectation value of the Hamiltonian will be an upper bound to the exact eigenvalue. In this case one must resort to what is sometimes called the linear variational principle (LVP). Implementation of the LVP requires that a basis set of N independent functions φ_i is chosen, and N trial wave functions are constructed from a linear combination of these basis functions:

$$\Psi_j = \sum_{i=1}^N c_{ij} \varphi_i. \quad (2)$$

The coefficients c_{ij} can then be determined by solving the secular linear equations

$$\sum_{i=1}^N c_{ij} (\langle \varphi_j | \hat{H} | \varphi_i \rangle - \varepsilon_j \langle \varphi_j | \varphi_i \rangle) = 0, \quad j = 1, 2, \dots, N. \quad (3)$$

Nontrivial solutions to Eq. (3) exist if and only if the secular determinant vanishes, namely,

$$\det |\mathbf{H} - \varepsilon \mathbf{S}| = 0, \quad (4)$$

where \mathbf{H} is the Hamiltonian matrix with elements $H_{ij} = \langle \varphi_i | \hat{H} | \varphi_j \rangle$ and \mathbf{S} is the overlap matrix with elements $S_{ij} = \langle \varphi_i | \varphi_j \rangle$.

According to the MacDonald-Hylleraas-Undheim theorem [1,2], the resulting ordered set of N roots ε_j of the secular determinant are upper bounds to the corresponding lowest- N eigenvalues of the Hamiltonian. Furthermore, these approximate eigenvalues approach the exact eigenvalues monotonically as the basis set is increased. In this way, using a collection of basis functions, it is possible to build an upper bound to the desired excited state. What has been lost is the ability to construct a simple trial wave function with a simple physical interpretation.

It would be desirable to have a generalized variational principle with the property that, given an arbitrary trial wave function approximating the desired state, it could be possible to extract an upper bound to the exact energy of the state without the necessity of resorting to the above basis-set expansion method. It is the purpose of this paper to show such a generalization. We prove the generalization for one-dimensional problems by taking advantage of information on the nodal structure of wave functions. Furthermore, we show how in some cases this can be generalized to

*dario.bressanini@uninsubria.it

†peter.reynolds@us.army.mil

higher-dimensional systems if additional information on the nodal structure of the specific system under study is available.

II. THEORY

Consider the spectrum of a one-dimensional (1D) Hamiltonian

$$\hat{H}\phi_n = E_n\phi_n, \quad n = 0, 1, 2, \dots \quad (5)$$

It is well known [3] that for 1D Hamiltonians the so-called nodal theorem holds, i.e., the ground state has no nodes and a nondegenerate M th excited state has exactly M nodes, dividing the real axis into $M + 1$ regions Ω_i , which are usually called *nodal regions* or *nodal domains*.

Consider now a trial function Ψ approximating the exact M th excited state. This trial function must have M nodes, but it need not be a linear combination of basis functions. Let us introduce the expectation values of the Hamiltonian within each nodal region Ω_i ,

$$\langle \hat{H} \rangle_{\Omega_i} = E_{\Omega_i} = \frac{\langle \Psi | \hat{H} | \Psi \rangle_{\Omega_i}}{\langle \Psi | \Psi \rangle_{\Omega_i}} \equiv \frac{H_{\Omega_i}}{S_{\Omega_i}}. \quad (6)$$

The subscript indicates that the integration has been performed only over the region Ω_i . It is easy to see that the full expectation value of the Hamiltonian over the whole space is a weighted average of the expectation values of the different nodal regions

$$\langle \hat{H} \rangle = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_i^{M+1} \langle \Psi | \hat{H} | \Psi \rangle_{\Omega_i}}{\sum_i^{M+1} \langle \Psi | \Psi \rangle_{\Omega_i}} = \frac{\sum_i^{M+1} S_{\Omega_i} E_{\Omega_i}}{\sum_i^{M+1} S_{\Omega_i}}. \quad (7)$$

For a generic excited state, the expectation value [Eq. (7)] can be above or below the exact eigenvalue. We now will prove the following theorem, which is a generalization of the ground-state variational principle in one dimension.

Generalized variational principle. Given a 1D trial wave function Ψ with M nodes, the maximum of the energies $\langle \hat{H} \rangle_{\Omega_i}$ of the nodal regions is an upper bound to the energy of the M th excited state.

To prove this result we construct a basis set of $M + 1$ auxiliary functions f_i ,

$$f_i(x) = \begin{cases} \Psi(x), & x \in \Omega_i \\ 0, & x \notin \Omega_i, \end{cases} \quad (8)$$

which are extracted piecewise out of the original trial wave function. Each f_i is equal to the original wave function within the nodal region Ω_i and zero outside.

Since the nodal regions do not overlap it is easy to show that

$$\langle f_i | f_j \rangle = 0, \quad \langle f_i | f_i \rangle = \langle \Psi | \Psi \rangle_{\Omega_i} = S_{\Omega_i} \quad (9)$$

and

$$\langle f_i | \hat{H} | f_j \rangle = 0, \quad \langle f_i | \hat{H} | f_i \rangle = \langle \Psi | \hat{H} | \Psi \rangle_{\Omega_i} = H_{\Omega_i}. \quad (10)$$

The latter integrals involving the Hamiltonian must be treated with care since there is a derivative discontinuity of the wave function at the boundaries of the nodal regions and this produces a δ function. However, since the functions f_i are zero at their boundaries, the contribution to these integrals is zero. More rigorously, we can take the limit of the integrals,

moving the boundaries from inside the nodal region up to the nodes, and one indeed gets zero contribution.

In order to prove the theorem, let us now build a trial function with the basis set just constructed and call it Ψ_T to distinguish it from the original:

$$\Psi_T = \sum_{i=1}^{M+1} c_i f_i. \quad (11)$$

Of course, if we take all the coefficients to be unity, we recover the original trial function. However, we can instead return to the formalism of Eq. (3) and optimize the linear coefficients c_i by solving the secular determinant of Eq. (4). Given the orthogonality of the above basis functions,

$$\begin{vmatrix} H_{\Omega_1} - \varepsilon S_{\Omega_1} & 0 & \dots & 0 \\ 0 & H_{\Omega_2} - \varepsilon S_{\Omega_2} & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & H_{\Omega_{M+1}} - \varepsilon S_{\Omega_{M+1}} \end{vmatrix} = 0. \quad (12)$$

The solutions of this equation are trivially

$$\varepsilon_i = \frac{H_{\Omega_i}}{S_{\Omega_i}} = \langle H \rangle_{\Omega_i}. \quad (13)$$

In other words, the eigenvalues are equal to the expectation values of the Hamiltonian within each nodal region. By the MacDonald-Hylleraas-Undheim theorem, the ordered set of these values are the upper bounds of the corresponding eigenvalues of the Hamiltonian. Since we are interested in the M th state, this means that the largest eigenvalue is an upper bound¹ to the M th state, i.e.,

$$\max_i \langle H \rangle_{\Omega_i} \geq E_M. \quad (14)$$

This completes the proof. Note that if we start with a positive-everywhere wave function, we recover the usual ground-state variational principle.

To apply the generalized variational principle, we need a method to specify the nodes, or otherwise be able to distinguish different nodal regions, in order to compute the expectation value of the Hamiltonian in each nodal region. For one-dimensional systems this is usually easy.

III. METHOD

To illustrate the theorem, let us consider the $2s$ state of the hydrogen atom. Note that, for a spherically symmetric state, after the integration of the angular coordinates the three-dimensional problem reduces to a one-dimensional one. The trial wave function $\Psi = (1 - ar)e^{-br}$ has two nodal domains, as needed to get the first excited state. We note that this form has sufficient flexibility to correctly describe, for different values of the parameters a and b , both the $1s$ and $2s$ states. However, the standard variational principle cannot

¹It also means that the i th eigenvalue is an upper bound to the i th state. However, these lower eigenstates will in general be poor bounds. Better ones can be constructed as described here, using the proper number of nodes for the state of interest.

guarantee that the expectation value of the energy of the above function is an upper bound to the $2s$ energy, only that it is an upper bound to the $1s$ energy. This failure of the standard variational principle is because the $2s$ state has the same symmetry as the ground state. Since our trial function has a spherical node with radius $r = 1/a$, it is easy to analytically compute the two nodal expectation values, corresponding to the two nodal regions: $0 < r < 1/a$ and $r > 1/a$. By our theorem, the larger of the two must be an upper bound to the exact energy of the $2s$ state, namely, -0.125 a.u. This is indeed easy to show with some algebra. It produces the exact energy when $b = 1/2$ and $a = 1/2$, where the two regions produce the same result.

For higher-dimensional systems, the nodal theorem on which our proof relies is no longer valid. Courant [3] was able to prove a weaker version: The nodes of the M th excited state divide the space into at most $M + 1$ nodal domains. This immediately shows that generalizing our theorem to higher-dimensional systems is not straightforward. The M th excited state does not in general have $M + 1$ nodal regions, which means that we might not have $M + 1$ basis functions as needed to apply the MacDonald-Hylleraas-Undheim theorem. However, we will show that if additional information about the nodal structure is available, the theorem can still be applied.

To see how this can be done, let us consider a nontrivial example that cannot be reduced to a one-dimensional problem, namely, the $1s2s^1S$ excited state of the helium atom. As in our earlier example, being an excited state of the same symmetry as the ground state means that the usual variational principle is not able to give an upper bound. Nevertheless, we can successfully exploit our nodal variational principle here. We choose to describe the $1s2s$ state with the physically descriptive and compact trial wave function

$$\Psi = (r_1^2 + r_2^2 - k^2)(e^{-2r_1 - br_2} + e^{-2r_2 - br_1}) \exp\left(\frac{r_{12}/2}{1 + dr_{12}}\right). \quad (15)$$

This choice is motivated by the success of the hyperspherical treatment of the helium atom. However, other simple forms may be chosen as well.

The exact shape of the node of the $1s2s^1S$ eigenstate is unknown [4]. However, the trial function above has a single nodal hypersphere of hyperradius $R^2 = r_1^2 + r_2^2 = k^2$ dividing the space into two nodal regions. Note that the trial wave function [Eq. (15)] was chosen to satisfy all cusp conditions and that the parameters b and d have no impact on the node. The position of the node is determined entirely by the choice of the constant k . Note that formally [see Eq. (8)] we will be using two trial wave functions built out of the Ψ in Eq. (15): Ψ_{outer} , considering only the region where $R > k$, and similarly Ψ_{inner} , where $R < k$.

As indicated, in order to exploit our theorem, additional knowledge about the nodal structure is necessary. Specifically, we exploit the knowledge that we need to build an upper bound to an excited state of 1S symmetry and that it is known to have a single nodal surface [4]. This requires that the trial functions in *all* nodal regions separately must be orthogonal to lower-energy states of different symmetry. In this case it means that it must be orthogonal to the $1s2s^3S$ state. Since both Ψ_{outer}

TABLE I. VMC expectation values of the energy, in hartrees, for different values of the parameter k , over the entire space (labeled full) and separately in the two nodal regions (labeled A and B). Error bars on the last digit are shown in parentheses. The maximum of the latter two columns is shown in bold.

k	VMC (full)	VMC outer (A)	VMC inner (B)
2.8	-2.7400(2)	-1.8652(2)	-2.807 16(8)
2.6	-2.6236(3)	-1.9197(1)	-2.763 89(8)
2.4	-2.4276(3)	-1.9813(1)	-2.6951(1)
2.3	-2.3420(3)	-1.9882(1)	-2.6469(1)
2.2	-2.2417(3)	-2.0175(1)	-2.5858(1)
2.1	-2.1535(2)	-2.0704(1)	-2.5029(1)
2.0	-2.127 64(7)	-2.109 53(4)	-2.3921(1)
1.9	-2.127 96(4)	-2.124 36(4)	-2.2599(1)
1.8	-2.130 16(4)	-2.130 68(4)	-2.0979(1)
1.7	-2.133 64(4)	-2.135 96(4)	-1.8972(2)
1.6	-2.138 13(4)	-2.141 40(3)	-1.6351(2)

and Ψ_{inner} have 1S symmetry, this is automatically satisfied and ensures that our theorem provides a true upper bound to the $1s2s^1S$ state. This kind of argument could be applied to other selected excited states of many-electron systems as well.

In order to compute the expectation values of the Hamiltonian for this correlated trial wave function, since analytical integration is not possible, we resort to variational quantum Monte Carlo (VMC) methods, which are well suited for the task [5]. The VMC method also readily provides expectation values of the energy within each nodal region [see Eq. (6)]. To illustrate our nodal variational principle, we performed several calculations using different radii for the hyperspherical node (i.e., different values of k). In each case we optimized the other variational parameters by minimizing the total energy.

Table I and Fig. 1 show the expectation values of the energy for different values of the parameter k over the entire space and separately in the two nodal regions, labeled A and B. Note that the two nodal regions are not related by a symmetry operation and so give different values for the expectation value of the Hamiltonian. For any given k , the trial function is an approximation to the first excited 1S state. However, because the usual variational principle is no longer valid, the VMC energy can be above or below the exact value. These energies can be seen in the VMC (full) energy column. There is no *a priori* way to know which values of k are good or bad. The two other columns show the energies in the two nodal regions (respectively outside and inside the nodal hypersphere). It is easy to check (numerically) that the theorem is valid, since for each value of k the maximum of the two nodal energies is always an upper bound to the exact energy [6], which is $-2.145 974$.

The best upper bound is obtained when the energies are the same in the two nodal regions. From Fig. 1 we can see that this crossing occurs for k between 1.8 and 1.9. It is also the case that, as in the standard variational principle, k can be optimized to obtain the best energy bound. Doing so in this case gives an upper bound for the $1s2s^1S$ energy of -2.129 a.u.

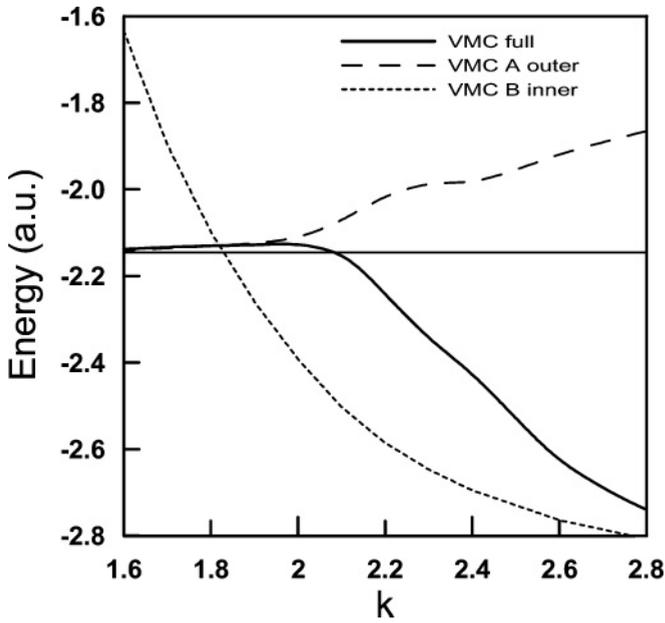


FIG. 1. Full variational and nodal energies for the $1s2s^1S$ excited state of the helium atom, with the trial function described in the text, for different values of k . The solid horizontal line indicates the exact energy. The bold line indicates the total VMC energy, the long-dashed and short-dashed lines indicate, respectively, the VMC energy of the outer and inner nodal cells. Unlike the full VMC energy, which plummets at large k and fails to be variational in this situation, the larger valued of the two broken curves always lies above the horizontal line, thereby satisfying the generalized principle.

Having used our variational formalism, we step back to note that the wave function in Eq. (15) could have been treated, in principle, by the MacDonald-Hylleraas-Undheim theorem because it can be cast into a linear combination of two basis functions. In practice, however, this would require the computation of the relevant matrix elements that are not available in analytical form for an explicitly correlated wave function such as this one. Moreover, with our formalism, we need not be restricted to wave functions that can be decomposed in this way.

To show that our result for $D > 1$ is not an artifact of our choice of Eq. (15), we consider a more sophisticated correlated wave function built using hydrogenic orbitals multiplied by a Jastrow factor

$$[1s(r_1)2s(r_2) + 1s(r_2)2s(r_1)] \exp\left(\frac{r_{12}/2}{1 + dr_{12}}\right), \quad (16)$$

where $1s(r) = e^{-2r}$ and $2s(r) = (r - c)e^{-br}$. The shape of the node is no longer a hypersphere, but one can still distinguish two nodal regions, one where the wave function is positive and the other where the wave function is negative. As before, an upper bound to the exact energy can be similarly extracted. We optimized the latter wave function by hand, and by using the generalized variational principle obtained an upper bound of -2.1423 a.u. It would be possible to employ a still more sophisticated wave function to obtain a still better upper bound.

IV. CONCLUSION

The properties of the nodal structure of wave functions are important in many different fields, ranging, e.g., from the study of quantum dots to the fixed node diffusion Monte Carlo method and from the quantum Hall effect to quantum chaos. Despite the fundamental importance of wave-function nodes (as opposed to orbital nodes), only recently have a few studies [4,7–12] begun to investigate the properties of nodes of both exact and approximate wave functions. In this paper we presented a theorem that actually exploits the nodes and could potentially be useful in many diverse fields. We have shown that information about the nodal structure of the wave function can be exploited in a variational principle that is able to give an upper bound of an arbitrary excited state of a 1D system. Straightforwardly, the theorem does not hold for an arbitrary n -dimensional system since by Courant's theorem the M th excited state might have fewer than $M + 1$ nodal regions. However, we showed that the theorem could still be used in higher-dimensional systems if additional information on the nodal structure is available. We illustrated the higher-dimensional method by applying it to the first excited singlet state of the helium atom.

-
- [1] J. K. L. MacDonald, *Phys. Rev.* **43**, 830 (1933).
 - [2] E. A. Hylleraas and B. Undheim, *Z. Phys.* **65**, 759 (1930).
 - [3] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Wiley, New York, 1953), Vol. I.
 - [4] D. Bressanini and P. J. Reynolds, *Phys. Rev. Lett.* **95**, 110201 (2005).
 - [5] D. Bressanini and P. J. Reynolds, *Adv. Chem. Phys.* **105**, 37 (1999).
 - [6] G. W. F. Drake and Zong-Chao Van, *Chem. Phys. Lett.* **229**, 486 (1994).
 - [7] D. Bressanini and G. Morosi, *J. Chem. Phys.* **129**, 054103 (2008).
 - [8] T. C. Scott, A. Luchow, D. Bressanini, and J. D. Morgan, *Phys. Rev. A* **75**, 060101(R) (2007).
 - [9] L. Mitas, *Phys. Rev. Lett.* **96**, 240402 (2006).
 - [10] M. Bajdich, L. Mitas, G. Drobný, and L. K. Wagner, *Phys. Rev. B* **72**, 075131 (2005).
 - [11] D. M. Ceperley, *J. Stat. Phys.* **63**, 1237 (1991).
 - [12] D. Bressanini, D. M. Ceperley, and P. J. Reynolds, *Recent Progress in QMC* (World Scientific, Singapore, 2002).