# Introduction and statement of purpose

The theory of quantum mechanics is a remarkable achievement in science which leaves the chemist with a straightforward procedure to get information about molecules and reactions: solve the Schrodinger equation. Unfortunately, for all but a very few systems, the Schrodinger equation is not exactly soluble and we must resort to finding approximate solutions. These are commonly found using some type of variation method. In this research, we do not seek to establish the validity of the theory of quantum mechanics; instead, we wish to help validate the implementation of quantum mechanics computationally.

The variation method (linear or non-linear) is the method of choice for approximating solutions to Schrodinger's equation. This method gives only an upper bound to the energy eigenvalues and makes no claim as to the quality of the eigenfunctions (i.e. improvements in the energy do not necessarily imply improvement in the eigenfunctions). In principle, one can get arbitrarily close to the true energy eigenvalues. Unfortunately, the size of problems usually demands more than the available computational power and time; and improvements in computer technology and programming techniques cannot keep up with the interest in larger and larger problems. Thus in many cases convergence to the true eigenvalues is not achieved, and even if it is, one has no way to know that they have indeed converged.

Lower bounds are thus important to bound the true eigenvalues in a certain interval. But they are also difficult to calculate. The difficulty lies in the lack of a general and straightforward procedure (c.f. the variation method for upper bounds) and the complexity of the calculations (usually requiring the square of the Hamiltonian operator). This research is a small part of a long and continuing effort by many researchers to improve the ease and accuracy of lower bound calculations.

We will do three things in this thesis:

1. Present the variation method and show WHY, in principle, lower bound calculations are much more difficult than upper bound calculations.

2. Review some of the standard lower bound methods and point out HOW, computationally, they are more difficult than the variation method.

3. Introduce the method of arbitrary choice without truncation and show how it fails on application to the helium atom.

# Notes for the reader

1. The thesis is dedicated to the development and testing of the method of arbitrary choice without truncation. My research, however, dealt more generally with lower bound methods of any kind. An extension of a lower bound methed by Calogero and Marchioro is touched upon briefly in section 2.4.2. This extension is dealt with in detail in appendix G, where it is written as a paper. The computer code used to perform the Hylleraas variational calculations in that paper is recorded in appendix F. Some comments about lower bounds by the effective field method are also written in the form of a paper (mentioned in section 2.5) and are located in appendix H. A third paper, a side result of the research, is presented in appendix I. At the time of the defense, these papers were in the submission process to various journals. To avoid copyright infringements, access to these three appendices (G, H, and I) has been permanently limited to the Virginia Tech campus. In the case that Virginia Tech places these appendices freely on the internet, Virginia Tech is solely responsible for copyright violations.

2. The inner product commonly used by chemists is denoted $\langle\psi|H|\psi\rangle$ which represents the integral of $\psi^*H\psi$. We will abandon this notation and use instead the notation used by mathematicians:

$$(H\psi,\psi).$$

This is conjugate linear in the second argument, instead of the first, which is more natural since it is reminiscent of complex numbers: $z = a + i\, b = (a,b)$

3. Atomic units are used throughout: $h/(2\pi) = e = m = 1$, where $h = 6.6262\times10^{-34}$ J·s is Planck's constant; $e$=$1.6022\times10^{-19}$ C is the magnitude of an electronic charge; and $m$=$9.1095\times10^{-31}$ kg is the mass of an electron.

4. The infinite-nuclear mass approximation and spinless Hamiltonian is used throughout.

5. Summation indices are omitted when obvious. $\sum$ without denotation denotes either an infinite sum (usually over a countable complete set of functions) or a finite sum (usually over the number of electrons) which should be clear in context. $\sum''$ specifically denotes the double sum $\sum_{i=1}^{N}\sum_{j=i+1}^{N}$ .

6. "Q.E.D." marks the end of a proof.

7. Bold unitalicized capital letters denote spaces, e.g. **M**, **H**.

# Chapter One
# Introduction to upper and lower bounds

## 1.1 The variation theorem and method

### 1.1.1 The variation theorem (non-rigorously)

We first present a common form of the variation theorem introduced by Eckart in 1930 [Ek] and its simple proof.

Variation theorem: Given that the self-adjoint Hamiltonian, $H$, has a complete set of eigenfunctions, and there exist a lowest eigenvalue, $\lambda_1$, then the minimum of the Rayleigh quotient, defined as $R(\psi_n) = (H\psi_n,\psi_n)/(\psi_n,\psi_n)$, over a set of nonzero trial functions $\{\psi_n\}$, gives an eigenvalue greater than or equal to the lowest eigenvalue of $H$.

Proof: Let $\{\lambda_n\}$ be the corresponding set of eigenvalues ordered from smallest to largest, and $\{\phi_n\}$ be the complete set of corresponding orthonormal eigenfunctions for the Hamiltonian $H$. Any trial function, $\psi$, can be written as a finite or infinite linear combination of the $\{\phi_n\}$:

$$\psi = \sum c_n\phi_n$$

Taking the Rayleigh quotient of $\psi$ we find:

$$(H\psi,\psi)/(\psi,\psi) \;=\; (H \sum c_m\phi_m, \sum c_n\phi_n) / (\sum c_m\phi_m, \sum c_n\phi_n)$$

$$= \; \sum \sum c_m c_n^*(H\phi_m,\phi_n) / \sum |c_m|^2 \;=\; \sum \sum c_m c_n^*(\lambda_m\phi_m,\phi_n) / \sum |c_m|^2$$

$$= \sum |c_m|^2 \lambda_m / \sum |c_m|^2 \;\geq\; \sum |c_m|^2 \lambda_1 / \sum |c_m|^2 \;=\; \lambda_1$$

Q.E.D.

This form of the variation theorem is very simple and is presented in many texts (e.g. [Ch, PW, Sh]). However, the theorem in the form above is useless to the quantum chemist because all atomic and probably all molecular systems may not have a complete set of eigenfunctions, e.g. it is well known that the hydrogenic eigenfunctions are not a complete set [SL]. By this we mean that an arbitrary $\psi$ in the domain of the operator may not be represented as the limit of a linear combination of eigenfunctions. Another use of the term completeness is used in spectral theory - it is that notion that is guaranteed for self-adjoint operators, not the former. We now present a proof that all atomic Hamiltonians do not have a complete set of eigenfunctions.

Theorem: Any atomic Hamiltonian, $H$, in the spinless infinite-nuclear mass approximation does not possess a total orthonormal set of eigenfunctions.

Proof: We start with three facts:

1. The spectrum of $H$ contains the interval $[0,\infty)$ wherein it is purely continuous, i.e. there are no eigenvalues [Wd].

2. For a self-adjoint operator ($H$ is such), for any spectral point $\lambda$, given $\varepsilon>0$, there exists a function $u\in \mathrm{Dom}(\mathbf{H})$ of unit norm such that [HS]

$$\|(H-\lambda)u\| < \varepsilon \,.$$

3. For any closed operator $A$: $\mathrm{Dom}(A)\subset\mathbf{X}\rightarrow\mathbf{Y}$, where $\mathbf{X}$ and $\mathbf{Y}$ are normed linear spaces, we have that given a sequence $\{x_n\}\subset\mathrm{Dom}(A)$ such that $x_n\rightarrow x\in\mathbf{X}$ and $Ax_n \rightarrow y\in\mathbf{Y}$, then $x\in\mathrm{Dom}(A)$ and $Ax = y$ [Kr].

We start with fact 2. Let $\sqrt{e} = \lambda > 0$, $\lambda$ a real spectral point (by fact 1). Then for there exist a $u\in\mathrm{Dom}(\mathbf{H})$ of unit norm such that:

$$0 \;\leq\; (H^2u,u) - 2\,\lambda(Hu,u) + \lambda^2 \;=\; \|(H-\lambda)u\| \;<\; \lambda^2 \;.$$

This implies that

$$(H^2u,u) \;<\; 2\,\lambda(Hu,u) \,.$$

Since $(H^2u,u) > 0$ for all non-zero $u$, and $\lambda > 0$, we know that $(Hu,u) > 0$.

We now assume that unit norm eigenfunctions $\{\psi_n\}$ (whose eigenvalues we denote as $E_n$) form a total orthonormal set. Then $u$ can be written as

$$u = \lim_{N \to \infty} u_N \quad \text{where } u_N = \sum_{n=1}^{N} c_n y_n \quad .$$

Then for $M < N < \infty$, we have

$$\|Hu_M - Hu_N\| = \|H \sum_{n=M}^{N} c_n y_n\| = \|\sum_{n=M}^{N} c_n H y_n\| = \|\sum_{n=M}^{N} c_n E_n y_n\| \to 0 \text{ as } M,N \to \infty$$

since $E_n$ is bounded between 0 (see fact 1) and some finite negative number (for the finite lower bound see [Lb,Kt1] and Appendix A). Thus the sequence $\{Hu_n\}$ is cauchy in **H** and because **H** is complete, this implies that the sequence converges. Since $H$ is a closed operator, fact 3 implies that

$$Hu = \lim_{N \to \infty} Hu_N \quad \text{where } Hu_N = \sum_{n=1}^{N} c_n H y_n = \sum_{n=1}^{N} c_n E_n y_n \quad .$$

We now want to show that $(Hu,u) = \lim_{N \to \infty} (Hu_N, u_N)$.

$$|(Hu,u) - (Hu_N, u_N)| = |(Hu, u - u_N) + (Hu - Hu_N, u_N)|$$

$$\leq |(Hu, u - u_N)| + |(Hu - Hu_N, u_N)|$$

$$\leq \|Hu\| \cdot \|u - u_N\| + \|Hu - Hu_N\| \cdot \|u_N\|$$

$$\leq c\|u - u_N\| + \|Hu - Hu_N\|$$

$$\to 0 \text{ as } N \to \infty \text{ where } c \text{ is some constant.}$$

$$(\text{both } \|u - u_N\| \text{ and } \|Hu - Hu_N\| \to 0 \text{ as } N \to \infty)$$

The integral $(Hu_N, u_N)$ will always be negative since it will be a linear combination of the eigenvalues which are all negative. This implies that $(Hu,u)$ will also be negative or zero. This contradicts the assumption that $(Hu,u) > 0$. Thus $u$ cannot be contained in the closed linear span of the eigenfunctions, which implies that the eigenfunctions do not form a total orthonormal set.

Q.E.D.

### 1.1.2 The variation theorem (rigorously)

Fortunately, the variation theorem, without the hypothesis of completeness, can still be proven; in fact, it falls out naturally from obtaining the spectral resolution of the operator. Unfortunately, spectral theory is not so common among chemists, so we present here a different proof.

The definition of a positive semi-definite operator, $A$, is

$$A \geq 0 \iff (Au,u) \geq 0 \ \forall\, u \in \mathrm{Dom}(A) .$$

In addition, if $A$ is also a self-adjoint operator, we have that

$$A \geq 0 \iff \sigma(A) \subset [0,\infty) ,$$

where $\sigma(A)$ is the set of all eigenvalues and points in the continuous spectrum, if any.

<u>Theorem:</u> For any self-adjoint operator, $A$, with $\sigma_1$ as the infimum of the spectrum, then

$$\inf_{\substack{u \in Dom(A) \\ u \neq 0}} (Au,u)/(u,u) \geq \sigma_1 .$$

<u>Proof:</u> The spectrum of $A$ is contained in the interval $[\sigma_1,\infty)$. Then the spectrum of $A$-$\sigma_1$ is contained in the interval $[0,\infty)$. Thus $A$-$\sigma_1$ is a positive semi-definite operator and

$$\inf_{\substack{u \in Dom(A) \\ u \neq 0}} (\{A-\sigma_1\}u,u)/(u,u) \geq 0.$$

From this it follows that

$$\inf_{\substack{u \in Dom(A) \\ u \neq 0}} (Au,u)/(u,u) \geq \sigma_1.$$

Q.E.D.

This states only that the Rayleigh quotient of a specific $u \in \mathrm{Dom}(A)$ is always greater than or equal to the infimum of the spectrum - it says nothing about whether the infimum of the spectrum is an eigenvalue or part of the continuous spectrum. For a general atomic or molecular Hamiltonian, one does not know whether or not there are any bound states (i.e. eigenvalues below the continuous spectrum) so a variational calculation may be converging to a number, but *a priori* one does not know if this is even an eigenvalue

Even more troublesome is the possibility that the infimum of the spectrum is $-\infty$; thus a variational calculation would never converge. This possibility has been eliminated as it is known that the spectrum of atomic and molecular Hamiltonians are bounded below [Kt1, Lb]. A simpler proof of this fact can be obtained using the united atom theorem (see appendix). The united atom theorem states that the infimum of the spectrum of a molecule consisting of $N$ electrons and $M$ positive charges is greater than or equal to the infimum of the spectrum of a single $N$-electron atom with nuclear charge $M$. The spectrum of an atom can be bounded below by summing the appropriate hydrogenic eigenvalues according to the Pauli principle (thus neglecting electron-electron repulsion). This is completely rigorous since the hydrogenic spectrum is known exactly.

### 1.1.3 The linear variation method

A naive, but rigorous, implementation of the variation theorem is immediately obvious: one just thinks of trial functions and takes the lowest Rayleigh quotient. A more efficient implementation is to construct a finite set of linearly independent trial functions, $\{\phi_n\}$, called a basis set, and determine the eigenvalues of a generalized eigenvalue problem:

$$[H]x_n = \Lambda_n[S]x_n \, ,$$

where $[H]$ and $[S]$ are matrices with elements $(H\phi_m,\phi_n)$ and $(\phi_m,\phi_n)$, respectively; $\Lambda_n$ and $x_n$ are the approximate eigenvalues and eigenvectors of the operator $H$. (In general we differentiate the *operator* $A$ from the *matrix* $[A]$, which is a representation of the operator $A$ projected onto the finite subspace spanned by the basis set $\{\phi_n\}$). This is called the linear variation method. If the basis set $\{\phi_n\}$ is orthonormal, then $[S]$ becomes the identity matrix $[I]$ and the generalized eigenvalue problem reduces to:

$$[H]x_n = \Lambda_n x_n \, .$$

The linear variation method is especially efficient because not only does it provide an upper bound to the ground state energy, but it also provides upper bounds to the first $M$-1 excited state energies, where $M$ is the dimension of the basis set. This result is commonly known as MacDonald's theorem in chemistry [Ch,Mc].

## 1.2 The min-max and max-min principles

The following characterizations assume that we are dealing with a self-adjoint operator whose spectrum consists of finitely or infinitely many discrete eigenvalues. If continuous spectrum exists, then the characterizations apply only to those discrete eigenvalues below the continuous spectrum.

The Rayleigh quotient gets its name from Lord Rayleigh, whom Weinstein and Stenger [WS] acknowledge with the first presentation [Ry] of the variation theorem (called the Rayleigh principle outside chemistry circles). The Rayleigh principle characterizes the lowest eigenvalue, $\lambda_1$, of an operator, $A$, as:

$$\lambda_1 = \min_{\substack{y \in Dom(A) \\ y \neq 0}} R(\psi) = \min_{\substack{y \in Dom(A) \\ y \neq 0}} (A\psi,\psi)/(\psi,\psi) \, ,$$

where the minimum is taken over all non-zero functions, $\psi$, in the domain of the operator, Dom($A$). This was extended by Weber [Wb] to give a characterization of higher eigenvalues:

$$\lambda_n = \min_{\substack{y \in Dom(A) \\ (y, y_{k<n})=0 \\ y \neq 0}} R(\psi) \, ,$$

where the minimum is now taken over all non-zero functions in the domain of the operator that are orthogonal to the first $n$-1 eigenfunctions, $\psi_k$, $k$=1,...,$n$. Another characterization of the $n$-th eigenvalue without reference to the first $n$-1 eigenfunctions is possible. This is called the min-max principle [WS], generalized to operators on infinite-dimensional domains by Polya [Po]:

$$\lambda_n = \min_{\mathbf{M}_n \subset Dom(A)} \max_{\substack{y \in \mathbf{M}_n \\ y \neq 0}} R(\psi) \, ,$$

where the minimum is over all $n$-dimensional subspaces, $\mathbf{M}_n$, of the domain of the operator, Dom($H$), and the maximum is taken over the functions in $\mathbf{M}_n$. Note that the min-max principle reduces to Rayleigh's principle for $n$=1

Instead of looking at all possible $\mathbf{M}_n$, we may choose to look at only a few, so the minimum may not be reached - thus an upper bound to $\lambda_n$ is obtained. The maximization of the Rayleigh quotient presents no significant difficulty since the trial functions come from an $n$-dimensional space, which means we need only find the largest eigenvalue of an $n$-dimensional matrix. The min-max principle thus immediately leads to a practical algorithm involving one or more *n*-dimensional matrix eigenvalue problems which guarantees an upper bound to $\lambda_n$ (c.f. the linear variation method).

Yet another characterization is possible, called the max-min principle [WS] presented first by Weyl [Wy ]:

$$\lambda_n = \max_{M_{n-1} \subset \mathbf{H}} \min_{\substack{y \in Dom(A) \\ (y, M_{n-1}) = 0 \\ y \neq 0}} R(\psi) ,$$

where the maximum is taken over all *n*-1 dimensional subspaces, $\mathbf{M}_{n-1}$, and the minimum is taken over all functions in the domain of the operator orthogonal to the $\mathbf{M}_{n-1}$ subspace. If we only look at some of the $\mathbf{M}_{n-1}$ subspaces, then the maximum may not be reached - and thus a lower bound results.

However, we run into a computational problem when we try to minimize the Rayleigh quotient. This is not equivalent to a matrix eigenvalue problem because the space of available trial functions is infinite dimensional (for atomic and molecular Hamiltonians). Thus the lower bound problem does not reduce to a finite dimensional problem as is the case for the upper bound problem. This is the essential difference between the two problems and why lower bounds are so difficult to obtain.

Note: The inner product is sesquilinear and all the operators we will consider are linear. If we let $\alpha$ be any complex number and $\psi$ any function in the domain of the operator *A*, then:

$$(A\alpha\psi,\alpha\psi)/(\alpha\psi,\alpha\psi) = (\alpha A\psi,\alpha\psi)/(\alpha\psi,\alpha\psi) = \alpha\alpha^*(A\psi,\psi)/\alpha\alpha^*(\psi,\psi) = (A\psi,\psi)/(\psi,\psi)$$

which means we can restrict our sampling of functions in maximums and minimums to functions of unit norm. The Rayleigh quotients then become simply $R(\psi) = (A\psi,\psi)$.

## 1.3. Summary

Upper bounds are easily determined by the variation method. This is because upper bounds can be calculated using a finite-dimensional subspace of the domain of the operator.  Lower bounds, however, are very difficult to obtain, and this difficulty arises because lower bounds must be calculated in infinite-dimensional subspaces of the domain of the operator; thus no finite-dimensional matrix eigenvalue problem *by itself* will yield lower bounds.

In the next chapter, we see that many lower bound methods do *end up* as a matrix eigenvalue problem. However, these matrix eigenvalue problems generally do not resemble the upper bound variational ones, are much more difficult to solve, and can only be formed after certain *a priori* information about the operator has been utilized.

# Chapter Two

# Review of standard lower bound methods

## 2.1 Weyl's inequality

Given an operator $H = A + B$, the eigenvalues of $A$ and $B$ unfortunately do not simply add to give the eigenvalues of $H$. There is however a relationship known as Weyl's inequality that relates them non-linearly. Recalling that $\mathrm{Dom}(H) = \mathrm{Dom}(A) \cap \mathrm{Dom}(B)$ implies that $\mathrm{Dom}(H) \subseteq \mathrm{Dom}(A)$ and $\mathrm{Dom}(H) \subseteq \mathrm{Dom}(B)$, we can derive Weyl's inequality from the max-min principle as follows:

$$\lambda_{i+j-1}(H) \;=\; \max_{M_{i+j-2}\subset \mathbf{H}} \; \min_{\substack{u\in Dom(H)\\ (u,M_{i+j-2})=0\\ \|u\|=1}} (\{A+B\}u,u) \;=\; \max_{M_{i+j-2}\subset \mathbf{H}} \; \min_{\substack{u\in Dom(H)\\ (u,M_{i+j-2})=0\\ \|u\|=1}} [\,(Au,u) + (Bu,u)\,]$$

$$\geq \max_{M_{i+j-2}\subset \mathbf{H}} \Big[ \min_{\substack{u\in Dom(H)\\ (u,M_{i+j-2})=0\\ \|u\|=1}} (Au,u) + \min_{\substack{u\in Dom(H)\\ (u,M_{i+j-2})=0\\ \|u\|=1}} (Bu,u) \Big] \;\geq\; \max_{M_{i+j-2}\subset \mathbf{H}} \Big[ \min_{\substack{u\in Dom(A)\\ (u,M_{i+j-2})=0\\ \|u\|=1}} (Au,u) + \min_{\substack{u\in Dom(B)\\ (u,M_{i+j-2})=0\\ \|u\|=1}} (Bu,u) \Big]$$

$$\geq \min_{\substack{u\in Dom(A)\\ (u,\mathbf{A}_{i-1}\mathbf{B}_{j-1})=0\\ \|u\|=1}} (Au,u) \;+\; \min_{\substack{u\in Dom(B)\\ (u,\mathbf{A}_{i-1}\mathbf{B}_{j-1})=0\\ \|u\|=1}} (Bu,u) \;\geq\; \lambda_i(A) \;+\; \lambda_j(B) \;\;,$$

where $\mathbf{A}_{i-1}\mathbf{B}_{j-1}$ is the space spanned by the first $i$-1 eigenvectors of $A$ and the first $j$-1 eigenvectors of $B$. The dimension of this space is at most $i+j$-2. Weyl's inequality in its most general form applies to $H = A_1 + A_2 + ... + A_N$, for any finite $N$.

## 2.2 Temple-like bounds

### 2.2.1 Temple's formula

We start with the equation,

$$(\{H-E_1\}\{H-E_2\}\psi,\psi) \geq 0 \;,$$

where $E_1$ and $E_2$ are the first and second eigenvalues (not equal) of the Hamiltonian $H$. This equation is valid if $E_1 \leq (H\psi,\psi) \leq E_2$ . This constraint on $\psi$ is sufficient, but it may not be necessary. Christoffersen offers a proof of this inequality without any constraint on $\psi$, but he incorrectly assumes that $H$ has a complete set of eigenfunctions [Ch].

This can be rewritten to give a lower bound to $E_1$:

$$(E_1\psi,\psi) \;\geq\; \frac{(H^2 y\,,y) - E_2(Hy\,,y)}{(Hy\,,y) - E_2(y\,,y)} \;.$$

The drawback of this formula is that it requires a separation of the ground and first excited state, i.e. one must have a lower bound, $L$, to $E_2$ that is also an upper bound to $E_1$ and $(H\psi,\psi)$ in order to get a lower bound to $E_1$:

$$E_1 \le (H\psi,\psi) \le L \le E_2.$$

Obtaining such a bound may be more difficult than finding a lower bound to $E_1$ directly. Some authors have taken experimental values for $E_2$ and set $L = E_2$ [Ki].

Aside from this, one encounters computational difficulty. Temple's formula requires solving integrals of the form $(H\psi,H\psi)$ which are much more difficult than the $(H\psi,\psi)$ integrals involved in upper bounds calculations. Harder integrals are a key problem for lower bound methods, and we will run into this repeatedly.

### 2.2.2 Weinstein's formula

In 1934, Weinstein [Wn] introduced the following inequality for an operator, $A$, with a complete set of eigenfunctions with eigenvalues, $\lambda_n$:

$$\sqrt{I_2 - I_1^2} + I_1 \ \ge\ \lambda_n \ \ge\ I_1 - \sqrt{I_2 - I_1^2}\ ,$$

$$\text{where } I_1 = (A\psi,\psi) > \lambda_1 \text{ and } I_2 = (A\psi,A\psi)\ .$$

So given an energy $I_1$ and corresponding $\psi$, Weinstein's formula allows construction of a bracket which holds at least one true eigenvalue. In general, one cannot know which eigenvalue(s) is in this bracket, so the upper and lower bounds are not very useful. Thus $I_2 \to I_1^2$ is necessary, but not sufficient, for convergence to a *specific* eigenvalue.

Weinstein's result can easily be modified to give a more Temple-like inequality for $\lambda_n$ if a lower bound to $\lambda_{n+1}$ is known (that is also an upper bound to $\lambda_n$) [Kn]. Since chemical Hamiltonians do not posses complete sets of eigenfunctions, Weinstein's result seem not apply. A slight modification of the result applies for any self-adjoint operator [Kt2]:

$$\sqrt{I_2 - I_1^2} + I_1 \ \ge\ \lambda \ \ge\ I_1 - \sqrt{I_2 - I_1^2}\,,$$

where $\lambda$ is now any spectral point (i.e. it may be a point in the continuous spectrum instead of an eigenvalue).

### 2.2.3 A generalized eigenvalue problem

If one knows a lower bound, $L$, to the $n$-th eigenvalue ($n \geq 2$) of an operator $A$ that is also an upper bound to the $n$-1 lower eigenvalues (c.f. Temple's method), then one can form a variational procedure to determine lower bounds to the $n$-1 lowest eigenvalues. To describe this procedure we refer to figure 1 below.

We assume that $A$ is bounded below with finitely or infinitely many eigenvalues, where the eigenvalues of interest, and all those below, have finite multiplicity. If $A$ has essential spectrum, we require that it lies above the eigenvalues of interest. Figure 1 depicts an operator with three eigenvalues and continuous spectrum above.

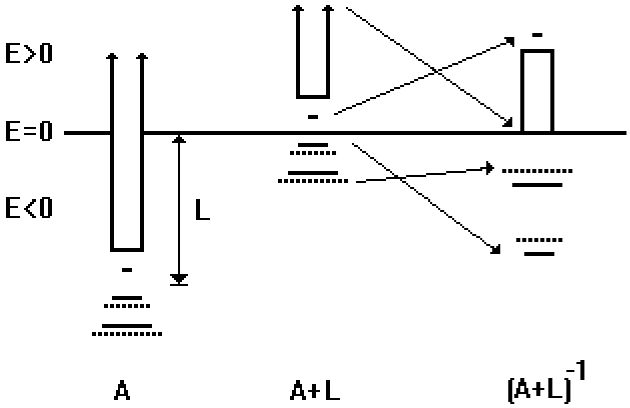

**Figure 1.** The spectrum of the operator $A$ is raised by a constant $L$ so that only discrete spectrum lies below $E$=0. The spectrum is then inverted. Solid lines represent the eigenvalues and dotted lines represent desired lower bounds (adapted from [Sc]).

We define a new operator, $A+L$, whose spectrum is entirely discrete below $E = 0$. Inverting this operator gives yet another operator, $(A+L)^{-1}$, whose spectrum below $E = 0$ is also discrete. It is vital that the negative portion of the spectrum of $(A+L)^{-1}$ be purely discrete with finitely many eigenvalues (see below), which is why we first raised the spectrum by $L$. This being the base, we may calculate upper bounds to the eigenvalues of $(A+L)^{-1}$ using the variation method. These upper bounds for the eigenvalues of $(A+L)^{-1}$ can be easily translated into lower bounds to the eigenvalues of $A$.

The finite-dimensional eigenvalue problem to be solved is

$$[(A+L)^{-1}]x_n = \zeta_n x_n ,$$

where an orthonormal basis set $\{\phi_n\}$ has been chosen. Solution of this requires knowledge of the form of the inverse operator, which may not be known. A simpler option is to choose $\{(A+L)\phi_n\}$ as the basis set, which transforms the above problem into:

$$[(A+L)]y_n = \xi_n[(A+L)^2]y_n .$$

This form avoids the need to know the form of $(A+L)^{-1}$, but forces one to evaluate integrals of the form $(\{A+L\}^2\phi_m,\phi_n)$ which as mentioned before, can be very difficult.

One should note that if $(A+L)^{-1}$ has $n$ eigenvalues below $E=0$, then the variational calculation must be large enough to reveal negative upper bounds to all $n$ eigenvalues. If not, then a lower bound to the first eigenvalue of $(A+L)$ will not be found. For instance, for the problem in figure 1, if only one negative eigenvalue of $(A+L)^{-1}$ is found, then only a lower bound to the second eigenvalue of $(A+L)$ results. This is why it is necessary to have only a finite number of eigenvalues in the negative portion of the spectrum.

## 2.3 Intermediate problems

### 2.3.1 Aronszajn intermediate problems

Many eigenvalue problems have an operator of the form $A = A_o + B$ where $A_o$ is an exactly soluble system (or soluble to any desired accuracy) and $B$ is a positive definite operator (i.e. $(B\psi,\psi) > 0$ for any function $\psi$ in the domain of $B$). We usually call $A_o$ the base problem, and $B$ the perturbation.

The following theorem will allow us to relate simply the eigenvalues of $A$ and $A_o$.

<u>Theorem:</u> A positive definite perturbation, $B$, always increases the eigenvalues of a base problem, $A_o$.

Proof: Using the min-max principle and the positivity of $B$, we have:

$$\lambda_n(A_o+B) = \min_{M_n \subset Dom(A)} \max_{\substack{y \in M_n \\ \|y\|=1}} (\{A_o+B\}\psi,\psi) > \min_{M_n \subset Dom(A)} \max_{\substack{y \in M_n \\ \|y\|=1}} (A_o\psi,\psi)$$

$$\geq \min_{M_n \subset Dom(A_o)} \max_{\substack{y \in M_n \\ \|y\|=1}} (A_o\psi,\psi) = \lambda_n(A_o)$$

Q.E.D.

This tells us that we already have lower bounds to eigenvalues of $A$; i.e. the eigenvalues of $A_o$ are themselves lower bounds to the eigenvalues of $A$.

Aronszajn initiated the method of intermediate problems for operators of the form $A = A_o + B$, where $A_o$ is solvable to desired precision and $B$ is positive definite. He envisioned a sequence of positive definite operators $\{B_i\}$ such that $\{B_i\}$ converges to $B$ and $B_{i+1}-B_i > 0$. This would allow one to construct a sequence of eigenvalue problems:

$$A_o < A_o + B_1 < A_o + B_2 < \ldots < A_o + B = A .$$

A simple example of such a sequence is $B_i = (½)^i B$ . This example is not useful, however, since each intermediate operator is just as difficult to solve as the original operator - i.e. all the difficulties of $B$ are still present. Aronszajn found a way to construct a useful sequence $\{B_i\}$ where each $B_i$ was formed using projection operators.

In the Hilbert space of quantum chemistry, there exists the standard inner product,

$$(u,x) = \int (u)(x^*) .$$

We define a new inner product [Bz,FR] of the form

$$((u,x)) = (Bu,x) = \int (Bu)(x^*) ,$$

where $B$ is a positive-definite Hermitian perturbation to $A_o$.

Let $P_{\mathbf{K}}$ be a projection operator onto the subspace $\mathbf{K} \subset \mathbf{H}$, in the new inner product (but not in the original inner product). For simplicity of presentation, we will initially restrict $\mathbf{K}$ to be one-dimensional, i.e. the span of a single normalized vector $v$. Thus, for all $u \in \mathbf{H}$,

$$P_{\mathbf{K}}(u) = ((u,v))v .$$

We define a new operator, $BP_{\mathbf{K}}$, which makes $A+BP_{\mathbf{K}}$ an intermediate problem:

$$(A_0u,u) < (A_0+BP_{\mathbf{K}}u,u) < (A_0+Bu,u) .$$

The first inequality follows from the idempotency of projection operators and the fact that $B$ is Hermitian in the original inner product and $P_{\mathbf{K}}$ is Hermitian in the new inner product:

$$(BP_{\mathbf{K}}u,u) = ((P_{\mathbf{K}}u,u)) = ((P_{\mathbf{K}}P_{\mathbf{K}}u,u)) = ((P_{\mathbf{K}}u,P_{\mathbf{K}}u)) \geq 0$$

The second inequality follows merely from the Hermitian nature of $B$:

$$(BP_{\mathbf{K}}u,u) = ((P_{\mathbf{K}}u,u)) \leq ((u,u)) = (Bu,u)$$

Thus $A_0+BP_{\mathbf{K}}$ is an intermediate problem. If we remove the restriction that **K** be one-dimensional, then a whole sequence of increasing intermediate problems can be constructed by increasing the dimension of **K**. In general, however, an arbitrary choice of **K** will not create an easily soluble intermediate problem.

### 2.3.2 Bazley's special choice

Bazley's claim to fame is his famous "special choice". By choosing **K** cleverly, one can form a sequence of finite-dimensional problems that give increasing lower bounds to $A_0+B$. We illustrate this when **K** is only one-dimensional.

If the base problem $A_0$ is exactly solvable, with eigenpairs $\{\lambda_n,u_n\}$, then choose $\mathbf{K} = B^{-1}\mathbf{M} = \text{span}[N^{-1}B^{-1}u_1]$, where $N = \sqrt{(B^{-1}u_1,u_1)}$ is the normalization factor in the new inner product. Given an arbitrary function $u \in \mathbf{H}$ of unit norm, we investigate the Rayleigh quotients of our intermediate problem.

$$(\{A_0+BP_{B^{-1}\mathbf{M}}\}u,u) = (A_0u,u) + (BP_{B^{-1}\mathbf{M}}u,u) = (A_0u,u) + ((P_{B^{-1}\mathbf{M}}u,u))$$

$$= (A_0u,u) + ((\,((u,N^{-1}B^{-1}u_1))N^{-1}B^{-1}u_1 , u)) = (A_0u,u) + N^{-2}\,((u,B^{-1}u_1))((B^{-1}u_1,u))$$

$$= (A_0u,u) + N^{-2}\,(u,u_1)(u_1,u) = (A_0u,u) + (u_1,u)^2 / (B^{-1}u_1,u_1).$$

Thus if $u$ is orthogonal to $u_1$ (i.e. $(u,u_1)=0$ ), then we have

$$(\{A_0+BP_{B^{-1}\mathbf{M}}\}u,u) = (A_0u,u)$$

If $u$ is not orthogonal to $u_1$, then $u = u_1$. This gives:

$$(\{A_0+BP_{B^{-1}\mathbf{M}}\}u_1,u_1) = (A_0u_1,u_1) + 1/(B^{-1}u_1,u_1)$$

Using Rayleigh's principle, we have that the minimum over all $u\in\mathbf{H}$ of unit norm simplifies to a minimum of only two functions, $u_1$ and $u_2$.

$$\min_{\substack{y\in Dom(A)\\ \|y\|=1}} R(\mathrm{u}) = \min[\,(A_0u_2,u_2)\,,\,(A_0u_1,u_1) + 1/(B^{-1}u_1,u_1)\,]$$

$$= \min[\,\lambda_2\,,\,\lambda_1 + 1/(B^{-1}u_1,u_1)\,]$$

So the only change in the spectrum of $A_0$ caused by $BP_{B^{-1}\mathbf{M}}$ is the raising of $\lambda_1$ by the amount $(B^{-1}u_1,u_1)$. We then have that the set $\{\lambda_1+1/(B^{-1}u_1,u_1), \lambda_2, \lambda_3, \ldots\}$, when ordered, gives lower bounds to the eigenvalues of $A_0+B$.

When the dimension of **M** is increased by adding more eigenfunctions of the base problem, we find similar things happening. If **M** is *m*-dimensional, then *m* of the eigenvalues of $A_0$ will increase. To find the new eigenvalues, an *m*-dimensional variational problem must be solved. Thus, with Bazley's special choice, an infinite-dimensional problem *effectively* reduces to a finite dimensional problem. We stress the word "effectively" since we are really still finding the eigenvalues of an infinite-dimensional problem, but the infinite-dimensional part comes from the base problem, and this is exactly soluble by hypothesis. If the subspace **M** is not formed from the eigenfunctions of the base problem, then the benefit is lost. We explain this pictorially below.

We divide Hilbert space into two parts: $\mathbf{H}=\mathbf{M}\oplus\mathbf{M}^{\perp}$, where **M** is the finite-dimensional subspace formed from linear combinations of a finite number of base problem eigenvectors and $\mathbf{M}^{\perp}$ is infinite-dimensional orthogonal complement. Bazley's special choice admits an infinite-dimensional matrix decomposition of the following form:

| | |
|---|---|
| **M×M**<br>finite-dimensional matrix | **M×M$^\perp$**<br>0 |
| **M$^\perp$×M**<br>0 | **M$^\perp$×M$^\perp$**<br>infinite-dimensional matrix |

Pictorially, Bazley's special choice reduces a general infinite-dimensional intermediate eigenvalue problem to a block diagonal problem. The infinite-dimensional matrix corresponds to part of the exactly soluble base problem, and the finite-dimensional matrix is solved variationally. The eigenvalues of the *entire* matrix (finite- and infinite-dimensional parts) form the lower bound spectrum.

Without Bazley's special choice (meaning that the subspace **M** is not solely composed of linear combinations of a finite set of base problem eigenvectors), the block diagonal form above is not achieved, and the overlap of the finite- and infinite-dimensional matrices leaves an insoluble infinite-dimensional matrix where we write $\mathbf{H}=\mathbf{M}\oplus\mathbf{R}\oplus\infty$ (**R** and ∞ are finite- and infinite-dimensional subspaces, respectively, such that $\mathbf{M}^\perp=\mathbf{R}\oplus\infty$).

| | | |
|---|---|---|
| **M×M**<br>finite-dimensional matrix | **M×R**<br>finite-dimen. matrix | **M×∞**<br>0 |
| **R×M**<br>finite-dimensional matrix | **R×R**<br>finite-dimen. matrix | **R×∞**<br>finite/infinite-dimensional matrix |
| **∞×M**<br>0 | **∞×R**<br>finite/infinite-dimen. matrix | ∞×∞<br>infinite-dimensional matrix |

## 2.3.3 Truncation

Instead of constructing an intermediate problem by changing the perturbation, the base problem can be modified instead. Given the operator $A = A_0 + B$, where $A_0$ is exactly soluble, one can define the truncated base problem operator [BF] as:

$$A_o^k = \sum_{i=1}^{k} l_i (\cdot, u_i) u_i \ + \ l_{k+1} I \ - l_{k+1} \sum_{i=1}^{k} (\cdot, u_i) u_i \quad ,$$

so that an intermediate problem can be hypothesized as:

$$A^k = A_o^k + B < A_0 + B = A \,.$$

To prove this we construct a complete orthonormal set $\{u_1, u_2, ..., u_k, v_{k+1}, v_{k+2},...\}$ using the Gram-Schmidt procedure on the first $k$ eigenfunctions of $A_0$. Then for any function $u \in \mathbf{H}$ of unit norm:

$$(A_0 u,u) = \left( A_o \sum_{i=1}^{k} c_i u_i + A_o \sum_{i=k+1}^{\infty} c_i v_i \ , \ \sum_{i=1}^{k} c_i u_i + \sum_{i=k+1}^{\infty} c_i v_i \right)$$

$$= \left(A_o \sum_{i=1}^{k} c_i u_i , \sum_{i=1}^{k} c_i u_i\right) + \left(A_o \sum_{i=1}^{k} c_i u_i , \sum_{i=k+1}^{\infty} c_i v_i\right) + \left(A_o \sum_{i=k+1}^{\infty} c_i v_i , \sum_{i=1}^{k} c_i u_i\right) + \left(A_o \sum_{i=k+1}^{\infty} c_i v_i , \sum_{i=k+1}^{\infty} c_i v_i\right)$$

$$= \sum_{i=1}^{k} l_i |c_i|^2 + \left(A_o \sum_{i=k+1}^{\infty} c_i v_i , \sum_{i=k+1}^{\infty} c_i v_i\right) \geq \sum_{i=1}^{k} l_i |c_i|^2 + l_{k+1}\left(1 - \sum_{i=1}^{k} |c_i|^2\right) = (A_o^k u, u) \ .$$

The inequality comes from Weber's extension of Rayleigh's principle, i.e. the Span[$v_{k+1}$, $v_{k+2}$,...] is orthogonal to Span[$u_1$, $u_2$, ..., $u_k$] so that min $R(v) = \lambda_{k+1}$, where the minimum is taken only over functions in Span[$v_{k+1}$, $v_{k+2}$,...].

Thus $A_o^k + B$ is indeed an intermediate problem. By itself, however, it is not very useful since $B$ is usually difficult to handle.

### 2.3.4 Miller's method

Starting with the truncated operator,

$$A^k = A_o^k + B \ ,$$

defined in section 2.2.3., as an intermediate problem, we set up the eigenvalue problem.

$$A^k \psi = \sum_{i=1}^{k} l_i (y, u_i) u_i + l_{k+1} y - l_{k+1} \sum_{i=1}^{k} (y, u_i) u_i + B\psi = \Lambda\psi$$

or

$$(l_{k+1} - \Lambda + B) y = \sum_{i=1}^{k} (l_{k+1} - l_i)(y, u_i) u_i$$

If $B$ is a multiplicative operator (as is the case for the electron repulsion perturbation), then it is obvious that

$$y = \sum_{n=1}^{k} \frac{c_n u_n}{l_{k+1} - \Lambda + B} \ .$$

Substituting this form of $\psi$ into the previous equation gives $k$ simultaneous equations:

$$c_1 = [\, l^0_{k+1} - l^0_1 ]\sum_{j=1}^{k}\Big(\frac{c_j}{l^0_{k+1} - \Lambda + B} u_j , u_1\Big)$$

$$c_2 = [\, l^0_{k+1} - l^0_2 ]\sum_{j=1}^{k}\Big(\frac{c_j}{l^0_{k+1} - \Lambda + B} u_j , u_2\Big)$$

$$\ldots$$

$$c_k = [\, l^0_{k+1} - l^0_k ]\sum_{j=1}^{k}\Big(\frac{c_j}{l^0_{k+1} - \Lambda + B} u_j , u_k\Big)$$

This gives rise to the linear algebra problem of finding the coefficients. The resulting matrix problem of the form $Mc=0$ will have a solution only when $\Lambda$ is an eigenvalue of $A_k$. The resulting eigenvalues $\Lambda_i$ ($i=1,...,k$) will then be lower bounds to the first $k$ eigenvalues of the untruncated operator.

Miller's procedure allows one to avoid the complicated integrals involving the square of the operator; however, in place of these, one must calculate integrals of basically the inverse of the operator plus a constant. For chemical applications, these integrals turn out to be over the entire space since the inverse of $B = \sum'' 1/r_{ij}$ is an $N$-electron operator even though $B$ can be decomposed into $N(N\text{-}1)/2$ two-electron operators.

### 2.3.5 Truncation and arbitrary choice

Truncation by itself is certainly not useful, and the integrals needed in Miller's modification of truncation are generally much too difficult. Bazley and Fox proposed combining truncation of the base problem with the Aronszajn method [BF].

We start with the truncated and projected operator, $A^{k,m}$ , defined as

$$A^{k,m} = A_o^k + BP_{\mathbf{K}} ,$$

where $A_o^k$ is the truncated base problem and $P_{\mathbf{K}}$ is a projection operator of dimension $m$ in the inner product $((\cdot,\cdot)) = (B\cdot,\cdot)$. We set up the eigenvalue problem as we did in section 2.2.4. in order to determine the form of the eigenfunctions.

$$A^{k,m}\psi = \sum_{i=1}^{k} l_i (y, u_i) u_i + l_{k+1} y - l_{k+1}\sum_{i=1}^{k}(y, u_i)u_i + BP_{\mathbf{K}}\psi = \Lambda\psi$$

which gives

$$y = \frac{\sum_{i=1}^{k} ( l_{k+1} - l_{i} )( y , u_{i} ) u_{i} + BP_{\mathbf{K}}\, y}{( l_{k+1} - \Lambda )}$$

As was done for Miller's method, we can again set up a linear algebra problem to solve for the eigenvalues. In this case, however, we have the option of setting up a finite-dimensional matrix eigenvalue problem since we know the eigenfunction lies in the subspace $\mathbf{S}$ = Span[$u_1, u_2, ..., u_k$, $B\mathbf{K}$], where the dimension of $\mathbf{S}$ is less than or equal to $k + m$. The truncation of the base problem frees us from needing to use Bazley's special choice to generate the projection operator $P_{\mathbf{K}}$.

## 2.3.6 The relative Lehmann method

The relative Lehmann method has only recently been presented [BG]. Originally introduced as the "eigenvector-free method", it was named such because only the eigenvalues of a base problem are needed, unlike the previous methods which required knowledge of the eigenfunctions.

Given an operator $A = A_o + B$ (with ordered eigenvalues $\{ l_n \}$) where $A_o$ (with ordered eigenvalues $\{ l^{o}_{n} \}$) is the base problem and $B$ is a positive-definite perturbation, one must first choose two finite-dimensional spaces:

$$\mathbf{Q} \subset \mathrm{Dom}(A_o) \text{ and } \mathbf{P} \subset \mathrm{Dom}(B) .$$

One then sets up the following finite-dimensional matrix eigenvalue problem:

$$[F]x_n = \zeta_n [G] x_n ,$$

$$[F] = \begin{array}{|c|c|} \hline \mathbf{Q}\times\mathbf{Q} & \mathbf{Q}\times\mathbf{P} \\ A_o - \mu & 0 \\ \hline \mathbf{P}\times\mathbf{Q} & \mathbf{P}\times\mathbf{P} \\ 0 & B \\ \hline \end{array} \qquad [G] = \begin{array}{|c|c|} \hline \mathbf{Q}\times\mathbf{Q} & \mathbf{Q}\times\mathbf{P} \\ (A_o - \mu)^2 & B(A_o - \mu) \\ \hline \mathbf{P}\times\mathbf{Q} & \mathbf{P}\times\mathbf{P} \\ (A_o - \mu)B & B^2 \\ \hline \end{array}$$

where the constant $\mu \leq l_r^o$ for some known $r$. If this eigenvalue problem yields $k$ negative eigenvalues,

$$\zeta_1 \leq \zeta_2 \leq \ldots \leq \zeta_k < 0 ,$$

then $k$ lower bounds to the eigenvalues of $A$ can be formed as

$$\mu + \frac{1}{V_p} \leq \lambda_{r-p} \quad \text{where } p \leq k .$$

### 2.3.6 Weinberger's method

Weinberger's method [Wb,BG] is particularly interesting because it shares a similariy with the method of arbitrary choice without truncation (which we introduce in chapter three). Given a self-adjoint operator $A$, Weinberger assumes the knowledge of a finite dimensional subspace $\mathbf{P}$ (so that the Hilbert space can be written as $\mathbf{H}=\mathbf{P}\oplus\mathbf{P}^{\perp}$) outside of which we have:

$$(Au,u) \geq \mu \quad \forall\, u \in \mathbf{P}^{\perp}.$$

We now proceed similarly to the relative Lehmann method choosing two finite-dimensional spaces, one of which, $\mathbf{P}$, is already defined:

$$\mathbf{Q} \subset \mathrm{Dom}(A) \text{ and } \mathbf{P}.$$

One then sets up the following finite-dimensional matrix eigenvalue problem:

$$[F]x_n = \zeta_n [G] x_n ,$$

$$[F] = \begin{array}{|c|c|} \hline \mathbf{Q}\times\mathbf{Q} & \mathbf{Q}\times\mathbf{P} \\ A-\mu & "I" \\ \hline \mathbf{P}\times\mathbf{Q} & \mathbf{P}\times\mathbf{P} \\ "I" & 0 \\ \hline \end{array} \qquad [G] = \begin{array}{|c|c|} \hline \mathbf{Q}\times\mathbf{Q} & \mathbf{Q}\times\mathbf{P} \\ (A-\mu)^2 & (A-\mu) \\ \hline \mathbf{P}\times\mathbf{Q} & \mathbf{P}\times\mathbf{P} \\ (A-\mu) & "I" \\ \hline \end{array}$$

where we distinguish "$I$" as the identity *operator* instead of the identity *matrix*, so that the submatrix resulting from "$I$" may be dense.

If this eigenvalue problem yields *k* negative eigenvalues,

$$\zeta_1 \leq \zeta_2 \leq \ldots \leq \zeta_k < 0 ,$$

then *k* lower bounds to the eigenvalues of *A* can be formed as

$$\mu + \frac{1}{V_p} \leq \lambda_{1+d+k-r-p}$$

where $p \leq k$ , $d$ = Dim(**P**), and $r$ = Rank("*I*").

## 2.4 Partitioning and reduced operators

### 2.4.1 Invariance of expectation values

Solutions of the Schrodinger equation for atomic and molecular systems must be antisymmetric in the electron coordinates. In the max-min and min-max characterizations of eigenvalues, one thus need only take the Rayleigh quotient of trial functions that are antisymmetric (we neglect the nuclear coordinates - assume the Born-Oppenheimer approximation). Since each electron is indistinguishable, the expectation value of a given operator $A(i)$ (acting on the $i$-th electron) is equivalent to the expectation value of $A(j)$ for any $j=1,..,N$ where $N$ is the number of electrons, i.e. $(A(i)\psi,\psi) = (A(j)\psi,\psi)$ provided $\psi$ is antisymmetric.

This fact allows us to partition the electronic Hamiltonian in almost arbitrary fashion and the expectation value will remain constant if the wavefunction $\psi$ is antisymmetric [CM]:

$$( [\sum a_i T(i) + \sum b_i V(i) + \sum{}'' c_{i,j} R(i,j)] \psi , \psi ) \text{ is constant}$$

$$\text{as long as } \sum a_i = \sum b_i = N , \sum{}'' c_{i,j} = N(N-1)/2$$

## 2.4.2 Reduced operators

The invariance of the expectation value with regard to partitioning is useless if we stay in the given antisymmetric $N$-particle space of the original Hamiltonian. If we denote such a space by $\mathbf{H}^{N,N}$ where the first superscript denotes the particle dimension of the space and the second superscript denotes that space is antisymmetric in the first $N$ particles, then we have the following:

$$\mathbf{H}^{N,N} \subset \mathbf{H}^{N,N\text{-}1} = \mathbf{H}^{N\text{-}1,N\text{-}1} \otimes \mathbf{H}^{1,1}$$

If we partition an $N$-particle Hamiltonian $H^N$, so that we can write it as $R^{N\text{-}1}\otimes I$, we have the following inequality:

$$E_1 = \min_{\substack{y \in Dom(A) \\ \|y\|=1}} (H^N\psi,\psi) = \min (R^{N\text{-}1}\otimes I\ \psi,\psi)$$

$$\geq \min (R^{N\text{-}1}\otimes I\ \phi,\phi) = \min (R^{N\text{-}1}\gamma,\gamma) = \lambda_1$$

$$\text{where } \psi \in \mathbf{H}^{N,N}, \quad \phi \in \mathbf{H}^{N,N\text{-}1}, \quad \text{and } \gamma \in \mathbf{H}^{N\text{-}1,N\text{-}1}$$

The first inequality results merely from $\mathbf{H}^{N,N} \subset \mathbf{H}^{N,N\text{-}1}$. The last equality is discussed in detail in [MSB and appendix F]. This shows that ground state of a reduced Hamiltonian, $R^{N\text{-}1}$, is a lower bound to the ground state of the original Hamiltonian, $H^N$. The benefit of this comparison is that the reduced Hamiltonian operates in a smaller, and thus simpler space. One may then use one of the previously described techniques to get a lower bound to the reduced Hamiltonian ground state, which will then also be a lower bound to the original Hamiltonian ground state.

For the original presentation for constructing reduced Hamiltonians the reader should refer to [CM]. An application to the ground state of lithium is given in [MSB and appendix F] where there is also given an extension of this method to excited states which also allows for an improvement of the lower bound to the ground state. This is achieved by splitting the original operator $H$ into two parts - one with interparticle interactions, $H_A$, and one without, $H_B$. The operator $H_A$ is then reduced to a simpler operator while $H_B$ is not. The eigenvalues of $H_B$ and the reduced $H_A$ are combined using Weyl's inequality to obtain lower bounds to the eigenvalues of the original $H$.

## 2.5 The effective field method

An excitingly different method of obtaining lower bounds, called the effective field method, was introduced in the 70's by Hertel, Lieb, and Thirring [HLT]. For an atomic Hamiltonian, $H = H_0 + B$, the idea is to form a simple operator, $S$, lower than the electron-electron repulsion operator, $B$. This operator, $S$, is constructed as a sum of single-particle operators and a constant. One can then solve $H_0 + S$ very accurately since it reduces to $N$ one-electron equations, where $N$ is the number of electrons in the atom. This will give lower bounds to the atomic energies since

$$H_0 + S \;<\; H_0 + B \;.$$

Unfortunately, the success of this method is very limited since it seems that $S$ cannot get very close to $B$ and still retain its simple form. The relationship of the effective field method to Weyl's inequality is examined in [MBS and appendix H] where it is revealed that a slightly different formulation allows one to remove the restriction that $S$ be less than $B$ and also shows that lower bounds by the effective field method suffer from an inherant limitation. Elsewhere, a special form for S was constructed that gave an exactly soluble intermediate problem [GR].

## 2.6 Barta/Duffin bounds

In 1947 R. J. Duffin [Df] generalized for the Schrodinger equation a lower bound formula introduced by Barta [Br] ten years earlier:

$$E_1 \;\geq\; \inf_x \frac{Hy(x)}{y(x)} \;,$$

where $E_1$ is the lowest eigenvalue and $x$ represents all the coordinates. Note that this is an infimum over the coordinates of the function and not an infimum over all functions. This formula is not well known. Recently an application to helium has been reported [PFK]. We present a proof of this formula and a similar upper bound formula based on Thirring's proof [Th].

The non-degenerate ground state eigenfunction, $\psi_1$, is known to have no nodes, i.e. its sign is constant [CH]. Thus we can restrict our trial functions, $\psi$, to those of constant sign. Defining $E[x] = H\psi(x)/\psi(x)$ we have:

$$E_1(\psi_1,\psi) \;=\; (H\psi_1,\psi) \;=(\psi_1,H\psi) \;=\; (\psi_1,\psi E[x]) \;.$$

If $(\psi_1,\psi) > 0$, then $\psi_1(x)\psi(x) > 0$ for all $x$ since both $\psi_1$ and $\psi$ have constant sign, thus:

$$E_1(\psi_1,\psi) = (\psi_1,\psi E[x]) \geq (\psi_1,\psi) \inf_x E[x]$$

and

$$E_1(\psi_1,\psi) = (\psi_1,\psi E[x]) \leq (\psi_1,\psi) \sup_x E[x] .$$

The integral $(\psi_1,\psi)$ can never be zero since $\psi$ is restricted to have no nodes. Thus we may divide by it and get the bounds:

$$\sup_x E[x] \geq E_1 \geq \inf_x E[x] .$$

If $(\psi_1,\psi) < 0$, we get the same result.

These two formulas are advantageous because they do not require any integration; however, they are difficult to work with because there is no way to systematically improve the bounds. Also only the ground state can be bounded in this manner.

## 2.7 Lower bound methods for molecules

The united atom has been used as a starting point in perturbation calculations on molecules. Lee and Das [LD] used the neon atom as a starting point in a many-body perturbation procedure to calculate energies for $H_2O$, $NH_3$ and $CH_4$ (all isoelectronic with neon). They let the united atom centered at the heavy atom be the zero-order Hamiltonian and let the perturbation be the hydrogen potentials and the potential difference between the heavy and united atoms. This separation of the base problem and perturbation is not suitable for intermediate lower bound problems since the perturbation is positive definite.

As early as 1963, Bazley and Fox proposed an intermediate problem for molecules treating the electron-electron repulsion as the perturbation and the sum of truncated atomic Hamiltonians, similar to $H_{\text{Atom}}$ in Appendix A, as the base problem. No calculations were done until recently when Seto and Stankevich performed lower bound calculations for $H_2^+$ [SS].

Using $H_2^+$ as a specific example, Bazley and Fox also proposed the intermediate problem:

$$H > H_o^k + (H - H_o^k)P_{\mathbf{K}} \ .$$

where $H_o^k$ is the *k*-th order truncation of $H_O$ and $P_{\mathbf{K}}$ a projection operator onto the subspace **K**. Seto and Stankevich also performed calculations on $H_2^+$ using a variant of this intermediate problem [SS].

We should note that for the molecular hydrogen ion, $H_2^+$, the united atom theorem gives a lower bound spectrum,

$$\{-2.00, -1.25, ..., -1.00\}\text{a.u.},$$

that overlaps the lowest electronic eigenvalue, -1.103 a.u. [W], of the true molecule. Thus we can use the generalized method of section 2.2.3 to obtain an arbitrarily good lower bound to the lowest molecular eigenvalue. This seems unlikely to be applicable to other molecules; indeed it fails for the simple hydrogen molecule, $H_2$. In this case the united atom theorem will give a lower bound spectrum equivalent to that of the helium atom,

$$\{-2.9037, ..., -2.000\} \text{ a.u.} \ .$$

Large variational calculations for $H_2$ show that the electronic ground state is -1.8889 a.u. [KW] and is embedded in the continuous spectrum of the united atom rendering this method useless.

## 2.8 Summary

The following table summarizes the lower bound methods discussed in this chapter with respect to upper bound variational calculations. It lists the *a priori* information that would not be needed in upper bound calculations and also difficulties that exceed those found in upper bound calculations.

**Table 2.1** Summary of lower bound methods discussed in chapter two.

| Method | Goal | *A priori* information | Difficulties |
|---|---|---|---|
| **Temple** | lower bound to $\lambda_1$ | lower bound to $\lambda_2$ | calculation of $[(A^2\psi_m,\psi_n)]$ |
| **Weinstein** | lower bound to $\lambda_n$ | lower bound to $\lambda_{n+1}$<br>complete set of eigenfunctions | calculation of $[(A^2\psi_m,\psi_n)]$ |
| **Generalized** | *n*-1 lower bound(s) | lower bound to $\lambda_n$ | calculation of $[(A^2\psi_m,\psi_n)]$ |
| **Bazley** | lower bound(s) | base problem eigenvectors | inversion of the matrix $[(A\psi_m,\psi_n)]$ |
| **Miller** | lower bound(s) | base problem eigenvectors | calculation of $[(\{cI+A\}^{-1}\psi_m,\psi_n)]$ |
| **Truncation & arbitrary choice** | lower bound(s) | base problem eigenvectors | inversion of the matrix $[(A\psi_m,\psi_n)]$ |
| **Relative-Lehmann** | lower bound(s) | base problem eigen*values* | calculation of $[(A^2\psi_m,\psi_n)]$ |
| **Weinberger** | lower bound(s) | lower bound to lowest eigenvalue on an infinite-dim. subspace | calculation of $[(A^2\psi_m,\psi_n)]$ |
| **Effective Field** | comparison operator | none | must be coupled with another method |
| **Reduced Operator** | comparison operator | none | must be coupled with another method |
| **Barta/Duffin Bound** | lower bound to $\lambda_1$ | none | no method of systematic improvement |

# Chapter Three

# Arbitrary choice without truncation

## 3.1 Limitations of special choice

As wonderful as Bazley's special choice is, it has its drawbacks:

1. The major disappointment is that one needs a base problem whose eigenvalues and eigenfunctions are exactly known. For atomic systems, the eigenvalues and eigenfunctions of the base problem $H_0 = T + V$ are easily obtainable from the hydrogenic eigenvalues and eigenfunctions. For molecular systems, however, the obvious base problem (the neglect of electron-electron repulsion) is not exactly soluble, and thus lower bounds to the vast majority of chemical systems of interest are inaccessible by Bazley's special choice.

2. Another drawback is that one is restricted to using the base problem's eigenfunctions in creating the intermediate problem. Since in general these eigenfunctions form an incomplete set, we may not be able to form a sequence of intermediate problems whose eigenvalues converge to the true eigenvalues [Rd2] (the reference actually refers to an extension of Bazley's special choice method designed to eliminate drawback #3 listed below).

3. Bazley's intermediate problems do not raise the essential spectrum of the base problem. For the lithium atom, the essential spectrum of the base problem (neglect of repulsion) is at -9 hartree. Since the ground state of lithium is -7.48 hartree, we will always get very poor lower bounds. The reason that the essential spectrum is not raised is that Bazley's intermediate problems have a spectrum identical to that of the base problem, except for a finite number $M$ of eigenvalues. $M$ is also equal to the number of base problem eigenvectors used in creating $P_{B^{-1}\mathbf{M}}$. If only $M$ eigenvalues differ between the two spectrums, then there are still an infinite number of eigenvalues below the essential spectrum of the base problem (even if the $M$ eigenvalues are themselves actually raised into the essential spectrum). This is not a problem for two-electron systems since the essential spectrums of the base and true problems are the same. For systems with more electrons (e.g. lithium), however, this is a serious problem. Fortunately, this was problem was solved early on [F,Rd1]. Our research will deal only with two-electron systems, so we will not detail this solution; however, we do note that our generalization of Bazley's method has the same problem with the essential spectrum. Luckily, it also shares the same solution.

## 3.2 Generalization to arbitrary choice

### 3.2.1 Goals

We develop a generalization of Bazley's special choice, by removing the restriction that the subspace $\mathbf{M}$ of the projection operator $P_{B^{-1}\mathbf{M}}$ be formed from linear combinations of eigenfunctions of the base problem. Our goals in doing so are two-fold (basically, to solve problems #1 and #2 in section 3.1):

1. Remove the requirement that the eigenfunctions of the base problem be known thereby admitting molecular systems to lower bound study.
2. Increase the options for $P_{B^{-1}\mathbf{M}}$ to allow faster convergence of the lower bound to the desired eigenvalue, or to get convergence where there was none before.

### 3.2.2 Kahan-Weinberger bounds

Given an operator on an infinite-dimensional space, one can generally obtain rigorous upper bounds to eigenvalues using the variation method. Kahan [Pr] and Weinberger [Wb] introduced a method for obtaining lower bounds that results directly from Weyl's' inequality. These bounds are calculated from a finite-dimensional matrix eigenvalue problem that incorporates some *a priori* information about the operator on an infinite-dimensional *subspace* of Hilbert space.

We divide Hilbert space into three parts: $\mathbf{H}=\mathbf{M}\oplus\mathbf{R}\oplus\infty$**,** where **M** and **R** are of finite dimension and $\infty$ is infinite-dimensional. We assume that the operator admits an infinite-dimensional matrix decomposition of the following form (where to avoid clutter, we have represented zero-submatrices by points " · "):

$$\begin{array}{|c|c|c|}\hline \begin{array}{c}\mathbf{M}\times\mathbf{M}\\ A=A|_{\mathbf{M}}\end{array} & \begin{array}{c}\mathbf{M}\times\mathbf{R}\\ B\end{array} & \begin{array}{c}\mathbf{M}\times\infty\\ \cdot\end{array}\\ \hline \begin{array}{c}\mathbf{R}\times\mathbf{M}\\ B^*\end{array} & \begin{array}{c}\mathbf{R}\times\mathbf{R}\\ C=A|_{\mathbf{R}}\end{array} & \begin{array}{c}\mathbf{R}\times\infty\\ D\end{array}\\ \hline \begin{array}{c}\infty\times\mathbf{M}\\ \cdot\end{array} & \begin{array}{c}\infty\times\mathbf{R}\\ D^*\end{array} & \begin{array}{c}\infty\times\infty\\ E=A|_{\infty}\end{array}\\ \hline \end{array}$$

The submatrices on the diagonal represent restrictions of the operator $A$ to the different subspaces. The matrix need not be Hermitian or infinite-dimensional as a whole for the lower bound method of Kahan and Weinberger; however, in our applications they will be. The zero-submatrices in the upper-right and lower-left are necessary.

For any constant, $\lambda$, we can write the matrix above as a sum of two matrices:

$$\begin{array}{|c|c|c|}\hline \begin{array}{c}\mathbf{M}\times\mathbf{M}\\ A\end{array} & \begin{array}{c}\mathbf{M}\times\mathbf{R}\\ B\end{array} & \begin{array}{c}\mathbf{M}\times\infty\\ \cdot\end{array}\\ \hline \begin{array}{c}\mathbf{R}\times\mathbf{M}\\ B^*\end{array} & \begin{array}{c}\mathbf{R}\times\mathbf{R}\\ \lambda I\end{array} & \begin{array}{c}\mathbf{R}\times\infty\\ \cdot\end{array}\\ \hline \begin{array}{c}\infty\times\mathbf{M}\\ \cdot\end{array} & \begin{array}{c}\infty\times\mathbf{R}\\ \cdot\end{array} & \begin{array}{c}\infty\times\infty\\ \lambda I\end{array}\\ \hline \end{array} + \begin{array}{|c|c|c|}\hline \begin{array}{c}\mathbf{M}\times\mathbf{M}\\ \cdot\end{array} & \begin{array}{c}\mathbf{M}\times\mathbf{R}\\ \cdot\end{array} & \begin{array}{c}\mathbf{M}\times\infty\\ \cdot\end{array}\\ \hline \begin{array}{c}\mathbf{R}\times\mathbf{M}\\ \cdot\end{array} & \begin{array}{c}\mathbf{R}\times\mathbf{R}\\ C-\lambda I\end{array} & \begin{array}{c}\mathbf{R}\times\infty\\ D\end{array}\\ \hline \begin{array}{c}\infty\times\mathbf{M}\\ \cdot\end{array} & \begin{array}{c}\infty\times\mathbf{R}\\ D^*\end{array} & \begin{array}{c}\infty\times\infty\\ E-\lambda I\end{array}\\ \hline \end{array}$$

If $\lambda$ is a lower bound to the lowest eigenvalue of the operator $A$ on the subspace $\mathbf{R}\oplus\infty$, i.e. $\lambda\leq A\big|_{\mathbf{M}^{\perp}=\mathbf{R}\oplus\infty}$, then the second matrix is positive semi-definite. We can thus ignore it in Weyl's formula and the eigenvalues of the first matrix, by itself, will give lower bounds to the original matrix. Note the similarity of the logic used here to that in Temple's formula: in Temple's formula we assumed a lower bound to the second eigenvalue in order to get a lower bound to the first; here we are assuming a lower bound to $A$ in a subspace to get a lower bound on $A$ in the whole space. It is not necessary that $\lambda$ be a lower bound; however, it seems that the greatest lower bound is the optimal choice.

Another option that avoids *explicit* introduction of a parameter $\lambda$ is to split the matrix as follows:

$$\begin{array}{|c|c|c|}\hline \begin{matrix}\mathbf{M}\times\mathbf{M}\\ A\end{matrix} & \begin{matrix}\mathbf{M}\times\mathbf{R}\\ .\end{matrix} & \begin{matrix}\mathbf{M}\times\infty\\ .\end{matrix} \\ \hline \begin{matrix}\mathbf{R}\times\mathbf{M}\\ .\end{matrix} & \begin{matrix}\mathbf{R}\times\mathbf{R}\\ C\end{matrix} & \begin{matrix}\mathbf{R}\times\infty\\ D\end{matrix} \\ \hline \begin{matrix}\infty\times\mathbf{M}\\ .\end{matrix} & \begin{matrix}\infty\times\mathbf{R}\\ D^*\end{matrix} & \begin{matrix}\infty\times\infty\\ E\end{matrix} \\ \hline \end{array} + \begin{array}{|c|c|c|}\hline \begin{matrix}\mathbf{M}\times\mathbf{M}\\ .\end{matrix} & \begin{matrix}\mathbf{M}\times\mathbf{R}\\ B\end{matrix} & \begin{matrix}\mathbf{M}\times\infty\\ .\end{matrix} \\ \hline \begin{matrix}\mathbf{R}\times\mathbf{M}\\ B^*\end{matrix} & \begin{matrix}\mathbf{R}\times\mathbf{R}\\ .\end{matrix} & \begin{matrix}\mathbf{R}\times\infty\\ .\end{matrix} \\ \hline \begin{matrix}\infty\times\mathbf{M}\\ .\end{matrix} & \begin{matrix}\infty\times\mathbf{R}\\ .\end{matrix} & \begin{matrix}\infty\times\infty\\ .\end{matrix} \\ \hline \end{array}$$

In this case, knowledge of the lower bound $\lambda\leq A\big|_{\mathbf{M}^{\perp}=\mathbf{R}\oplus\infty}$ is still necessary in order to evaluate the eigenvalues of the left matrix, but we do not use it further. Computationally, this decomposition yielded worse results so we will only use the first decomposition from now on.

### 3.2.3 Optimizing the bounds

In order to calculate lower bounds to the intermediate problem we need to first bound the lowest eigenvalue of:

$$\begin{array}{|c|c|c|}\hline \begin{matrix}\mathbf{M}\times\mathbf{M}\\ .\end{matrix} & \begin{matrix}\mathbf{M}\times\mathbf{R}\\ .\end{matrix} & \begin{matrix}\mathbf{M}\times\infty\\ .\end{matrix} \\ \hline \begin{matrix}\mathbf{R}\times\mathbf{M}\\ .\end{matrix} & \begin{matrix}\mathbf{R}\times\mathbf{R}\\ C\end{matrix} & \begin{matrix}\mathbf{R}\times\infty\\ D\end{matrix} \\ \hline \begin{matrix}\infty\times\mathbf{M}\\ .\end{matrix} & \begin{matrix}\infty\times\mathbf{R}\\ D^*\end{matrix} & \begin{matrix}\infty\times\infty\\ E\end{matrix} \\ \hline \end{array}$$

Using Weyl's inequality, a lower bound will naturally be the min[0,μ] where μ is the lowest eigenvalue of the submatrix:

| | |
|---|---|
| **R**×**R**<br>$C$ | **R**×∞<br>$D$ |
| ∞×**R**<br>$D^*$ | ∞×∞<br>$E$ |

Writing the subspace **R**⊕∞ as $\mathbf{M}^\perp$, we realize that μ can also be defined as:

$$\mu \;=\; \min_{\substack{u\in Dom[A]\cap\mathbf{M}^\perp \\ \|u\|=1}} (Au,u)\ .$$

Using the max-min principle we find an upper bound for μ:

$$\Lambda_{n+1} \;=\; \max_{\substack{\mathbf{M}\subset\mathbf{H} \\ \dim(\mathbf{M})=n}} \;\min_{\substack{u\in Dom[A]\cap\mathbf{M}^\perp \\ \|u\|=1}} (Au,u) \;\geq\; \min_{\substack{u\in Dom[A]\cap\mathbf{M}^\perp \\ \|u\|=1}} (Au,u) \;=\; \mu\ ,$$

where $\Lambda_{n+1}$ is the ($n$+1)-th eigenvalue of $A$ in the normal Hilbert space.

Using min[0,$\Lambda_{n+1}$] as the incorrect (i.e. too high) lower bound removes the rigor from our calculation. However, we are only testing the method of arbitrary choice, and if our method fails to give good results with this "better than best" bound, then it will surely fail for the true bound. As long as the dimension of **M** is finite, we are guaranteed to have min[0,$\Lambda_{n+1}$] = $\Lambda_{n+1}$.

## 3.3 Arbitrary choice for the helium atom

### 3.3.1 Description of the basis set

In Bazley's special choice method, the vectors added to **M** are formed from the eigenfunctions of the base problem. For arbitrary choice, we can add any functions. We choose to add functions from a complete orthonormal set.

The complete orthonormal set that we use is:

$$\sqrt{\tfrac{(2a)^3}{(n_1+1)(n_1+2)}}\, L^2_{n_1}(2ar_1)e^{-ar_1}Y_{l_1,m_1}(q_1,f_1)\times\sqrt{\tfrac{(2a)^3}{(n_2+1)(n_2+2)}}\, L^2_{n_2}(2ar_2)e^{-ar_2}Y_{l_2,m_2}(q_2,f_2)\ ,$$

where $Y_{l,m}(q,f)$ are the orthonormal spherical harmonics and $L^2_n(x)$ are orthonormal associated Laguerre polynomials with the explicit representation:

$$L^2_n(x)\ =\sum_{k=0}^{n}\frac{(n+2)!(-x)^k}{(k+2)!(n-k)!k!}\ \ ,\ n = 0,1,2,3,...\ ,$$

that allows all integrals to be calculated in closed form as sums of the simple one- or two-dimensional integrals (see Appendix C).

For simplicity we choose only radial functions in forming **M**. All integrals in the computations then reduce to radial integrals only. This allows us to neglect the angular variables and use the radially complete basis:

$$\sqrt{\tfrac{(2a)^3}{(n_1+1)(n_1+2)}}\, L^2_{n_1}(2ar_1)e^{-ar_1}\times\sqrt{\tfrac{(2a)^3}{(n_2+1)(n_2+2)}}\, L^2_{n_2}(2ar_2)e^{-ar_2}$$

### 3.3.2 Constructing $P_{B^{-1}\mathbf{M}}$

The rank of an operator is the dimension of its range. If $P_{B^{-1}\mathbf{M}}$ is defined with **M** being a one-dimensional space (i.e. the span of a single vector), then all vectors orthogonal to **M** in the new inner product are annihilated and all other vectors are transformed into their component in **M**. Thus $P_{B^{-1}\mathbf{M}}$ is rank one. We now increase the dimension of **M**, thereby increasing the rank of $P_{B^{-1}\mathbf{M}}$. When $P_{B^{-1}\mathbf{M}}$ is rank one, the form of the projection is very simple and was explicitly shown in section 2.3.2; however, for rank two or higher, the form of the projection is more complex. This is addressed in Appendix B.

For simplicity, higher dimensional spaces **M** are defined by

$$\mathbf{M}(\mathrm{n}) = \mathrm{Span}[\{m_1,m_2,...,m_\mathrm{n}\}\otimes\{m_1,m_2,...,m_\mathrm{n}\}]\ ,$$

where $m_\mathrm{i}$ = $L^2_n(2ar)e^{-ar}$. Thus the dimension of **M**(n) increases as $n^2$.

With these spaces defined, we can determine the "better-than-best" lower bounds to $H_0$ on $\mathbf{M}^{\perp}$.

**Table 3.1** Better-than-best approximations for $\mu$.

| **Energy level #** | **Energy level name** | **$\mu \leq$ ...** |
|---|---|---|
| 1 | 1*S* 1*S* | NA |
| **2** | **1*S* 2*S*** | **-2.500** |
| 3 | 1*S* 2*P* | -2.500 |
| 4 | 1*S* 2*P* | -2.500 |
| **5** | **1*S* 2*P*** | **-2.500** |
| 6 | 1*S* 3*S* | -2.222 |
| 7 | 1*S* 3*P* | -2.222 |
| 8 | 1*S* 3*P* | -2.222 |
| 9 | 1*S* 3*P* | -2.222 |
| **10** | **1*S* 3*D*** | **-2.222** |
| 11 | 1*S* 3*D* | -2.222 |
| 12 | 1*S* 3*D* | -2.222 |
| 13 | 1*S* *3D* | -2.222 |
| 14 | 1*S* 3*D* | -2.222 |
| 15 | 1*S* 4*S* | -2.125 |
| 16 | 1*S* 4*P* | -2.125 |
| **17** | **1*S* 4*P*** | **-2.125** |

Thus for the subspaces **M**(1), **M**(2), **M**(3), and **M**(4) we have $\mu \leq \Lambda_2 = -2.500$, $\mu \leq \Lambda_5 = -2.500$, $\mu \leq \Lambda_{10} = -2.222$, and $\mu \leq \Lambda_{17} = -2.125$, respectively.

### 3.3.3 Determining the variational functions

In this method of arbitrary choice, we obtain lower bounds to the eigenvalues of the intermediate problems on a finite-dimensional subspace. Part of this subspace is **M**. The rest of it of course is a finite-dimensional part of $\mathbf{M}^{\perp}$. In this section we explicitly determine what this part is.

Given a finite-dimensional subspace **M**, we first form the space $H_0\mathbf{M}$. This will be a finite-dimensional subspace of **H** since **M** is finite-dimensional, with dimension less than or equal to the dimension of **M**. We must know this subspace explicitly in order to decompose the block-triangular form desired. We will define the subspace **R** as

$$\mathbf{R} = (H_0\mathbf{M})\backslash\mathbf{M} = P_{\mathbf{M}^\perp}(H_0\mathbf{M}),$$

which means **R** is the space of all functions in $H_0$**M** that are orthogonal to **M**, i.e. the **R** is orthogonal to **M**. We will then have $\mathbf{H}=\mathbf{M}\oplus\mathbf{R}\oplus\infty$. Note that in special choice, $H_0\mathbf{M} = \mathbf{M}$ so that **R** is empty.

To determine **R** we introduce the following notation:

$$[m]\times[n] = \mathrm{Exp}(-ar_1)(m\text{-th degree polynomial in } r_1) \\ \times \mathrm{Exp}(-ar_2)(n\text{-th degree polynomial in } r_2) .$$

where $[m] = \mathrm{Exp}(-ar_1)(m\text{-th degree polynomial in } r_1)$ and likewise for $n$.

We also define $[m']$ and $[-1]$ as

$$[m'] = [m] + \mathrm{Exp}(-ar_1)(\text{a term of } 1/r_1) ,$$

$$[-1] = \mathrm{Exp}[-ar_1](\text{a term of } 1/r_1) ,$$

and likewise for $[n']$.

It is straight-forward to see that the one-electron kinetic energy operator plus the nuclear attraction operator take an $[m]$ into $[m']$. It then follows that

$$H_0([m]\times[n]) = [m']\times[n] + [m]\times[n'] .$$

Our basis set for **M** is the form of $[m]\times[n]$. We now see how each basis function is transformed by $H_0$.

**Table 3.1** Transformation of **M** under $H_O$.

| $[m]\times[n]$ | $H_0([m]\times[n])$ | $P_{\mathbf{M}^\perp} H_0([m]\times[n])$ |
|---|---|---|
| $[0]\times[0]$ | $[0]\times[0'] + [0']\times[0]$ | $[0]\times[-1] + [-1]\times[0]$ |
| $[0]\times[1]$ | $[0]\times[1'] + [0']\times[1]$ | $[0]\times[-1] + [-1]\times[1]$ |
| $[0]\times[2]$ | $[0]\times[2'] + [0']\times[2]$ | $[0]\times[-1] + [-1]\times[2]$ |
| $[0]\times[n]$ | $[0]\times[n'] + [0']\times[n]$ | $[0]\times[-1] + [-1]\times[n]$ |
| $[1]\times[0]$ | $[1]\times[0'] + [1']\times[0]$ | $[1]\times[-1] + [-1]\times[0]$ |
| $[2]\times[0]$ | $[2]\times[0'] + [2']\times[0]$ | $[2]\times[-1] + [-1]\times[0]$ |
| $[m]\times[0]$ | $[m]\times[0'] + [m']\times[0]$ | $[m]\times[-1] + [-1]\times[0]$ |
| $[1]\times[1]$ | $[1]\times[1'] + [1']\times[1]$ | $[1]\times[-1] + [-1]\times[1]$ |
| $[1]\times[2]$ | $[1]\times[2'] + [1']\times[2]$ | $[1]\times[-1] + [-1]\times[2]$ |
| $[2]\times[1]$ | $[2]\times[1'] + [2']\times[1]$ | $[2]\times[-1] + [-1]\times[1]$ |

Using only the functions $\{P_{\mathbf{M}^\perp} H_O([m]\times[0]), P_{\mathbf{M}^\perp} H_O([0]\times[n])\}$ we can use the Gram-Schmidt procedure to generate an orthogonal set consisting of the functions

$$\mathrm{Exp}[-ar_1](1/r_1)\times\mathrm{Exp}[-ar_2]\,(r_2)^k \quad \text{and} \quad \mathrm{Exp}[-ar_1](r_1)^k\times\mathrm{Exp}[-ar_2]\,(1/r_2)\ ,\ k = 0,1,2,...$$

This set can then be used to construct any function of the form $P_{\mathbf{M}^\perp} H_O([m]\times[n])$ where $[m]\neq[0]$ and $[n]\neq[0]$. Thus

$$\mathrm{Span}\{P_{\mathbf{M}^\perp} H_O([m]\times[0]),\ P_{\mathbf{M}^\perp} H_O([0]\times[n])\} \ = \ P_{\mathbf{M}^\perp}(H_O\mathbf{M}) \ = \ \mathbf{R}\ .$$

Thus if use $\mathbf{M}(n)$ of dimension $n^2$, **R** is of dimension $2n$-1, so that we perform a finite-dimensional calculation on a subspace of dimension $n^2+2n-1$.

In the actual calculations we will not Gram-Schmidt orthonormalize the vectors. Instead we will use the $2n$-1 functions $\{P_{\mathbf{M}^\perp} H_O([m]\times[0]),\ P_{\mathbf{M}^\perp} H_O([0]\times[n])\}$ which are not orthogonal nor normalized and invert the overlap matrix. Such a procedure is suitable for small calculations which we will restrict ourselves to; however, for large calculations, matrix inversion is very expensive and can introduce numerical error. In that case it would be advantageous to first Gram-Schmidt orthonormalize the vectors so that the overlap matrix is the identity matrix.

### 3.3.4 Method of solution

We wish to solve an infinite-dimensional intermediate matrix eigenvalue problem for the helium atom. Having chosen a subspace $\mathbf{M}(n)$ of dimension $n^2$ spanned by the $n^2$ orthonormal functions $m_i$, $i=1,...,n^2$, we obtain the subspace $\mathbf{R}$ of dimension $2n-1$ spanned by the non-orthonormal functions $P_{\mathbf{M}^{\perp}} H_0 m_i = r_i$, $i=1,...,2n-1$. The matrix and the overlap matrix are depicted below.

$$H > H_I = \left[\begin{array}{c|c|c} \mathbf{M}\times\mathbf{M} & \mathbf{M}\times\mathbf{R} & \mathbf{M}\times\infty \\ A_1+A_2 & C & \cdot \\ \hline \mathbf{R}\times\mathbf{M} & \mathbf{R}\times\mathbf{R} & \mathbf{R}\times\infty \\ C^T & D & E \\ \hline \infty\times\mathbf{M} & \infty\times\mathbf{R} & \infty\times\infty \\ \cdot & E & F \end{array}\right]$$

where $H_I$ stands for "intermediate Hamiltonian". The overlap matrix is:

$$S = \left[\begin{array}{c|c|c} \mathbf{M}\times\mathbf{M} & \mathbf{M}\times\mathbf{R} & \mathbf{M}\times\infty \\ I & \cdot & \cdot \\ \hline \mathbf{R}\times\mathbf{M} & \mathbf{R}\times\mathbf{R} & \mathbf{R}\times\infty \\ \cdot & J & \cdot \\ \hline \infty\times\mathbf{M} & \infty\times\mathbf{R} & \infty\times\infty \\ \cdot & \cdot & I \end{array}\right]$$

where $I$ is the identity matrix.  This gives the matrix eigenvalue problem:

$$H_I x_n = \lambda_n S x_n \,.$$

We pre-multiply by $S^{-1}$ to obtain

$$S^{-1} H_I x_n = \lambda_n x_n \,.$$

where

$$S^{-1}H_I \;=\; \left[\begin{array}{c|c|c}
\begin{matrix}\mathbf{M}\times\mathbf{M}\\ A_1+A_2\end{matrix} & \begin{matrix}\mathbf{M}\times\mathbf{R}\\ C\end{matrix} & \begin{matrix}\mathbf{M}\times\infty\\ \cdot\end{matrix} \\
\hline
\begin{matrix}\mathbf{R}\times\mathbf{M}\\ K\end{matrix} & \begin{matrix}\mathbf{R}\times\mathbf{R}\\ D'\end{matrix} & \begin{matrix}\mathbf{R}\times\infty\\ E'\end{matrix} \\
\hline
\begin{matrix}\infty\times\mathbf{M}\\ \cdot\end{matrix} & \begin{matrix}\infty\times\mathbf{R}\\ E^{T'}\end{matrix} & \begin{matrix}\infty\times\infty\\ F'\end{matrix}
\end{array}\right]$$

We now replace the submatrix on $\mathbf{M}^{\perp}\times\mathbf{M}^{\perp}$ with $L = (-\Lambda_{n^2+1})I$ to obtain the final matrix eigenvalue problem which will give lower bounds to the intermediate problem.

$$S^{-1}H_{\mathrm{I}} \;=\; \left[\begin{array}{c|c|c}
\begin{matrix}\mathbf{M}\times\mathbf{M}\\ A_1+A_2\end{matrix} & \begin{matrix}\mathbf{M}\times\mathbf{R}\\ C\end{matrix} & \begin{matrix}\mathbf{M}\times\infty\\ \cdot\end{matrix} \\ \hline
\begin{matrix}\mathbf{R}\times\mathbf{M}\\ K\end{matrix} & \begin{matrix}\mathbf{R}\times\mathbf{R}\\ L\end{matrix} & \begin{matrix}\mathbf{R}\times\infty\\ \cdot\end{matrix} \\ \hline
\begin{matrix}\infty\times\mathbf{M}\\ \cdot\end{matrix} & \begin{matrix}\infty\times\mathbf{R}\\ \cdot\end{matrix} & \begin{matrix}\infty\times\infty\\ L\end{matrix}
\end{array}\right]$$

Computationally, of course, we will not work with infinite-dimensional matrices. We work only in the **M⊕R** subspace. The eignevalues of this finite-dimensional sub-matrix,

$$\left[\begin{array}{c|c}
\begin{matrix}\mathbf{M}\times\mathbf{M}\\ A_1+A_2\end{matrix} & \begin{matrix}\mathbf{M}\times\mathbf{R}\\ C\end{matrix} \\ \hline
\begin{matrix}\mathbf{R}\times\mathbf{M}\\ K\end{matrix} & \begin{matrix}\mathbf{R}\times\mathbf{R}\\ L\end{matrix}
\end{array}\right]$$

give lower bounds to the intermediate problem, which in turn give lower bounds to the full problem. We need to form only the following six submatrices:

$$\begin{aligned}
A_1 &= [(H_0 m_i, m_j)] \\
A_2 &= [(B P_{B^{-1}\mathbf{M}} m_i, m_j)] \\
C &= [(H_0 m_i, r_j)] \\
C^{\mathrm{T}} &= [(H_0 r_i, m_j)] \\
J &= [(r_i, r_j)] \\
K &= K(J, C^{\mathrm{T}})
\end{aligned}$$

Formation of theses matrices is described in Appendix D.

## 3.4 Results and discussion

Since we did not employ a rigorous lower bound for $H_0$ on $\mathbf{M}^{\perp}$, all the "lower bounds" we report are not really lower bounds at all; however, the main goal is to test the method of arbitrary choice to see if it is a viable method for calculating lower bounds.

Figure 3.1 shows the results for the one-dimensional $\mathbf{M}(1)$, four-dimensional $\mathbf{M}(2)$, nine-dimensional $\mathbf{M}(3)$, and sixteen-dimensional $\mathbf{M}(4)$, all varied with the adjustable parameter '$a$' in the exponential part of the basis functions. The inverted data point on the plot of $\mathbf{M}(1)$ marks the one true lower bound for the ground state of helium because at this point, $a = 2$, $\mathbf{M}(1)$ is the $1S1S$ function for the repulsionless helium atom, i.e. a special choice function. In this case $\mu = \Lambda_2 = -2.500$ so the lower bound of -3.08 a.u. is rigorous. At points close to $a = 2$, we would expect $\mu$ not to deviate too far from -2.500 so that similar lower bounds would result. Whether or not these would improve the special choice bound we cannot determine without rigorous values for $\mu$.

As the dimension of our matrix eigenvalue problem grows, we find that the lower bounds do not necessarily get better, and for many choices of $a$ the bound actually gets worse. This is very disappointing since it means that convergence to the true lower bound may not ever occur even if we keep increasing the dimension of the problem. This is very much unlike upper bound variational problems where the bounds never worsen and in practice almost always get better. Lack of data points in the plots as the dimension of $\mathbf{M}$ increases is merely because they are off-scale - i.e. too low.

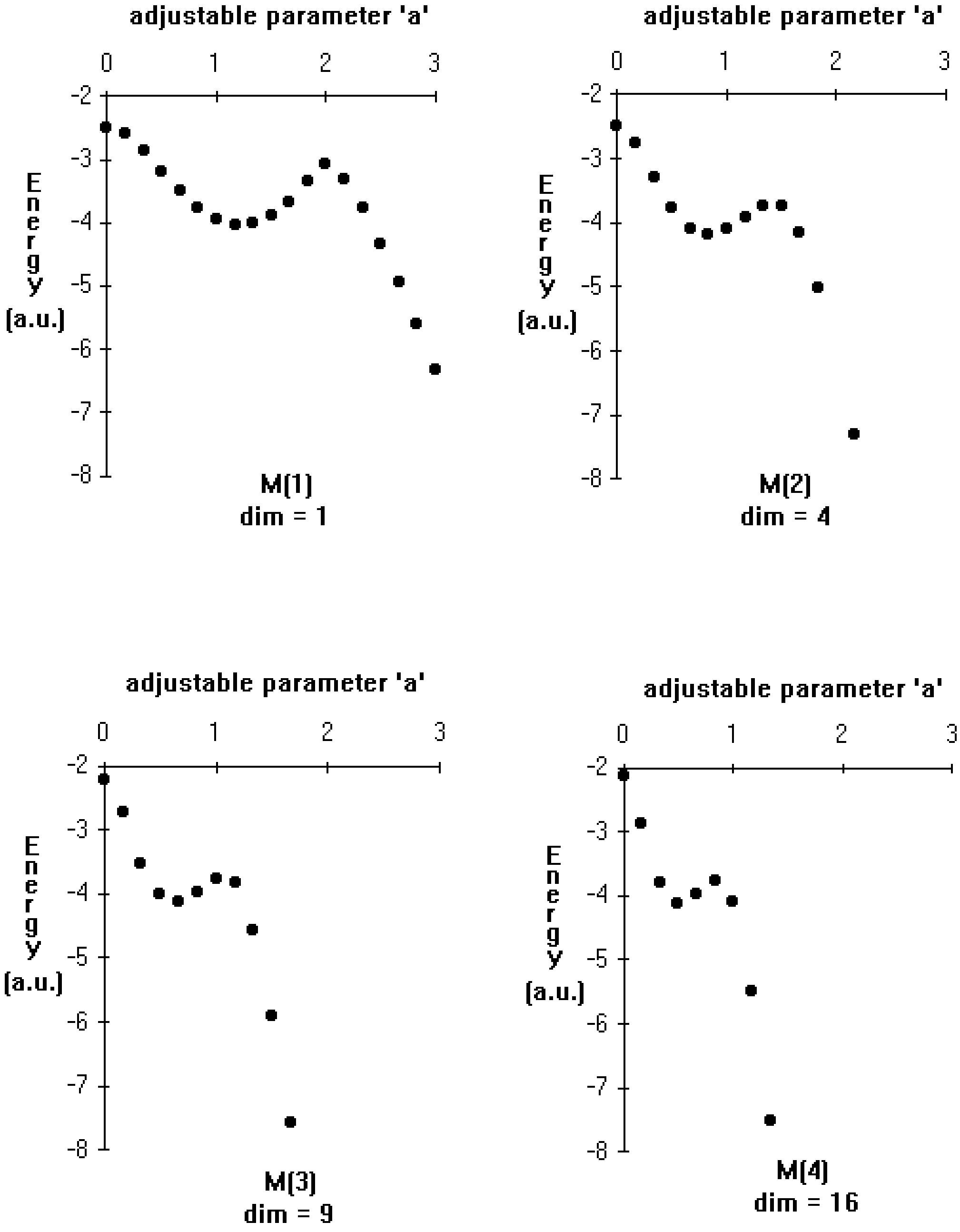

**Figure 3.1** Lower bound to the ground state of the helium atom

The reason that the lower bound in many cases does not improve or worsens when we increase the size of the problem is due to the fact that although the intermediate problem is approaching the true problem, we are not solving the intermediate problem exactly; instead we are calculating lower bounds to the intermediate problem which gives us lower bounds to lower bounds.

We are forced to only calculate lower bounds to the intermediate problem because we have mixing between the finite-dimensional submatrix on **M** and the infinite-dimensional submatrix on $\mathbf{M}^{\perp}$, i.e. the off-block diagonal submatrices are non-zero. If they were zero, we would have special choice and could solve the intermediate problem exactly. Thus, the fact that the off-block diagonal submatrices are non-zero is the problem. One way of estimating how "large" these submatrices are is to solve for the largest magnitude eigenvalue of:

| | | |
|---|---|---|
| **M×M**<br>. | **M×R**<br>$C$ | **M×∞**<br>. |
| **R×M**<br>$K$ | **R×R**<br>. | **R×∞**<br>. |
| **∞×M**<br>. | **∞×R**<br>. | **∞×∞**<br>. |

This value is reported in figure 3.2 for the different choices of **M**. Again, lack of data points in the plots as the dimension of **M** increases is merely because they are off-scale - in this case too high.

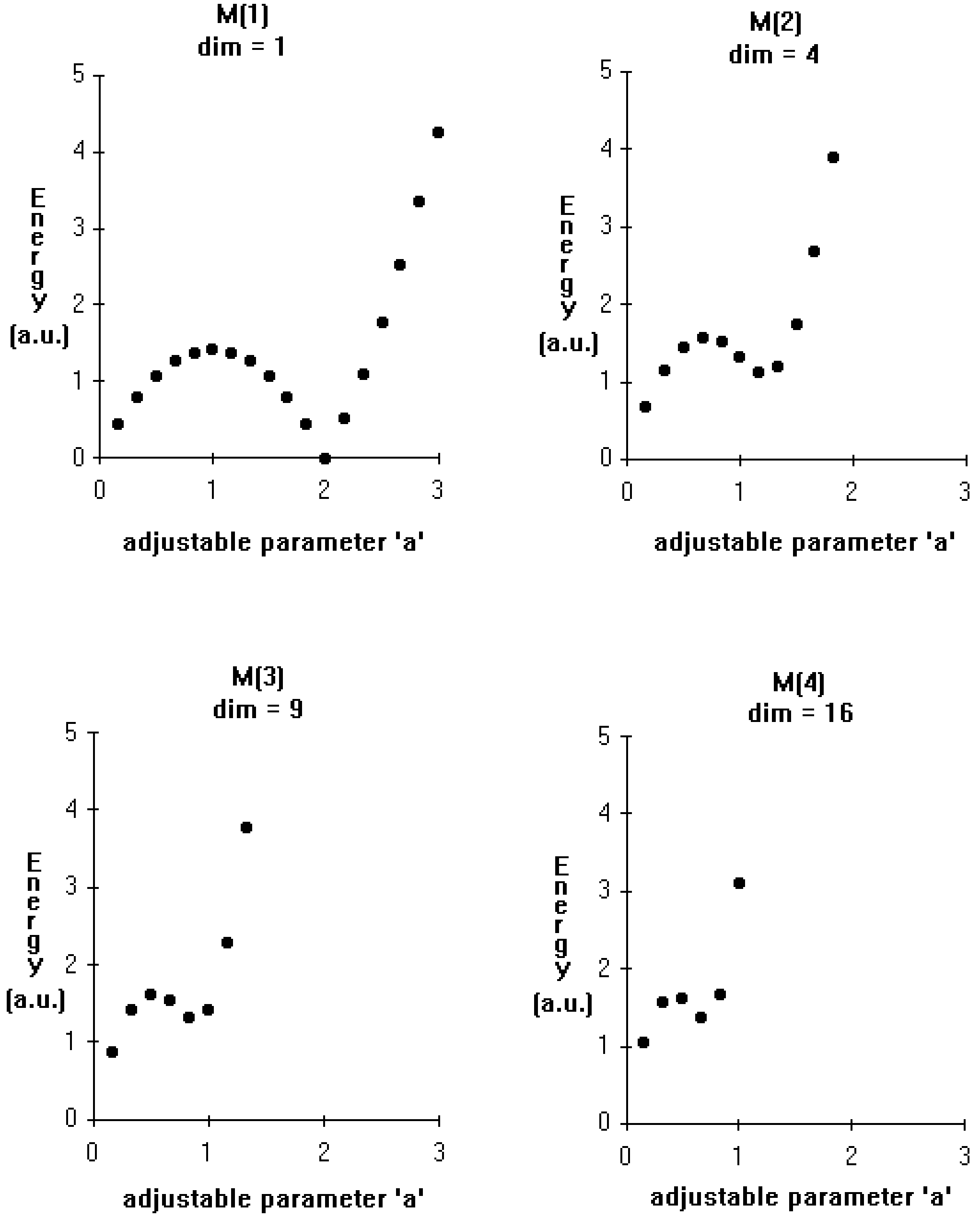

**Figure 3.2** Norms of "mixing" matrix

The only way to get better results is to force the off-block-diagonal submatrices to be as "small" as possible, i.e. close to zero-matrices. This again brings us back to special choice.

At present we have no way to determine a rigorous lower bound to $\mu$ and even if we are given this piece of data we cannot arbitrarily increase the size of the problem to get improved lower bounds. Fortunately there is some good news; that is $\mu$ is the *only* piece of information needed - no knowledge of the base problem's eigenfunctions is needed.

It should be pointed out that given $\mu$ or a lower bound to $\mu$, we also have a lower bound to one of the base problem eigenfunctions (which we used as a better-than-best estimate of $\mu$). This information is all that is needed in some other lower bound methods, e.g. Temple's method and the generalized version of it. Thus arbitrary choice must compete with these methods.

## 3.5 A new approach

### 3.5.1 Review of goals and progress

The main focus of this dissertation is the generation of an intermediate problem very similar to the special choice methods of Bazley and Fox; however, we desire a method that gives us arbitrary choice in making the projection to form $BP_{B^{-1}\mathbf{M}}$.

When we arbitrarily choose **M**, we are no longer able to solve the intermediate problem because of mixing between the finite- and infinite-dimensional matrices. Thus we used the Kahan-Weinberger bounds to approximate the intermediate problem. This approximation gave us a lower bound to the intermediate problem, and thus a lower bound to a lower bound of the full problem, and keeps us from getting guaranteed improvement of the lower bounds with increased computational effort.

Although the current approach seems to have reached a dead end, the goal is still highly desirable and Professor Beattie has suggested to attack the problem from a different angle, e.g. from Weinberger's method. This is discussed below.

### 3.5.2 Arbitrary choice and Weinberger's method

For a given self-adjoint operator $A$, Weinberger assumes the knowledge of a finite-dimensional subspace **P** (so that the Hilbert space can be written as $\mathbf{H}=\mathbf{P}\oplus\mathbf{P}^{\perp}$) outside of which we have:

$$(Au,u) \geq \mu \quad \forall\ u \in \mathbf{P}^{\perp}.$$

This *a priori* information is remarkably similar to what we need for arbitrary choice without truncation: using our usual decomposition of $A$ into a base problem $A_0$ and a positive-definite perturbation $B$, we need knowledge of a finite dimensional subspace **P** such that

$$(A_0 u,u) \geq \mu \quad \forall\ u \in \mathbf{P}^{\perp}.$$

This is exactly the same information that is needed in Weinberger's method except that we need a bound only for the base operator $A_0$- not for the full operator $A$. It is very tempting to try and modify Weinberger's method to fit our intermediate problem construction; however, when we try this, we surprisingly find that no modification is necessary at all! Everything falls into place naturally due to the form of our intermediate perturbation $BP_{B^{-1}\mathbf{M}}$.

Defining the intermediate problem $A' = A_0 + BP_{B^{-1}\mathbf{M}} < A$, we try to use Weinberger's method for $A'$ - again it happens that we will be obtaining lower bounds to lower bounds of the full problem $A$, i.e. we are not solving the intermediate problem exactly. We let

$$\mathbf{P} = \mathbf{M},$$

where **P** is the subspace in Weinberger's method such that

$$(A'u,u) \geq \mu \quad \forall\ u \in \mathbf{M}^{\perp} = \mathbf{P}^{\perp}.$$

Expanding $A'$ we find an equality due to the fact that **P** = **M**:

$$\mu \le (A'u,u) = (\{A_0+BP_{B^{-1}\mathbf{M}}\}u,u) = (A_0u,u) + (BP_{B^{-1}\mathbf{M}}u,u)$$

$$= (A_0u,u) + ((P_{B^{-1}\mathbf{M}}u,u)) = (A_0u,u) + (( ((u,B^{-1}\mathbf{M}))B^{-1}\mathbf{M} , u ))$$

$$= (A_0u,u) + ( B(Bu,B^{-1}\mathbf{M})B^{-1}\mathbf{M} , u ) = (A_0u,u) + (u,\mathbf{M})(u,\mathbf{M})$$

$$= (A_0u,u) .$$

Thus the *a priori* information needed is a lower bound to the base problem's spectrum on the infinite-dimensional subspace $\mathbf{P}^{\perp}$. This is exactly what we require for arbitrary choice without truncation. Our specific choice of the intermediate problem and **P** = **M** makes the lower bound to the base problem's spectrum and the intermediate problem's spectrum on $\mathbf{P}^{\perp}$ identical.

Why do we require that the *a priori* information come from the base problem instead of the full problem (or in this case the intermediate problem)? Perhaps we should just use Weinberger's method from the start. When the perturbation is added to the base problem, the resulting operator imediately gets much more complicated. It is very likely that a method for determining eigenvalues of an operator on a suspace of its domain would be much easier to employ on the base problem consisting of only single-particle operators rather than on the intermediate or full problem. This is the reason we want the *a priori* information to come from the base problem.

### 3.5.3 Future work

A variation of arbitrary choice *with* truncation has been proved [Be] to allow convergence of lower bounds generated by intermediate problems to the true eigenvalues of the full problem. However, the work required to obtain sufficiently good lower bounds is such that no one has bothered to calculated them. This also method requires an exactly solvable base problem.

It is possible that arbitrary choice without truncation using Weinberger's method may also allow convergence to the exact eigenvalues expanding the availability of lower bounds to problems without exactly solvable base problems; however, it is also possible that the amount of work necessary to do so will dissuade such investigations. Immediately, one notices that integrals involving the square of the intermediate problem operator are involved just like in many other lower bound methods. The success of arbitrary choice without truncation using Weinberger's method may rest in the effort needed to get the *apriori* information. Unpublished work of Professor Beattie of Virginia

Tech and the late Professor Goerisch of the Institut für Angewandte Mathematik may provide an efficient way of getting such information.

## 3.6 Summary

The goal of this work was to generate an improved method for calculating lower bounds to atomic and molecular energy eigenvalues. A slight improvement was made in Calogero-Marchioro's reduced operator method (section 2.4 and appendix G) althougth the results are still far from being significant.

We have also presented a method, arbitrary choice without truncation, that allows calculation of lower bounds to a system that does not possess an exactly soluble base problem. Unfortunately, practical application of this method has not been realized for two reasons:

> 1. Arbitrary choice without truncation sets up an intermediate problem without the need for an exactly soluble base problem. In doing so it mixes a finite-dimensional matrix eigenvalue problem with an infinite-dimensional matrix eigenvalue problem. This prevents us from solving the intermediate problem; instead, we are restricted to calculating lower bounds to the intermediate problem - i.e. lower bounds to lower bounds of the full problem. Weyl's inequality does not give lower bounds that are close to the intermediate problem's eigenvalues in the case of the helium atom and it seems unlikely that it will for any atom.
>
> 2. The method also fails from a theoretical standpoint similar to that in Temple's method. Temple's method requires a lower bound to the second eigenvalue in order to calculate a lower bound to the first. In our case, we need a lower bound to the base problem operator on a smaller subspace. More times than not, this *a priori* lower bound is not available for use in Temple's method; and for arbitrary choice this seems to also be true.

In conclusion, the method of arbitrary choice without truncation for determining lower bounds to eigenvalues fails for chemical applications in general, just like every other lower bound method known to date.

Fortunately, there exists another approach to our goal that may eliminate the first failure mentioned above. This approach uses the same intermediate problem formulated for arbitrary choice without truncation thus retaining the arbitrariness of **M**. We suggest using Weinberger's method to get lower bounds to this intermediate problem instead of using the Kahan-Weinberger bounds. The second failure mentioned above can possibly be overcome by unpublished work of Beattie and Goerisch.

# References

Br J. Barta. Comptes Rendus Acad. Sci. Paris. 71, 204 (1937) - note: we were unable to locate this article - this is the reference listed in Df and PFK.

Bz Norman W. Bazley. Lower bounds for eigenvalues with application to the helium atom. Phys. Rev., 120, 1, 144-149 (1960)

BF1 Norman W. Bazley and David W. Fox. Lower bounds for eigenvalues of Schrodinger's equation. Phys. Rev. 124, 2, 483-492 (1961)

BF2 Norman W. Bazley and David W. Fox. Lower bounds for energy levels of molecular systems. J. Math. Phys. 4, 9, 1147-1153 (1963)

Be Christopher Beattie. Convergence of essential spectra for intermediate Hamiltonians. SIAM Journal of Mathematical Analysis. 16, 3, 492-499 (1985)

BG Christopher Beattie and Friedrich Goerisch. Methods for computing lower bounds to eigenvalues of self-adjoint operators. Numerische Mathematik, 72, 143-172 (1995)

BBB Carlos F. Bunge, Jose A. Barrientos, and Annik Vivier Bunge. Roothaan-Hartree-Fock ground-state atomic wave functions: Slater-type orbital expansions and expectation value for $Z$=2-54. Atomic Data and Nuclear Data Tables, 53, 113-162 (1993)

Ch Ralph E. Christophersen. Basic principles and techniques of molecular quantum mechanics. New York: Springer-Verlag (1989)

CM F.Calogero and C. Marchioro. Lower bounds to the ground-state energy of systems of identical particles. J. Math. Phys. 10, 4, 562-569 (1969)

CH R. Courant and D. Hilbert. Methods of mathematical physics, vol.1. Interscience Publishers, Inc.: New York (1953)

Df R. J. Duffin. Lower bounds for eigenvalues. Physical Review. 71, 827-828 (1947)

Ek Carl Eckart. The theory and calculation of screening constants. Physical Review, 36, 878-892 (1930)

Fx David W. Fox. Lower bounds for eigenvalues with displacement of essential spectra. SIAM J. Math. Anal. 3, 4, 617-624 (1972)

FR David W. Fox and Werner C. Rheinboldt. Computational methods for determining lower bounds for eigenvalues of operators in Hilbert space. SIAM Review, 8, 4, 427-462 (1966)

FS David W. Fox and Vincent G. Sigillito. Bounds for energies of radial lithium. J. Appl. Math. & Phys. 23, 392-411 (1972)

GR W. M. Greenlee and D. M. Russell. Progress on estimation of energy levels for multi-electron atoms. ISNM Vol 69 Numerical Treatment of Eigenvalue Problems. 3, 89-105 (1983)

HLT P. Hertel, E. H. Lieb, and W. Thirring. Lower bound to the energy of complex atoms. J. Chem. Phys. 62, 8, 3355-3356 (1975)

HS P. D. Hislop and I. M. Sigal. Introduction to spectral theory. New York: Springer-Verlag, (1996)

Kt1 Tosio Kato. Fundamental properties of Hamiltonian operators of Schrodinger type. Trans. Amer. Math. Soc. 70, 195-211 (1951)

Kt2 Tosio Kato. On the upper and lower bounds of eigenvalues. Journal of the Physical Society of Japan. 4, 334-339 (1949)

Ki Toichiro Kinoshita. Ground state of the helium atom. Phys. Rev., 105, 5, 1490-1502 (1957)

Kn W. Kohn. A note on Weinstein's variational method. Phys. Rev., 71, 12 902-904 (1947)

KW Kolos and Wolniewicz. J. Chem. Phys. 41, 12, 3663-3673 (1964)

Kr Erwin Kreyszig. Introductory Functional Analysis With Applications. New York: John Wiley & Sons (1989)

Lv N. N. Lebedev. Special functions and their applications (translated by Richard A. Silverman). New York: Dover Publications, Inc. (1972)

LD Taesul Lee and T. P. Das. Correlation energies of ten-electron molecular systems by a united-atom many-body perturbation procedure: $H_2O$, $NH_3$, and $CH_4$. Phys. Rev. A, 6, 3, 968-976 (1972)

Lb Elliot H Lieb. The stability of matter. Rev. Mod. Phys. 48, 4, 553-569 (1976)

Mc J. K. L. MacDonald. Successive approximations by the Rayleigh-Ritz variation method. Phys. Rev. 43, 830-833 (1933)

MBS M. G. Marmorino, C. A. Beattie, and J. C. Schug. To be published.

MSB M. G. Marmorino, J. C. Schug, and C. A. Beattie. To be published.

Mi William H. Miller. New equation for lower bounds to eigenvalues with application to the helium atom. Journal of Chemical Physics, 42, 4305-4306 (1965)

Pr Beresford N. Parlett. The symmetric eigenvalue problem. New Jersey: Prentice-Hall (1980)

PW Linus Pauling and E. Bright Wilson, Jr. Introduction to quantum mechanics (reprint of of 1935 edition). New York: Dover Publications (1985)

Po G. Polya. Estimates for eigenvalues. Studies in Mathematics and Mechanis, presented to Richard von Mises. New York: Academic Press, 200-207 (1954)

PFK I. Porras, D. Matthew Feldmann, and Frederick W. King. A nonlinear programming approach to lower bounds for the ground state energy of helium. International Journal of Quantum Chemistry. 71, 455-463 (1999)

Ry Lord Rayleigh (J. W. Strutt). The theory of sound (reprint of the 1894 second editon). New York: Dover Publications (1945) section 88.

Rd1 C. E. Reid. Intermediate Hamiltonians for the lithium atom. Int. J. Quant. Chem. 6, 793-797 (1972)

Rd2 Charles E. Reid. Lower bounds for the energy levels of the lithium atom. Chemical Physics Letters, 26, 2, 243-245 (1974)

SRK R. A. Sack, C. C. J. Roothaan, and W. Kolos. Recursive evaluation of some atomic integrals. Journal of Mathematical Physics. 8, 5 1093-1094 (1967)

Sn G. Sansone. Orthogonal functions (translated by Ainsley H. Diamond). New York: Dover Publications, Inc. (1991)

Sc Armin Scrinzi. Lower bounds to the binding energies of $td\mu$. Phys. Rev. A, 45, 11, 7787-7791 (1992)

SS Rikuzo Seto and Ivan V. Stankevich. Int. J. Quan. Chem. 72, 101-107 (1999)

Sh R. Shankar. Principles of quantum mechanics (second edition). New York: Plenum Press (1994)

SL Harrison Shull and Per-Olov Lowdin. Role of the continuum in superposition of configurations. J. Chem. Phys. 23, 1362 (1956)

SP Lawrence C. Snyder and Robert G. Parr. Extraordinary basic functions in valence theory. J. Chem. Phys. 34, 5, 1661-1665 (1961)

Th Walter Thirring. A course in mathematical physics, vol. 3: quantum mechanics of atoms and molecules. New York: Springer-Verlag (1981)

Tm G. Temple. The theory of Rayleigh's principle as applied to continuous systems. Proc. Roy. Soc. 119A , 276-293 (1928)

Wd Joachim Weidmann. On the continuous spectrum of Schrodinger operators. Comm. Pure Appl. Math. 19, 107-110 (1966)

Wb Hans F. Weinberger. Variational methods for eigenvalue approximation. Philadelphia: Society for Industrial and Applied Mathematics (1974)

Wn D.H. Weinstein. "Modified Ritz Method." Proceedings of the National Academy of Sciences of the U.S. 20 (1934) 529-532.

WS Alexander Weinstein and William Stenger. Methods of intermediate problems for eigenvalues: theory and ramifications. New York: Academic Press (1972)

Wd H. Wind. J. Chem. Phys. 42, 7, 2371-2373 (1965)

Wf Wolfram Research, Inc., Mathematica, Version 3.0, Chanpaign, IL (1996)

# Appendix A

## The united atom theorem

United Atom Theorem: The infimal spectral point of an *N*-electron molecular Hamiltonian is greater than or equal to the infimal spectral point of an *N*-electron atomic Hamiltonian with nuclear charge equal to the sum of the nuclear charges in the molecular system.

Proof: We write the electronic molecular Hamiltonian as follows:

$$H = \sum_{m=1}^{M} \sum_{i=1}^{N} V_m(Z_m, i) + \sum_{i=1}^{N} T(i) + \sum_{i<j}^{N} R(i,j) \ ,$$

$$V_m(Z_m, i) = -\frac{Z_m}{|\mathbf{r}_m - \mathbf{r}_i|} \ ,$$

where $N$ is the number of electrons, $M$ is the number of nuclei, $Z_m$ is the nuclear charge of the $m$-th nucleus, and $\mathbf{r}_m$ and $\mathbf{r}_i$ are the vector coordinates of the $m$-th and $i$-th nucleus and electron, respectively. We also define $Z_{Tot} = \sum Z_m$.

We divide the Hamiltonian into $M$ parts:

$$\begin{aligned} H \ = \ & \Big[\sum_{i=1}^{N} V_1(Z_1, i) + \frac{Z_1}{Z_{Tot1}} \sum_{i=1}^{N} T(i) + \frac{Z_1}{Z_{Tot1}} \sum_{i<j}^{N} R(i,j)\Big] \\ & + \ldots . + \Big[\sum_{i=1}^{N} V_M(Z_M, i) + \frac{Z_M}{Z_{Tot}} \sum_{i=1}^{N} T(i) + \frac{Z_M}{Z_{Tot}} \sum_{i<j}^{N} R(i,j)\Big] \end{aligned}$$

$$\begin{aligned} = \ & \frac{Z_1}{Z_{Tot}} \Big[\frac{Z_{Tot}}{Z_1} \sum_{i=1}^{N} V_1(Z_1, i) + \sum_{i=1}^{N} T(i) + \sum_{i<j}^{N} R(i,j)\Big] \\ & + \ldots . + \frac{Z_M}{Z_{Tot}} \Big[\frac{Z_{Tot}}{Z_M} \sum_{i=1}^{N} V_M(Z_M, i) + \sum_{i=1}^{N} T(i) + \sum_{i<j}^{N} R(i,j)\Big] \end{aligned}$$

$$= \frac{Z_1}{Z_{Tot}}\Big[\sum_{i=1}^{N} V_1(Z_{Tot}, i) + \sum_{i=1}^{N} T(i) + \sum_{i<j}^{N} R(i,j)\Big]$$

$$+ \,....\, + \frac{Z_M}{Z_{Tot}}\Big[\sum_{i=1}^{N} V_M(Z_{Tot}, i) + \sum_{i=1}^{N} T(i) + \sum_{i<j}^{N} R(i,j)\Big]$$

$$= \frac{Z_1}{Z_{Tot}}[H_{Atom}(Z_{Tot}, N)] + \,....\, + \frac{Z_M}{Z_{Tot}}[H_{Atom}(Z_{Tot}, N)]$$

where $H_{\text{Atom}}(Z,N)$ is the Hamiltonian for an $N$-electron atom with nuclear charge $Z$.

Using Weyl's inequality [Wy] (see section 2.1), we solve for the lowest points of the spectra for each of the $M$ identical atomic Hamiltonians and combine them to get the molecular electronic spectral bound:

$$E \geq \frac{Z_1}{Z_{Tot}} E_{Atom}(Z_{Tot}, N) + \ldots + \frac{Z_M}{Z_{Tot}} E_{Atom}(Z_{Tot}, N) = E_{Atom}(Z_{Tot}, N) \quad ,$$

where $E$ and $E_{\text{Atom}}(Z,N)$ are the lowest points of the molecular electronic and atomic spectra, respectively.

Thus the electronic energy of any molecular system is bounded below by the energy of a simply related atomic system. Since the nuclear contributions (kinetic and nuclear-nuclear repulsion) are strictly positive, the total energy is also bounded below by the atomic energy.

This theorem can also be extended to give lower bounds to excited states. These bounds are very poor in general as seen in table 4.1. The Hartree-Fock molecular ground state energies are taken from [Ch] and the Hartree-Fock atomic ground state energies are taken from [BBB]. The Hartree-Fock (HF) values make the bounds non-rigorous, but they are accurate enough to show how poorly the united atom serves as a base problem.

**Table 4.1** Comparison of united atom bound with actual ground state energy

| Molecule | HF ground state (a.u.) | HF united atom bound (a.u.) |
|---|---|---|
| $H_2$ | -1.134 | He -2.862 |
| $CH_4$ | -40.23 | Ne -128.5 |
| $NH_3$ | -56.23 | Ne -128.5 |
| $H_2O$ | -76.07 | Ne -128.5 |
| HF | -100.1 | Ne -128.5 |
| $N_2$ | -109.0 | Si -288.9 |
| CO | -112.8 | Si -288.9 |
| $F_2$ | -198.8 | Ar -526.8 |

# Appendix B

## Orthogonal vs. non-orthogonal projections

It is widely known that given a linearly independent set $\{u_n\}$ of $N$ functions whose span is an $N$-dimensional Hilbert space with inner product $(\cdot,\cdot)$, there exists a unique representation of each element $u \in \mathbf{H}$ of the form:

$$u = a_1 u_1 + a_2 u_2 + \ldots + a_N u_N \,.$$

Perhaps not appreciated is the fact that $a_n = (u,u_n)$ if and only if $\{u_n\}$ is an orthonormal set in addition to being linearly independent. Take, for example the two-dimensional Hilbert space, $\mathbf{R}^2$:

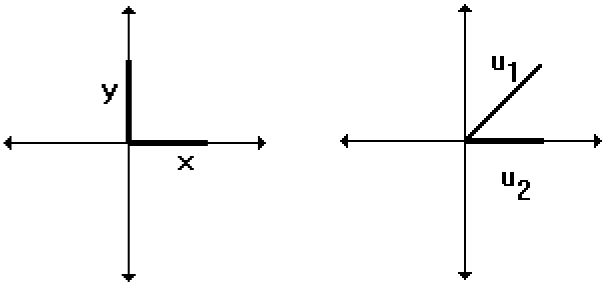

**Figure B.1**

Let $x$ and $y$ be the standard linearly independent orthonormal unit vectors and let $u_1$ and $u_2$ be the linearly independent, but not orthonormal, vectors: $u_1 = x + y$, $u_2 = x$ (see figure B.1). We take $u = u_1$ as a trial function for demonstration.

$$u = 1\,x + 1\,y = (u,x)\,x + (u,y)\,y = 1\,x + 1\,y$$

$$u = 1\,u_1 + 0\,u_2 \neq (u,u_1)\,u_1 + (u,u_2)\,u_2 = 2\,u_1 + 1\,u_2.$$

This causes some complications when building a projection operator. Usually one defines a projection onto a subspace $\mathbf{K} = \mathrm{Span}[u_1,u_2,\ldots,u_N] \subset \mathbf{H}$ of dimension $N$ as

$$P_{\mathbf{K}} = \Sigma\,(\cdot,u_n)u_n \,.$$

From our example above, we see that this is only valid if $\{u_1,u_2,\ldots,u_N\}$ is an orthonormal set. We now generalize, without proof, the formula for a projection operator.

We first write the usual projection operator in a slightly different form:

$$P_{\mathbf{K}} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ : \\ a_N \\ a_{N+1} \\ : \end{bmatrix} = I \begin{bmatrix} (u,u_1) = a_1 \\ (u,u_2) = a_2 \\ (u,u_3) = a_3 \\ : \\ (u,u_N) = a_N \\ (u,u_{N+1}) = 0 \\ : \end{bmatrix}$$

The more general form is:

$$P_{\mathbf{K}} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ : \\ a_N \\ a_{N+1} \\ : \end{bmatrix} = G^{-1} \begin{bmatrix} (u,u_1) \neq a_1 \\ (u,u_2) \neq a_2 \\ (u,u_3) \neq a_3 \\ : \\ (u,u_N) \neq a_N \\ (u,u_{N+1}) = 0 \\ : \end{bmatrix}$$

where $G^{-1}$ is the inverse of the gram matrix formed from the basis set $\{u_1,u_2,...,u_N\}$:

$$G = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ : \\ u_N \end{bmatrix} \mathbf{[}\, u_1, u_2, ..., u_N \,\mathbf{]} \ .$$

# Appendix C

## Radial integral formulas

The associated Laguerre polynomials, $L_n^2(x)$, are defined [Lv,Sn] equivalently as:

$$L_n^2(x) = \frac{e^x}{x^2 n!}\frac{d^n}{dx^n}(e^{-x}x^{n+2}) = \sum_{k=0}^{n}\frac{(n+2)!(-x)^k}{(k+2)!(n-k)!k!} \quad , n = 0,1,2,3,\ldots .$$

An complete orthogonal set of one-variable functions defined on $x\in[0,\infty)$ with integration weight $x^2$ is [Lv,Sn]

$$L_n^2(x)e^{-x/2} \;,\quad n = 0,1,2,3,\ldots .$$

This gives a complete orthogonal set of two-variable functions defined on $x\in[0,\infty)$ and $y\in[0,\infty)$ as

$$L_{n_1}^2(x)e^{-x/2} \times L_{n_2}^2(y)e^{-y/2} \;,\; n_1,n_2 = 0,1,2,3,\ldots .$$

We introduce a variable parameter, $2a$, and normalize the basis set to

$$\phi_{n1}(r_1)\times\phi_{n2}(r_2) = \sqrt{\tfrac{(2a)^3}{(n_1+1)(n_1+2)}}\, L_{n_1}^2(2ar_1)e^{-ar_1} \times \sqrt{\tfrac{(2a)^3}{(n_2+1)(n_2+2)}}\, L_{n_2}^2(2ar_2)e^{-ar_2} \;.$$

For our calculations, we need the following integrals (where $<\cdot,\cdot>$ denotes a one variable inner product):

$$(r_{12}\phi_k\times\phi_l,\phi_m\times\phi_n)$$
$$([T\otimes I]\phi_k\times\phi_l,\phi_m\times\phi_n) = <T\phi_k,\phi_m><\phi_l,\phi_n>$$
$$([T^2\otimes I]\phi_k\times\phi_l,\phi_m\times\phi_n) = <T^2\phi_k,\phi_m><\phi_l,\phi_n>$$
$$([I\otimes T]\phi_k\times\phi_l,\phi_m\times\phi_n) = <\phi_k,\phi_m><T\phi_l,\phi_n>$$
$$([I\otimes T^2]\phi_k\times\phi_l,\phi_m\times\phi_n) = <\phi_k,\phi_m><T^2\phi_l,\phi_n>$$
$$([V\otimes I]\phi_k\times\phi_l,\phi_m\times\phi_n) = <V\phi_k,\phi_m><\phi_l,\phi_n>$$
$$([V^2\otimes I]\phi_k\times\phi_l,\phi_m\times\phi_n) = <V^2\phi_k,\phi_m><\phi_l,\phi_n>$$
$$([I\otimes V]\phi_k\times\phi_l,\phi_m\times\phi_n) = <\phi_k,\phi_m><V\phi_l,\phi_n>$$
$$([I\otimes V^2]\phi_k\times\phi_l,\phi_m\times\phi_n) = <\phi_k,\phi_m><V^2\phi_l,\phi_n>$$

$$
\begin{aligned}
([T\otimes V]\phi_k\times\phi_l,\phi_m\times\phi_n) &= \langle T\phi_k,\phi_m\rangle\,\langle V\phi_l,\phi_n\rangle \\
([V\otimes T]\phi_k\times\phi_l,\phi_m\times\phi_n) &= \langle V\phi_k,\phi_m\rangle\,\langle T\phi_l,\phi_n\rangle \\
([VT\otimes I]\phi_k\times\phi_l,\phi_m\times\phi_n) &= \langle VT\phi_k,\phi_m\rangle\,\langle\phi_l,\phi_n\rangle \\
([TV\otimes I]\phi_k\times\phi_l,\phi_m\times\phi_n) &= \langle TV\phi_k,\phi_m\rangle\,\langle\phi_l,\phi_n\rangle \\
([I\otimes VT]\phi_k\times\phi_l,\phi_m\times\phi_n) &= \langle\phi_k,\phi_m\rangle\,\langle VT\phi_l,\phi_n\rangle \\
([I\otimes TV]\phi_k\times\phi_l,\phi_m\times\phi_n) &= \langle\phi_k,\phi_m\rangle\,\langle TV\phi_l,\phi_n\rangle
\end{aligned}
$$

where $T = -\frac{1}{2r^2}\frac{d}{dr}r^2\frac{d}{dr}$ and $V = -\frac{Z}{r}$. This amounts to only six basic integrals:

$$
\begin{gathered}
(r_{12}\phi_k\times\phi_l,\phi_m\times\phi_n) \\
\langle T\phi_k,\phi_m\rangle \\
\langle V\phi_k,\phi_m\rangle \\
\langle T^2\phi_k,\phi_m\rangle \\
\langle V^2\phi_k,\phi_m\rangle \\
\langle VT\phi_l,\phi_n\rangle
\end{gathered}
$$

The kinetic energy integral is simpler to solve in the form of

$$\langle T\phi_k,\phi_m\rangle = \tfrac{1}{2}\langle D\phi_k,D\phi_m\rangle \quad \text{where } D = \frac{d}{dr}.$$

This equality is valid if the functions are everywhere differentiable. If there are a finite number of points where the functions are not differentiable, then another term must be added [SP].

For the electron-repulsion integral, we have the operator identity [Rd2]

$$r_{12} = 2r_> + \frac{2r_<^2}{3r_>} .$$

Using the summation expression for the associate Laguerre polynomials we obtain the following formulas:

$$\langle V\phi_m,\phi_n\rangle = \frac{-Z(2a)(m+2)!(n+2)!}{\sqrt{(m+1)(m+2)(n+1)(n+2)}}\sum_{c=0}^{m}\sum_{k=0}^{n}\frac{(-1)^{c+k}(c+k+1)!}{(c+2)!(k+2)!(m-c)!(n-k)!c!k!}$$

$$\langle V^2\phi_m,\phi_n\rangle = \frac{(2aZ)^2(m+2)!(n+2)!}{\sqrt{(m+1)(m+2)(n+1)(n+2)}}\sum_{c=0}^{m}\sum_{k=0}^{n}\frac{(-1)^{c+k}(c+k)!}{(c+2)!(k+2)!(m-c)!(n-k)!c!k!}$$

$$\langle T\phi_m,\phi_n\rangle = \frac{(2a)^2(m+2)!(n+2)!}{2\sqrt{(m+1)(m+2)(n+1)(n+2)}}\sum_{c=0}^{m}\sum_{k=0}^{n}\frac{(-1)^{c+k}(c+k+2)!}{(c+2)!(k+2)!(m-c)!(n-k)!c!k!}\times$$

$$\left[\frac{(m+3)(n+3)}{(c+3)(k+3)}-\frac{1}{2}\frac{(m+3)}{(c+3)}-\frac{1}{2}\frac{(n+3)}{(k+3)}+\frac{1}{4}\right]$$

$$\langle T^2\phi_m,\phi_n\rangle = \frac{(2a)^4(m+2)!(n+2)!}{4\sqrt{(m+1)(m+2)(n+1)(n+2)}}\sum_{c=0}^{m}\sum_{k=0}^{n}\frac{(-1)^{c+k}(c+k)!}{(c+2)!(k+2)!(m-c)!(n-k)!c!k!}\times$$

$$\left[\frac{1}{16}\frac{(c+k+2)!}{(c+k)!}-\frac{1}{4}(m+n+4)(c+k+1)+\frac{1}{4}\frac{(m+3)}{(c+3)}(c+k+1)+\frac{1}{4}\frac{(n+3)}{(k+3)}(c+k+1)\right.$$

$$\left.+\frac{(m+3)(n+3)}{(c+3)(k+3)}+(m+2)(n+2)-\frac{(m+2)(n+3)}{(k+3)}-\frac{(n+2)(m+3)}{(c+3)}\right]$$

$$\langle VT\phi_m,\phi_n\rangle = \frac{-Z(2a)^3(m+2)!(n+2)!}{2\sqrt{(m+1)(m+2)(n+1)(n+2)}}\sum_{c=0}^{m}\sum_{k=0}^{n}\frac{(-1)^{c+k}(c+k)!}{(c+2)!(k+2)!(m-c)!(n-k)!c!k!}\times$$

$$\left[\frac{(m+3)(n+3)}{(c+3)(k+3)}(c+k+1)-\frac{1}{2}\frac{(m+3)}{(c+3)}(c+k+1)-\frac{1}{2}\frac{(n+3)}{(k+3)}(c+k+1)+\frac{(n+3)}{(k+3)}-\frac{1}{2}+\frac{(c+k+1)}{4}\right]$$

$$(r_{12}\phi_m \times \phi_s, \phi_n \times \phi_t) \;=\; \frac{(m+2)!(s+2)!(n+2)!(t+2)!}{3(2a)\sqrt{(m+1)(m+2)(s+1)(s+2)(n+1)(n+2)(t+1)(t+2)}} \times$$

$$\sum_{i=0}^{m}\sum_{j=0}^{s}\sum_{k=0}^{n}\sum_{l=0}^{t} \frac{1}{(i+2)!(j+2)!(k+2)!(l+2)!(m-i)!(s-j)!(n-k)!(t-l)!\,i!\,j!\,k!\,l!} \times$$

$$\Bigg[\; \frac{3(k+l+3)!}{2^{i+j+3}} \sum_{c=0}^{k+l+3} \frac{(i+j+2+c)!}{c!2^c} + \frac{(k+l+1)!}{2^{i+j+5}} \sum_{c=0}^{k+l+1} \frac{(i+j+4+c)!}{c!2^c}$$

$$+ \frac{3(i+j+3)!}{2^{k+l+3}} \sum_{c=0}^{i+j+3} \frac{(k+l+2+c)!}{c!2^c} + \frac{(i+j+1)!}{2^{k+l+5}} \sum_{c=0}^{i+j+1} \frac{(k+l+4+c)!}{c!2^c} \;\Bigg]$$

Fox and Sigillito [FS] reduced the summations in the kinetic energy and nuclear attraction integrals to yield the much simpler formulas

$$<V\phi_m,\phi_n> \;=\; -a\,Z\sqrt{\frac{(m+1)(m+2)}{(n+1)(n+2)}} \quad \text{for } m \le n \;\;,$$

$$<T\phi_m,\phi_n> \;=\; a^2\,\frac{4m+3}{6}\sqrt{\frac{(m+1)(m+2)}{(n+1)(n+2)}} + \frac{a^2 d_{m,n}}{2} \quad \text{for } m \le n \;.$$

# Appendix D

## Forming the matrices

We now show how to form the relevant matrices defined at the end of section 3.3.4 by reducing the matrix elements to integrals defined in Appendix C.

1) $A_1 = [(H_0 m_i, m_j)]$

$H_0(1,2) = T(1)\otimes I(2) + V(1)\otimes I(2) + I(1)\otimes T(2) + I(1)\otimes V(2)$

2) $A_2 = [(B P_{B^{-1}\mathbf{M}} m_i, m_j)]$

$$B P_{B^{-1}\mathbf{M}} = B G^{-1}((\cdot, B^{-1}\mathbf{M}))B^{-1}\mathbf{M} = G^{-1}(\cdot, \mathbf{M})\mathbf{M} = G^{-1}\sum_{i=1}^{n^2}(\cdot, m_i)m_i$$

where $G^{-1} = [(B^{-1}m_i, m_j)]^{-1}$ .

3) $C = [(H_0 m_i, r_j)] = [(H_0 m_i, P_{\mathbf{M}^\perp} H_0 m_{jj})]$ where we use the double subscript "*jj*" to remind us that there are only $2n$-1 $m_{jj}$ whereas there are $n^2$ $m_i$.

$$\begin{aligned}(H_0 m_i, P_{\mathbf{M}^\perp} H_0 m_{jj}) &= (H_0 m_i, P_{\mathbf{M}^\perp} P_{\mathbf{M}^\perp} H_0 m_{jj}) = (P_{\mathbf{M}^\perp} H_0 m_i, P_{\mathbf{M}^\perp} H_0 m_{jj}) \\ &= \Big( H_o m_i - \sum_{c=1}^{n^2}(H_o m_i, m_c)m_c \, , \, H_o m_{jj} - \sum_{k=1}^{n^2}(H_o m_{jj}, m_k)m_k \Big) \\ &= (H_o m_i, H_o m_{jj}) - \Big( H_o m_i \, , \sum_{k=1}^{n^2}(H_o m_{jj}, m_k)m_k \Big) - \Big( \sum_{c=1}^{n^2}(H_o m_i, m_c)m_c \, , \, H_o m_{jj} \Big) \\ &\quad + \Big( \sum_{c=1}^{n^2}(H_o m_i, m_c)m_c \, , \sum_{k=1}^{n^2}(H_o m_{jj}, m_k)m_k \Big)\end{aligned}$$

$$= (H_o m_{i,} H_o m_{jj}) - \sum_{k=1}^{n^2} (H_o m_{jj}, m_k)(H_o m_i, m_k) - \sum_{c=1}^{n^2} (H_o m_i, m_c)(m_c, H_o m_{jj}) + \sum_{c=1}^{n^2} \sum_{k=1}^{n^2} (H_o m_{jj}, m_k)(H_o m_i, m_c)(m_c, m_k)$$

$$= (H_o m_{i,} H_o m_{jj}) - \sum_{k=1}^{n^2} (H_o m_{jj}, m_k)(H_o m_i, m_k) - \sum_{k=1}^{n^2} (H_o m_i, m_k)(H_o m_{jj}, m_k) + \sum_{k=1}^{n^2} (H_o m_{jj}, m_k)(H_o m_i, m_k)$$

$$= (H_o^2 m_{i,} m_{jj}) - \sum_{k=1}^{n^2} (H_o m_{jj}, m_k)(H_o m_i, m_k)$$

$$\begin{aligned} H_o^2(1,2) = {} & T^2(1)\otimes I(2) + V^2(1)\otimes I(2) + I(1)\otimes T^2(2) + I(1)\otimes V^2(2) \\ & + TV(1)\otimes I(2) + VT(1)\otimes I(2) + I(1)\otimes TV(2) + I(1)\otimes VT(2) \\ & + 2T(1)\otimes T(2) + 2V(1)\otimes V(2) + 2T(1)\otimes V(2) + 2V(1)\otimes T(2) \end{aligned}$$

4) $C^{\mathrm{T}} = [(H_o r_i, m_j)]$ This is just the transpose of $C$.

5) $J = [(r_i, r_j)] = [(P_{\mathbf{M}^\perp} H_o m_{ii}, P_{\mathbf{M}^\perp} H_o m_{jj})]$

$$= (H_o^2 m_{ii,} m_{jj}) - \sum_{k=1}^{n^2} (H_o m_{jj}, m_k)(H_o m_{ii}, m_k)$$

(see matrix elements of $C$ for details)

6) $K = K(J, C^{\mathrm{T}})$

We let *Mathematica* [Wf] form $K$ by inverting $S$ and forming $S^{-1} H_{\mathrm{I}}$.