A modified version of the Waxman algorithm

W A Berger and H G Miller

Department of Physics, University of Pretoria, Pretoria 0002, South Africa

E-mail: wb@wberger.com and hmiller@maple.up.ac.za

Abstract
The iterative algorithm recently proposed by Waxman for solving eigenvalue
problems, which relies on the method of moments, has been modified
to improve its convergence considerably without sacrificing its benefits or
elegance. The suggested modification is based on methods to calculate low-
lying eigenpairs of large bounded Hermitian operators or matrices.

PACS numbers: 03.65.Ge, 02.60.Lj

Recently, Waxman [1] has proposed a convergent iterative algorithm for obtaining solutions
of the bound state eigenvalue problem, which does not involve a matrix diagonalization.
For operators which possess a continuum as well as a set of bound states, this is most
advantageous [2]. In the case of the ground state, for example, the eigenenergy, ϵ, is determined
numerically as a function of the coupling constant of the potential, λ, and inverted to yield the
ϵ corresponding to the required value of λ. The convergence rate of the algorithm, therefore,
depends on two factors: the number of iterations required to find an eigensolution for a
particular choice of ϵ and the number of times this must be repeated in order to determine the
value of ϵ which corresponds to the desired value of λ. In a recent paper, we have shown that
for many non-singular symmetric potentials which vanish asymptotically, a simple analytical
relationship between the coupling constant of the potential and the ground-state eigenvalue
exists which can be used to reduce the number of times λ has to be calculated for a given value
of ϵ [3]. Here we show that it is also possible to reduce the number of iterations necessary to
determine an eigensolution for a particular choice of ϵ. Furthermore, the existence of a simple
analytical relationship between ϵ and λ can be used to handle problematic situations that are
referred to as pseudoconvergence in other methods [4] and which can occur in the Waxman
algorithm as well.

In the Waxman algorithm, eigenpairs are determined as functions of the strength of the
potential in the following manner [1]. Here we shall use the abstract Hilbert space notation
for the most part and point out their meaning in a one-dimensional coordinate space where
appropriate. Starting from

$$(T - \lambda V)|u\rangle = \epsilon|u\rangle, \tag{1}$$

with

$$u(x) = \langle x|u\rangle, \quad \lim_{|x|->\infty} u(x) = 0, \tag{2}$$

where T is the kinetic energy operator (or more generally a suitable Hermitian operator for the unperturbed system), $\lambda > 0$ is the strength parameter of the potential ($\lambda V > 0$ and $V(x) \to 0$ *as* $|x| \to \infty$) and the energy eigenvalue, ϵ (with $\epsilon < 0$), is negative and corresponds to a bound state. Using Green's method, a solution to equation (1) is given by

$$|u\rangle = \lambda G_\epsilon V|u\rangle, \tag{3}$$

where the Green's operator, G_ϵ, can formally be defined as

$$G_\epsilon = (T - \epsilon)^{-1}, \tag{4}$$

and the corresponding Green's function satisfies

$$\lim_{|x|->\infty} G_\epsilon(x) = 0. \tag{5}$$

Normalizing $|u\rangle$ with a suitable reference state $|\text{ref}\rangle$

$$\langle \text{ref}|u\rangle = 1 \tag{6}$$

allows λ to be written as (see equation (3))

$$\lambda = \langle \text{ref}|G_\epsilon V|u\rangle^{-1}, \tag{7}$$

which can then be used to eliminate λ from equation (3)

$$|u\rangle = \frac{G_\epsilon V|u\rangle}{\langle \text{ref}|G_\epsilon V|u\rangle}. \tag{8}$$

From equations (7) and (8), λ can be determined as a function of ϵ in the following manner. For a particular choice of ϵ, equation (8) can be iterated

$$|n + 1\rangle = \frac{G_\epsilon V|n\rangle}{\langle \text{ref}|G_\epsilon V|n\rangle} \tag{9}$$

until it converges and λ can then be determined from equation (7). Repeating for different values of ϵ yields a set of different values of the potential strength λ. When enough points have been determined, a simple interpolation procedure can be used to determine the dependence of ϵ on λ.

If we assume, that u is square integrable and therefore can be normalized we can choose $|u\rangle$ as a reference vector in (8) yielding

$$|u\rangle\langle u|G_\epsilon V|u\rangle = G_\epsilon V|u\rangle \tag{10}$$

from which it is clearly seen that we are dealing with an eigenvalue problem for the operator $G_\epsilon V$ with eigenvalue $\lambda^{-1} = \langle u|G_\epsilon V|u\rangle$. The equation corresponding to (9) is now given by

$$|n + 1\rangle = \frac{G_\epsilon V|n\rangle}{\langle n|G_\epsilon V|n\rangle}. \tag{11}$$

From (11), it immediately follows $\langle n|n + 1\rangle = 1$ and therefore

$$|n + 1\rangle = |n\rangle + c_n^\perp |n_\perp\rangle, \tag{12}$$

where the (at present) unknown vector $|n_\perp\rangle$ obeys $\langle n|n_\perp\rangle = 0$, $\langle n_\perp|n_\perp\rangle = 1$ and $c_n^\perp \neq 0$ as long as $|n\rangle$ is not an eigenvector. Normalizing $|n + 1\rangle$ yields

$$|n + 1\rangle' = \left(1 + \left|c_n^\perp\right|^2\right)^{\frac{-1}{2}} \left(|n\rangle + c_n^\perp |n_\perp\rangle\right). \tag{13}$$

Equation (13) defines the start vector for the next iteration step, and it can be easily seen that the relative contribution of the 'leading' term $|n\rangle$ is reduced in each iteration step.

It is now essential to note that in the n+1 iteration step, we have to deal with two vectors: $|n\rangle$ and $|n_\perp\rangle$ which define (in general) a two-dimensional subspace. Therefore, one could ask, does equation (13) already define the best choice within this subspace to achieve optimal convergence. Remember, we are looking for a linear combination of the form

$$|n+1\rangle'' = c_1|n\rangle + c_2 G_\epsilon V|n\rangle = d_n|n\rangle + d_n^\perp|n_\perp\rangle, \tag{14}$$

where the c's and the d's are chosen to satisfy the normalization and the convergence requirements. Clearly, in the original Waxman scheme $c_1 = 0$. This corresponds to using the method of moments or power method [5, 6] to solve equation (10).

Another choice would be to diagonalize $G_\epsilon V$ projected onto the two-dimensional subspace spanned by $|n\rangle$ and $|n_\perp\rangle$, and take one of the two eigenvectors as the new start vector in the next iteration step. Such an approach has been proposed to calculate low-lying eigenvalues of a Hermitian operator [7] and generalized to low-lying eigenstates of unbounded Hermitian operators [8].

The iteration scheme is then the following. From $G_\epsilon V|n\rangle$, calculate a normalized vector orthogonal to the vector $|n\rangle$:

$$|n_\perp\rangle = \left(\frac{G_\epsilon V|n\rangle}{\langle n|G_\epsilon V|n\rangle} - |n\rangle \right) * \left| c_n^\perp \right|^{-1}, \tag{15}$$

such that $\langle n_\perp|n_\perp\rangle = 1$ and $\langle n|n_\perp\rangle = 0$. The matrix representation of $G_\epsilon V$ in the subspace spanned by $|n\rangle$ and $|n_\perp\rangle$ is

$$\begin{pmatrix} \epsilon_n & v_n \\ v_n^* & \alpha_n \end{pmatrix}, \tag{16}$$

with

$$\epsilon_n = \langle n|G_\epsilon V|n\rangle = \lambda_n^{-1}, \qquad v_n = \langle n|G_\epsilon V|n_\perp\rangle, \qquad \alpha_n = \langle n_\perp|G_\epsilon V|n_\perp\rangle. \tag{17}$$

Now the 2×2 matrix (16) can be diagonalized yielding two eigenvalues and the corresponding two eigenvectors. One of the eigenvalues always lies above ϵ_n, the other one below ϵ_n as long as $v_n \neq 0$. (Here we make use of the fact that λ is real.) If we always choose the upper eigenvalue, the ϵ_n form a monotonically increasing sequence bounded by the highest eigenvalue of $G_\epsilon V$. Therefore, the sequence is convergent. (Similar arguments hold, if one chooses the lower eigenvalue in each step in case the spectrum is bounded from below.) Since in each step the modified iteration step is an optimization with respect to the original power method step, one might expect an improved convergence rate.

In order to investigate the convergence properties, we have performed a number of calculations with matrices where a discrete spectrum for T was chosen and a random potential V was added. The signs in equation (1) were chosen such that λ is positive and the monotonically increasing iteration scheme (i.e. increasing in terms of $\epsilon_n = \lambda_n^{-1}$ and therefore decreasing in terms of λ_n) described above was chosen. Figure 1 shows a typical convergence pattern for a 20×20 matrix using the original Waxman method (power) compared to the modified scheme proposed here (2×2). The number of iterations needed until a specified convergence limit is achieved for the new scheme is roughly one half that required for the original Waxman algorithm. Figure 2 illustrates the convergence properties of the modified iteration scheme in another, more dramatic, case where the original algorithm has considerable difficulty to converge. It can be seen from both examples that the convergence rate improves significantly.

To judge the advantage of the faster convergence one has to take into account that because of the last term in equations (17) the new scheme is always one additional iteration step ($G_\epsilon V|n_\perp\rangle$) ahead, and in each iteration step there are more vector operations. While the latter
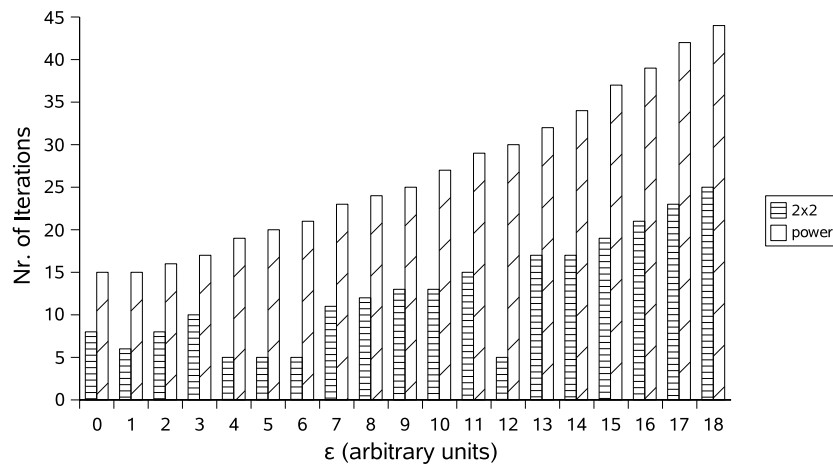
Figure 1. Typical convergence rates of the coupling constant λ for different ground-state energies ϵ for the Waxman algorithm (power) and the modified algorithm (2 × 2) using a 20 × 20 matrix model Hamiltonian.
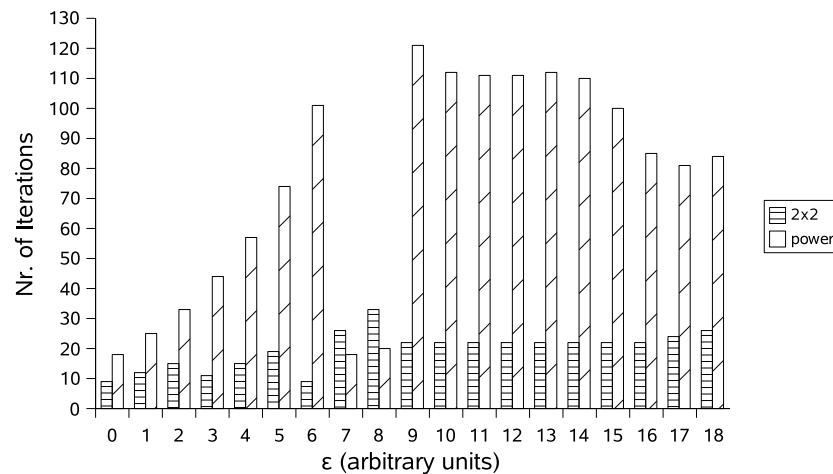

Figure 2. Convergence rates of the coupling constant λ of different ground-state energies ϵ for the Waxman algorithm (power) and the modified algorithm (2 × 2) using a 20 × 20 matrix which yields very poor convergence as the model Hamiltonian.

are negligible as far as computational resources are concerned, the additional iteration step has to be taken into account. Even then this increases the total number of iterations only by 1 and still results in a considerably accelerated convergence.

It is interesting to understand why achieving convergence in the second example is so slow. In figure 3, we show the convergence of λ_n for different energies as a function of iteration number n. It can be seen that convergence in certain cases is delayed by a number of steps in such a way that seems to indicate that ϵ (or λ) appears to converge to an incorrect value. Such behaviour, which we referred to as pseudoconvergence, is common in several iterative algorithms and has been discussed earlier for the Lanczos and modified-Lanczos algorithm [4].
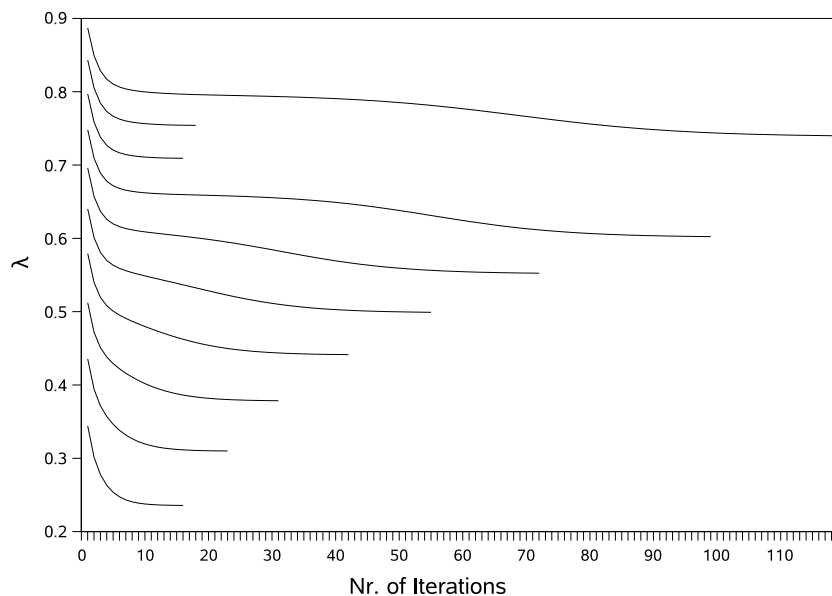
Figure 3. Convergence curves of the coupling constant λ for different ground-state energies ϵ for the Waxman algorithm using a 20×20 matrix as a model Hamiltonian which causes psuedoconvergence.
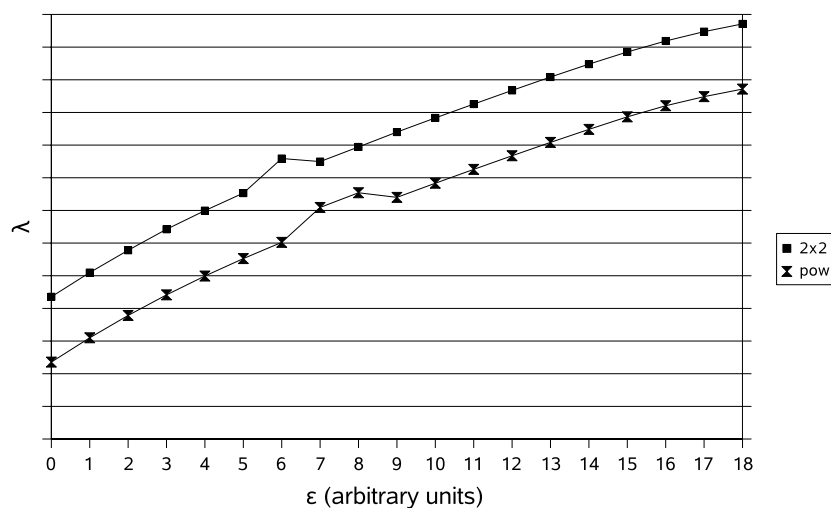

Figure 4. Dependence of the numerically determined coupling constant λ on the ground-state energy ϵ for the Waxman algorithm (power) and the modified algorithm (2×2) using a 20×20 matrix as a model Hamiltonian which causes pseudoconvergence. The two curves which are identical when full convergence is achieved have been vertically shifted apart from each other for better visibility.

In general, two problems arise with pseudoconvergence: how to detect and how to remedy it? The remedy in the present case may be to switch from the power to the 2×2 scheme. Since the 2×2 scheme itself suffers from pseudoconvergence, further measures as discussed in [4] may be appropriate.

The task of detecting pseudoconvergence is more difficult since if you look, e.g., at the rate of the change of the eigenvalue in the successive iteration steps, this can only indicate that there *may* be convergence. However, as has been shown recently for a wide class of potentials, smooth relations between ϵ and λ exist [3]. We therefore show the corresponding relations for the pseudoconvergence exhibited in figure 4. It can be seen that the power method exhibits marked deviations from the smooth behaviour at ϵ values 7 and 8. In figure 2, we can see that for these values the power method seemed to have converged rapidly. In fact, and we checked this with the exact results, pseudoconvergence occurred, which the algorithm had failed to detect. Similarly in figure 4 in the 2×2 scheme, one would suspect that pseudoconvergence occurred for $\epsilon = 6$ which is actually the case.

To summarize, the examples suggest that the modified algorithm has considerably improved convergence rates in general. In the case where pseudoconvergence occurs, the savings may become dramatic. The use of the Waxman algorithm enables one to detect the occurrence of pseudoconvergence reasonably quickly from the marked deviations from the smooth dependence of λ on ϵ. As in the power method scheme, the modified scheme does not require an explicit matrix representation in a large basis. Thus a major advantage of Waxman's method, namely the fact that the scheme can be applied directly to operators and wavefunctions in coordinate space, either using a discretized numerical or a parametrized analytical representation [3], is preserved when migrating to the modified algorithm.

References

[1] Waxman D 1998 *J. Phys. A: Math. Gen.* **31** 1329
[2] Andrew R A, Miller H G and Plastino A R 2006 *J. Phys. A: Math. Gen.* **39** L297
[3] Berger W A and Miller H G 2006 *J. Phys. A: Math. Gen.* **39** 14075
[4] Miller H G and Berger W A 1979 *J. Phys. A: Math. Gen.* **12** 1693
[5] Faddejew D K and Faddejewa W N 1964 *Numerische Methoden der linearen Algebra* (München: Oldenberg)
[6] Vorobyev U V 1962 *Moment Method in Applied Mathematics* (New York: Gordon and Breach)
[7] Berger W A, Miller H G and Kreuzer K G 1977 *J. Phys. A: Math. Gen.* **10** 1089
[8] Kreuzer K G, Miller H G, Dreizler R M and Berger W A 1980 *J. Phys. A: Math. Gen.* **12** 1645