Schwinger and anomaly-free Kohn variational principles and a generalized Lanczos algorithm for nonsymmetric operators

H.-D. Meyer
Theoretische Chemie, Physikalisch Chemisches Institut, Universität Heidelberg, Im Neuenheimer Feld 253, Federal Republic of Germany

J. Horáček
Faculty of Mathematics and Physics, Charles University, 1800 Prague 8, Czechoslovakia

L. S. Cederbaum
Theoretische Chemie, Physikalisch Chemisches Institute, Universität Heidelberg, Im Neuenheimer Feld 253, Federal Republic of Germany
(Received 26 September 1990)

The Schwinger variational principle is combined with the Lanczos algorithm. It is shown that the straightforward approach where the Lanczos basis is generated by $V-VG_0V$ is not very useful. The difficulties arise because of the unboundedness of the Schwinger operator $V-VG_0V$ for unbound potentials (such as, e.g., the Yukawa potential). A modification of the Schwinger operator is presented that resolves the problem. This approach leads to a *new* Lanczos recursion for nonsymmetric (and non-Hermitian) operators yielding nonorthogonal basis functions. The resulting approach is shown to be equivalent to the continued-fraction method of Horáček and Sasakawa. To illuminate the particular properties of the Schwinger-Lanczos basis we have applied it as well to other variational principles like the $\tilde{C}$ functional, the Newton variational principle, and the Kohn variational principle. To treat the multichannel problem we have adopted a modified band-Lanczos algorithm. This allows for a more efficient computation of the off-diagonal T-matrix elements than previous approaches.

I. INTRODUCTION

The Schwinger variational principle[1] (SVP) for computing T-matrix elements has found widespread application in particular when nonlocal interactions are present, as, e.g., in the field of electron-molecule scattering.[2,3] The computations are usually performed with linear variational parameters only, i.e., by introducing some basis set. Obviously the choice of this basis set is crucial and it seems very natural to combine the SVP with the Lanczos algorithm[4-6] because one expects the Lanczos vectors to yield an optimal basis set. It is the main purpose of this paper to investigate if and in what sense the Lanczos vectors provide an optimal basis set for the SVP. The combination of the SVP with the Lanczos algorithm is not new; it has been studied recently by Duneczky and Wyatt.[7] However, their approach is different from ours because they have adopted the more complicated biorthogonal Lanczos algorithm.

The present paper is organized as follows. In Sec. II we shall give a brief review on the Lanczos algorithm. In particular we investigate the convergence properties of the Lanczos algorithm when it is applied to compute the expectation value of an operator inverse. In Sec. III we combine the SVP and the Lanczos algorithm, and in Sec. IV we compare the Schwinger-Lanczos approach with the method of continued fractions, recently put forward by Horáček and Sasakawa.[8,9] In Sec. V we adopt the Schwinger-Lanczos basis set to evaluate the T matrix (or the K matrix) of other variational principles like the $\tilde{C}$ functional, the Newton variational principle, and the Kohn variational principle. In Sec. VI we introduce a band-Lanczos algorithm to treat the multichannel case, and in Sec. VII we finally conclude our investigation.

II. A BRIEF REVIEW ON THE LANCZOS METHOD

We want to apply the Lanczos algorithm[4-6] in order to compute the expectation value of an operator inverse, i.e., $\langle\phi|A^{-1}|\phi\rangle$. Here A shall be a Hermitian or a symmetric operator. We call an operator symmetric if its coordinate representation is symmetric. When A is symmetric (but non-Hermitian) we have to use the symmetric scalar product

$$\langle f|g\rangle=\int f(r)g(r)dr \tag{2.1}$$

(i.e., without complex conjugation) rather than the Hermitian one. This choice ensures that the matrix representation of a symmetric operator is complex symmetric. When A is Hermitian we use, of course, the usual Hermitian scalar product; i.e., the symmetry properties of A determine the choice of the scalar product.

If the operator inverse A^{-1} cannot be obtained analytically, one is forced to approximate A^{-1} by the inverse of some matrix representation of the operator A. We thus approximate the expectation value $\langle\phi|A^{-1}|\phi\rangle$ by

$$\langle\phi|A^{-1}|\phi\rangle \approx \sum_{i,j=1}^{N} \langle\phi|\chi_i\rangle(\langle\chi_i|A|\chi_j\rangle)^{-1}\langle\chi_j|\phi\rangle \ , \quad (2.2)$$

where $(\langle\chi_i|A|\chi_j\rangle)^{-1}$ is a shorthand notation for the (i,j)th matrix element of the inverse of the matrix representation of A. The choice of the basis set $\{|\chi_i\rangle\}$ is obviously crucial for the accuracy of this approximation.

The Lanczos algorithm provides us with a basis set which depends on the operator A as well as on the function ϕ. As shown below the Lanczos basis is an optimal basis set for solving our problem (compare also to Ref. 10). The first Lanczos vector, which is called the starting vector, is chosen to be just the normalized vector $|\phi\rangle$, i.e.,

$$|\chi_1\rangle = |\phi\rangle\langle\phi|\phi\rangle^{-1/2} \ . \quad (2.3)$$

The Lanczos recursion then reads

$$\beta_i|\chi_{i+1}\rangle = A|\chi_i\rangle - \alpha_i|\chi_i\rangle - \beta_{i-1}|\chi_{i-1}\rangle \ , \quad (2.4)$$

where, by definition, $\beta_0=0$. More explicitly, the recursion is performed as

$$|r_i\rangle = A|\chi_i\rangle - \beta_{i-1}|\chi_{i-1}\rangle \ , \quad (2.5a)$$

$$\alpha_i = \langle\chi_i|r_i\rangle \ , \quad (2.5b)$$

$$|s_i\rangle = |r_i\rangle - \alpha_i|\chi_i\rangle \ , \quad (2.5c)$$

$$\beta_i = \langle s_i|s_i\rangle^{1/2} \ , \quad (2.5d)$$

$$|\chi_{i+1}\rangle = \beta_i^{-1}|s_i\rangle \ . \quad (2.5e)$$

(Note that, if the Lanczos vectors are not needed for further computations, the recursion can be coded such that one needs to keep only two vectors in the fast memory.) The Lanczos vectors are orthogonal and they tridiagonalize the operator A, i.e.,

$$\langle\chi_i|\chi_j\rangle = \delta_{ij} \ , \quad (2.6)$$

$$\langle\chi_i|A|\chi_{i-1}\rangle = \beta_{i-1} \ , \quad (2.7a)$$

$$\langle\chi_i|A|\chi_i\rangle = \alpha_i \ , \quad (2.7b)$$

$$\langle\chi_i|A|\chi_{i+1}\rangle = \beta_i \ , \quad (2.7c)$$

$$\langle\chi_i|A|\chi_j\rangle = 0 \quad \text{if } |i-j| \geq 2 \ , \quad (2.7d)$$

or in a more compact form

$$\langle\chi_i|A|\chi_j\rangle = B_{ij} \ , \quad (2.8)$$

where B denotes the tridiagonal matrix where the α_i form the diagonal and the β_i the subdiagonal. The matrix $\underline{B}$ is real symmetric (complex symmetric) if A is Hermitian (complex symmetric). After N Lanczos recursions we will arrive at the $N\times N$ matrix

$$\underline{B}_N = \begin{pmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{N-1} \\ & & & \beta_{N-1} & \alpha_N \end{pmatrix} . \quad (2.9)$$

Next we define the finite rank operator A_N by

$$A_N = \sum_{i,j=1}^{N} |\chi_i\rangle\langle\chi_i|A|\chi_j\rangle\langle\chi_j| = \sum_{i,j=1}^{N} |\chi_i\rangle B_{ij}\langle\chi_j| \ . \quad (2.10)$$

This Lanczos approximation to the full operator A has the following remarkable properties:[10]

$$A_N^k|\chi_1\rangle = A^k|\chi_1\rangle \quad \text{for } 0\leq k\leq N-1 \ , \quad (2.11)$$

$$\langle\chi_1|A_N^k|\chi_1\rangle = \langle\chi_1|A^k|\chi_1\rangle \quad \text{for } 0\leq k\leq 2N-1 \ (N \text{ odd}) \ , \quad \text{for } 0\leq k\leq 2N-2 \ (N \text{ even}) \ . \quad (2.12)$$

Hence the $2N-2$ first moments of the operator A with respect to the starting vector $|\chi_1\rangle$ are exact within the Lanczos approximation. A (Hermitian or symmetric) approximation of rank N does not exist that provides more exact powers or exact moments than does the Lanczos approximation A_N. In this sense the Lanczos basis is an optimal basis set.

We return to our problems and find that the Lanczos approximation to the desired expectation value can now be written as

$$\langle\phi|A^{-1}|\phi\rangle = \langle\phi|\phi\rangle(\underline{B}_N^{-1})_{11} \ . \quad (2.13)$$

Note that the Lanczos vectors are not needed. The (1,1) element of the inverse of a tridiagonal matrix can be expressed as a continued fraction. In our case we have

$$(\underline{B}_N^{-1})_{11} = \cfrac{1}{\alpha_1 - \cfrac{\beta_1^2}{\alpha_2 - \cfrac{\beta_2^2}{\alpha_3 \cdots - \cfrac{\beta_{N-1}^2}{\alpha_N}}}} \ . \quad (2.14)$$

Hence the convergence properties of the Lanczos method can be related to the convergence properties of a continued fraction. The convergence properties of the Lanczos method are investigated in the Appendix. There we discuss which properties of the tridiagonal matirx $\underline{B}$ or the operator A guarantee a convergence. The speed of convergence is also discussed.

III. THE SCHWINGER-LANCZOS APPROACH

The T matrix determined by applying the SVP (Refs. 1–3) may be written as

$$T = \langle\phi|VP[P(V-VG_0V)P]^{-1}PV|\phi\rangle \ , \quad (3.1)$$

where ϕ denotes the free wave, V the interaction potential, G_0 the free Green's function, and P is a projector on some finite-dimensional subspace spanned by the basis set $\{|\chi_i\rangle\}$. The inversion of the operator $P(V-VG_0V)P$ is accomplished by the inversion of the matrix $\langle\chi_i|V-VG_0V|\chi_j\rangle$. The SVP provides the exact T matrix if the exact scattered wave $\Psi=\phi+G_0V\Psi$ is in the

space spanned by the basis set, i.e., if $P\Psi=\Psi$.

We now turn to the question of how to choose the basis set $\{|\chi_i\rangle\}$. This choice is obviously crucial for the accuracy of T in Eq. (3.1). As outlined above, a convenient choice is provided by the Lanczos basis. Hence we define the starting vector as

$$|\chi_1\rangle = V|\phi\rangle \cdot \langle\phi|V^2|\phi\rangle^{-1/2} , \tag{3.2}$$

and generate the other vectors $|\chi_2\rangle, |\chi_3\rangle, \ldots$ by the Lanczos recursion (2.4) and (2.5) setting

$$A = V - VG_0V . \tag{3.3}$$

Since G_0 and A are non-Hermitian but symmetric operators (in coordinate space), we adopt the symmetric scalar product (2.1) throughout the rest of the paper. Finally, after N Lanczos iterations we arrive at the following approximation for the T matrix:

$$T_N = \langle\phi|V^2|\phi\rangle (\underline{B}_N^{-1})_{11} \tag{3.4}$$

where, as before, $(\underline{B}_N^{-1})_{11}$ may be expressed as a continued fraction.

To continue our investigation we explicitly discuss specific forms for ϕ and G_0. We turn to the simplest case, s-wave scattering. Thus

$$\phi(r) = k^{-1}\sin kr , \tag{3.5}$$

$$G_0(r,r') = -2k^{-1}\sin kr_< \exp(ikr_>) , \tag{3.6}$$

where $r_<$ and $r_>$ denote the lesser and the greater of r and r', respectively. If one now assumes that the potential V has a Coulomb singularity at $r=0$, one finds that even the first moment

$$\begin{aligned}\alpha_1 &= \langle\chi_1|V - VG_0V|\chi_1\rangle \\ &= \langle\phi|V^3 - V^2G_0V^2|\phi\rangle / \langle\phi|V^2|\phi\rangle\end{aligned} \tag{3.7}$$

does not exist. More generally, the Lanczos method for $V - VG_0V$ works only if the potential is finite everywhere. But even when V is finite we do not expect this Lanczos method to be particularly useful. A physical interpretation of the method may help to understand this. The Lanczos vectors $|\chi_j\rangle$, $j=1,\ldots,N$, obviously span the same space as the vectors $(V - VG_0V)^{j-1}V|\phi\rangle$. Hence in this Schwinger-Lanczos method one implicitly approximates the exact wave function by

$$|\Psi\rangle \approx \sum_{j=1}^{N} c_j (V - VG_0V)^{j-1}V|\phi\rangle , \tag{3.8}$$

with variational constants c_j. This obviously is an ill-behaved expansion.

In order to overcome all these problems we recall that we want to compute the expectation value $\langle\chi_1|A^{-1}|\chi_1\rangle$. Introducing some symmetric operator D we may formally write

$$\begin{aligned}\langle\chi_1|A^{-1}|\chi_1\rangle &= \langle\chi_1|D(DAD)^{-1}D|\chi_1\rangle \\ &= \langle f_1|(DAD)^{-1}|f_1\rangle\end{aligned} \tag{3.9}$$

with

$$|f_1\rangle = D|\chi_1\rangle . \tag{3.10}$$

We now can choose D freely in order to make the operator DAD most suitable for the Lanczos method. The operator DAD must be bounded so that all its moments exist and its eigenvalues should be clustered in the way outlined in the Appendix. We will choose

$$D = V^{-1/2} \tag{3.11}$$

in the following and arrive at

$$DAD = 1 - V^{1/2}G_0V^{1/2} . \tag{3.12}$$

For a wide class of potentials (the so-called Rollnik class[11]) the operator $V^{1/2}G_0V^{1/2}$ is known to be a Hilbert-Schmidt operator,[12] i.e.,

$$\int |(V^{1/2}G_0V^{1/2})(r,r')|^2 dr\, dr' < \infty .$$

When we assume Eq. (3.6) for G_0 and further assume for simplicity that the potential V is local and that it obeys for some $\epsilon > 0$ the conditions

$$|V(r)| \le \text{const} \times r^{-2+\epsilon} \quad \text{as } r \to 0 , \tag{3.13a}$$

$$|V(r)| \le \text{const} \times r^{-2-\epsilon} \quad \text{as } r \to \infty , \tag{3.13b}$$

we can easily show that $V^{1/2}G_0V^{1/2}$ is a Hilbert-Schmidt operator. More generally, the condition

$$\int\int dr\, dr' r_> |V(r,r')| < \infty \tag{3.13c}$$

on the (radial) nonlocal potential $V(r,r')$ ensures that $V^{1/2}G_0V^{1/2}$ is a Hilbert-Schmidt operator. Hence, with the choice Eq. (3.11) for the operator D we have reached our goal. For a wide class of potentials all moments of DAD exist. Furthermore, the operator DAD has precisely the form which was found in the Appendix to be optimal for the Lanczos method.

Returning to the Lanczos algorithm, we now define the starting vector by

$$|f_1\rangle = V^{1/2}|\phi\rangle \langle\phi|V|\phi\rangle^{-1/2} . \tag{3.14}$$

We prefer to use $V^{1/2}G_0V^{1/2}$ as the Lanczos operator rather than $1 - V^{1/2}G_0V^{1/2}$. This choice for the Lanczos operator obviously does not change the Lanczos vectors. Hence we set

$$A = V^{1/2}G_0V^{1/2} \tag{3.15}$$

in the following and perform the Lanczos recursion as outlined in Eqs. (2.4)–(2.9) where, of course, the $|\chi_i\rangle$ are to be replaced by the $|f_i\rangle$. The resulting T matrix now reads

$$T = \cfrac{\langle\phi|V|\phi\rangle}{1-\alpha_1 - \cfrac{\beta_1^2}{1-\alpha_2 - \cfrac{\beta_2^2}{1-\alpha_3 - \cdots}}} . \tag{3.16}$$

The appearance of the operator $V^{1/2}$ within the Lanczos recursion is obviously very inconvenient. However, this operator can be eliminated by introducing a new

nonorthogonal basis set the vectors of which are defined as

$$|g_i\rangle = V^{-1/2}|f_i\rangle \tag{3.17}$$

and in particular

$$|g_1\rangle = |\phi\rangle\langle\phi|V|\phi\rangle^{-1/2} . \tag{3.18}$$

Since the $|f_i\rangle$ are orthonormal and tridiagonalize $V^{1/2}G_0V^{1/2}$, one finds that the $|g_i\rangle$ diagonalize V and tridiagonalize VG_0V, i.e.,

$$\langle g_i|V|g_j\rangle = \delta_{ij} , \tag{3.19}$$

$$\langle g_i|VG_0V|g_{i-1}\rangle = \beta_{i-1} , \tag{3.20a}$$

$$\langle g_i|VG_0V|g_i\rangle = \alpha_i , \tag{3.20b}$$

$$\langle g_i|VG_0V|g_{i+1}\rangle = \beta_i , \tag{3.20c}$$

$$\langle g_i|VG_0V|g_j\rangle = 0 \quad \text{for } |i-j| \geq 2 . \tag{3.20d}$$

From Eqs. (2.4) and (3.17) it now immediately follows that

$$\beta_{i+1}|g_{i+1}\rangle = G_0V|g_i\rangle - \alpha_i|g_i\rangle - \beta_{i-1}|g_{i-1}\rangle , \tag{3.21}$$

or explicitly

$$|r_i\rangle = G_0V|g_i\rangle - \beta_{i-1}|g_{i-1}\rangle , \tag{3.22a}$$

$$\alpha_i = \langle g_i|V|r_i\rangle , \tag{3.22b}$$

$$|s_i\rangle = |r_i\rangle - \alpha_i|g_i\rangle , \tag{3.22c}$$

$$\beta_i = \langle s_i|V|s_i\rangle^{1/2} , \tag{3.22d}$$

$$|g_{i+1}\rangle = \beta_i^{-1}|s_i\rangle . \tag{3.22e}$$

Note that the inconvenient operator $V^{1/2}$ has disappeared from all the working equations. Equations (3.21) and (3.22) can be interpreted as a generalized Lanczos algorithm for nonsymmetric operators. The vectors $|g_i\rangle$ generated by this generalized Lanczos algorithm are nonorthonormal and Eqs. (3.22b) and (3.22d) differ from Eqs. (2.5b) and (2.5d).

We want to emphasize that the generalized Lanczos algorithm is introduced for numerical convenience only. The generalized Lanczos algorithm subject to the operator G_0V and the (normal) Lanczos algorithm subject to the operator $V^{1/2}G_0V^{1/2}$ yield exactly the same tridiagonal matrix $\underline{B}$. In particular, the expression for the T matrix Eq. (3.16) remains valid as well as our considerations on the convergence of the Lanczos method. It is, however, convenient to reconsider the physical argument given above [cf. Eq. (3.8)] for the new basis set $\{|g_i\rangle\}$. The exact wave function $|\Psi\rangle$ is now expanded in the functions $|g_j\rangle$, which in turn span the same space as $(G_0V)^{j-1}|\phi\rangle$. Hence the SVP approximates Ψ by

$$|\Psi\rangle \approx \sum_{j=1}^{N} c_j(G_0V)^{j-1}|\phi\rangle , \tag{3.23}$$

where the c_j are variational parameters. The approximation (3.23) can be interpreted as a combination of perturbative and variational methods, which is known to work very well even in the case of strong potentials.[13]

The generalized Lanczos algorithm may have a much broader range of applications then just the SVP. Assume that one wants to diagonalize the non-Hermitian (nonsymmetric) operator A. We first seek for a Hermitian (symmetric) operator S such that

$$SA = A^{\dagger}S \tag{3.24}$$

holds, i.e., SA is Hermitian (symmetric). Recall that $A^{\dagger}$ is defined such that $\langle\phi|A\Psi\rangle = \langle A^{\dagger}\phi|\Psi\rangle$ holds for all $|\Psi\rangle$ and $|\phi\rangle$ out of the domain of A. Hence the meaning of $A^{\dagger}$ depends on the particular scalar product chosen (Hermitian or symmetric). From Eq. (3.24) it follows immediately that the operator C

$$C = S^{1/2}AS^{-1/2} \tag{3.25}$$

is also Hermitian (symmetric). The above equation is a similarity transformation and, as is well known, the operators A and C thus have identical eigenvalues. Furthermore, $S^{-1/2}$ times an eigenvector of C gives a (right-hand side) eigenvector of A. To compute the spectrum of the operator A one may apply the Lanczos algorithm to the Hermitian (symmetric) operator C. The key point now is that the complicated operator C does not appear in the working equations when the generalized Lanczos algorithm is adopted [see Eqs. (3.21) and (3.22) and replace G_0V by A and V by S]. But there still remains an unresolved problem: How can one find an operator S which satisfies Eq. (3.24)? For the special case that A is given as a product of two Hermitian (symmetric) operators, $A = A_1 \cdot A_2$, one finds that $S = A_2$ is a solution to Eq. (3.24). However, it is not clear to us how to find a solution to Eq. (3.24) in general.

IV. COMPARISON TO THE METHOD OF CONTINUED FRACTIONS

Horáček and Sasakawa have recently derived an expression for the T matrix[8,9] which, similar to the Schwinger-Lanczos approach, expresses the T matrix as a continued fraction. The so-called method of continued fractions[7,8] (MCFV) starts with introducing a family of potentials and wave functions which are recursively defined as

$$V_0 = 0 , \tag{4.1a}$$

$$V_1 = V , \tag{4.1b}$$

$$V_{i+1} = V_i - \frac{V_i|u_i\rangle\langle u_i|V_i}{\langle u_i|V_i|u_i\rangle} ; \tag{4.1c}$$

and

$$|u_1\rangle = |\phi\rangle , \tag{4.2a}$$

$$|u_{i+1}\rangle = G_0V_i|u_i\rangle . \tag{4.2b}$$

The authors finally arrived at the following continued fraction for the T matrix:

$$T=\frac{\langle u_1|V_1|u_1\rangle^2}{\langle u_1|V_1|u_1\rangle-\langle u_1|V_1|u_2\rangle-\dfrac{\langle u_2|V_1|u_2\rangle^2}{\langle u_2|V_2|u_2\rangle-\langle u_2|V_2|u_3\rangle-\cdots}} . \tag{4.3}$$

In the following we shall show that the MCFV approach is formally equivalent to the Schwinger-Lanczos method developed in the Sec. III, Eqs. (3.16)–(3.22). We first note that

$$V_j|u_i\rangle=0 \quad \text{for } 1\le i\le j-1 \tag{4.4}$$

holds because of the definition (4.1c). This in turn allows us to write

$$V|u_i\rangle=V_i|u_i\rangle+V_{i-1}|u_{i-1}\rangle\frac{\langle u_{i-1}|V_{i-1}|u_i\rangle}{\langle u_{i-1}|V_{i-1}|u_{i-1}\rangle}+V_{i-2}|u_{i-2}\rangle\frac{\langle u_{i-2}|V_{i-2}|u_i\rangle}{\langle u_{i-2}|V_{i-2}|u_{i-2}\rangle} . \tag{4.5}$$

To prove the equivalence we first note that the two basis sets $\{|u_i\rangle\}$ and $\{|g_i\rangle\}$ are related by

$$V_i|u_i\rangle=a_iV|g_i\rangle , \tag{4.6}$$

where the constants a_i are recursively defined as

$$a_1=\langle u_1|V|u_1\rangle^{1/2} , \tag{4.7a}$$

$$a_{i+1}=a_i\beta_i . \tag{4.7b}$$

We will finally show that the generalized Lanczos recursion follows from the recursion (4.2b). Multiplying Eq. (4.6) with G_0 yields

$$|u_{i+1}\rangle=a_iG_0V|g_i\rangle \tag{4.8}$$

and hence

$$\begin{aligned}\langle u_i|V_i|u_i\rangle&=a_i\langle g_i|V|u_i\rangle\\&=a_ia_{i-1}\langle g_i|VG_0V|g_{i-1}\rangle\\&=a_i^2 .\end{aligned} \tag{4.9a}$$

Similarly one obtains

$$\langle u_i|V_i|u_{i+1}\rangle=\alpha_ia_i^2 , \tag{4.9b}$$

$$\langle u_i|V_i|u_{i+2}\rangle=a_{i+1}^2 , \tag{4.9c}$$

$$\langle u_i|V_i|u_{i+k}\rangle=0 \quad \text{if } k>2 . \tag{4.9d}$$

Using these equations as well as Eq. (4.7) one can easily show that the two expressions for the T matrix [Eqs. (3.16) and (4.3)] are identical. Moreover, it now follows from Eqs. (4.5) and (4.9) that

$$V|u_{i+1}\rangle=V_{i+1}|u_{i+1}\rangle+\alpha_iV_i|u_i\rangle+\beta_{i-1}^2V_{i-1}|u_{i-1}\rangle \tag{4.10}$$

holds. Multiplying this equation with G_0 yields

$$G_0V|u_{i+1}\rangle=a_{i+1}G_0V|g_{i+1}\rangle+a_i\alpha_iG_0V|g_i\rangle+a_{i-1}\beta_{i-1}^2G_0V|g_{i-1}\rangle \tag{4.11}$$

or

$$|u_{i+1}\rangle=a_i(\beta_i|g_{i+1}\rangle+\alpha_i|g_i\rangle+\beta_{i-1}|g_{i-1}\rangle) . \tag{4.12}$$

With the aid of Eq. (4.8) we arrive at

$$G_0V|g_i\rangle=\beta_i|g_{i+1}\rangle+\alpha_i|g_i\rangle+\beta_{i-1}|g_{i-1}\rangle , \tag{4.13}$$

which is the desired generalized Lanczos recursion (3.21).

Although the Schwinger-Lanczos method and the MCFV approach are equivalent, they may behave differently in a numerical implementation. We believe that the Schwinger-Lanczos method shows a better numerical performance, in particular (selective) reorthonormalization[5,6] can there be easily incorporated if necessary. Moreover, the Schwinger-Lanczos recursion itself seems to be simpler than the MCFV recursion. However, we have not conducted a careful numerical comparison of the two methods.

V. OTHER VARIATIONAL PRINCIPLES

In this section we want to apply our Schwinger-Lanczos basis $\{|g_i\rangle\}$ to other variational principles such as the $\tilde{C}$ functional,[3,14] the Newton variational principle[15] (NVP), and the Kohn variational principle[16–18] (KVP). The former two variational principles are—similar to the SVP—based on the Lippmann-Schwinger equation. The KVP, on the other hand, is based on the Schrödinger equation. The application of the KVP requires that one is able to evaluate the matrix elements of the Hamiltonian.

Let us first turn to the $\tilde{C}$ functional. The T matrix and the full Green's function may be written as

$$T=V+VGV \tag{5.1}$$

and

$$G=G_0+G_0TG_0 . \tag{5.2}$$

Combining these two equations and replacing T on the right-hand side by its SVP expression yields

$$T=V+VG_0V+VG_0V(V-VG_0V)^{-1}VG_0V . \tag{5.3}$$

This becomes the $\tilde{C}$-functional expression, if the inversion of the operator $V-VG_0V$ is performed within a basis set. Using the Schwinger-Lanczos basis [cf. Eqs. (3.14)–(3.22)], we obtain

$$\begin{aligned}T_N^C&=\langle\phi|V|\phi\rangle[\underline{1}+\underline{B}_N+\underline{B}_N(1-\underline{B}_N)^{-1}\underline{B}_N]_{11}\\&=\langle\phi|V|\phi\rangle(\underline{1}-\underline{B}_N)_{11}^{-1}\\&=T_N^S .\end{aligned} \tag{5.4}$$

Hence the $\tilde{C}$ functional and the SVP yield exactly the same result for the T matrix when the Schwinger-

Lanczos basis is used. The superscripts C and S refer to the $\tilde{C}$ functional and the SVP, respectively.

The NVP is derived from the exact expression of the T matrix

$$T=V+VG_0V(VG_0V-VG_0VG_0V)^{-1}VG_0V \tag{5.5}$$

by performing the inversion within a basis set. When applying our Lanczos-Schwinger basis we first note that

$$\langle g_i|VG_0VG_0V|g_j\rangle=\sum_{k=1}^{N+1}B_{ik}B_{kj}$$
$$=(\underline{B}_N^2+\underline{Q}_N\beta_N^2)_{ij} \tag{5.6}$$

holds for $1\le i,j\le N$. Here $\underline{Q}_N$ denotes a $N\times N$ matrix, the (N,N) element of which is unity and all other elements of which are zero. The NVP expression for the T matrix now becomes

$$T_N^N=\langle\phi|V|\phi\rangle[\underline{1}+\underline{B}_N(\underline{B}_N-\underline{B}_N^2-\underline{Q}_N\beta_N^2)^{-1}\underline{B}_N]_{11}\ , \tag{5.7a}$$

i.e., the T matrix of the NVP is given entirely in terms of the elements of the Lanczos matrix $\underline{B}$. Developing the inverse with respect to β_N^2 yields

$$T_N^N=T_N^S+a_1^2(\underline{1}-\underline{B}_N)_{1N}^{-1}\beta_N^2(\underline{1}-\underline{B}_N)_{N1}^{-1}+0(\beta_N^4)\ . \tag{5.7b}$$

Hence the NVP result differs from the SVP result, but T_N^N is very close to T_{N+1}^S. This follows from Eq. (A6a). Since we now invert $\underline{1}-\underline{B}$ rather than $\underline{B}$ itself, we have to replace in Eq. (A6a) α_{N+1} by $1-\alpha_{N+1}$, which, in turn, is close to 1. Hence $T_N^N=T_{N+1}^S+0(\beta_N^4)+O(\beta_N^2\alpha_{N+1})$. This is a quite reasonable result because one implicitly performs one extra iteration due to the operator VG_0VG_0V when using the NVP.

Let us now turn to the KVP and begin with considering a particular variant of this variational principle, namely the complex Kohn method.[18] This method can be derived from the exact equation

$$T=V+VGV \tag{5.8}$$

by approximating the full Green's function G by a finite rank expression which is obtained by simply inverting some matrix representation of $E-H$. Surprisingly, such a brute-force approach to the full Green's function can be shown[18] to be variational stable, provided the basis functions satisfy certain (energy-dependent) boundary conditions. The Schwinger-Lanczos functions $|g_j\rangle$ do not satisfy the required boundary conditions. However, the closely related functions $|u_j\rangle$ defined in Sec. IV do so. Assuming Eqs. (3.5) and (3.6), and using quantities defined in Sec. IV, one finds

$$u_1(r)=k^{-1}\sin(kr)\ , \tag{5.9a}$$

$$u_2(r)\underset{r\to\infty}{\sim}-a_1^2\exp(ikr)\ , \tag{5.9b}$$

$$u_j(r)=\underset{r\to\infty}{\longrightarrow}0\quad\text{for } j\ge 3\ . \tag{5.9c}$$

Equation (5.9c) follows because

$$u_{j+1}(r)=\langle r|G_0V_j|u_j\rangle\to-2\exp(ikr)\langle u_1|V_j|u_j\rangle=0$$

[cf. Eq. (4.4)]. Remember that the functions $|u_j\rangle$ and $|g_j\rangle$ are related to each other by a simple transformation [see Eq. (4.12)]; *they span exactly the same space.* To avoid the inconvenient normalization factors a_j, however, we prefer to introduce the functions $|\tilde{u}_j\rangle$, which are defined as

$$|\tilde{u}_1\rangle=a_1^{-1}|u_1\rangle\ , \tag{5.10a}$$

$$|\tilde{u}_j\rangle=a_{j-1}^{-1}|u_j\rangle\quad\text{for } j>1\ . \tag{5.10b}$$

With these renormalized functions Eqs. (4.8) and (4.12) simplify to

$$|\tilde{u}_{j+1}\rangle=G_0V|g_j\rangle \tag{5.10c}$$

and

$$|\tilde{u}_1\rangle=|g_1\rangle\ , \tag{5.11a}$$

$$|\tilde{u}_{j+1}\rangle=\beta_j|g_{j+1}\rangle+\alpha_j|g_j\rangle+\beta_{j-1}|g_{j-1}\rangle$$
$$=\sum_k B_{jk}|g_k\rangle\ . \tag{5.11b}$$

To evaluate of the matrix elements of the Hamiltonian H we write $H=H_0+V$ and assume that

$$(H_0-E)|\tilde{u}_1\rangle=0 \tag{5.12}$$

and

$$(H_0-E)G_0=-1 \tag{5.13}$$

hold. Hence

$$(H_0-E)|\tilde{u}_{j+1}\rangle=(H_0-E)G_0V|g_j\rangle=-V|g_j\rangle \tag{5.14}$$

and with the aid of Eq. (5.11) the following results are obtained:

$$\langle\tilde{u}_j|H_0-E|\tilde{u}_1\rangle=0\ , \tag{5.15}$$

$$\langle\tilde{u}_1|H_0-E|\tilde{u}_j\rangle=-\delta_{j2}\ , \tag{5.16}$$

$$\langle\tilde{u}_{i+1}|H_0-E|\tilde{u}_{j+1}\rangle=-B_{ij}\ . \tag{5.17}$$

Similarly

$$\langle\tilde{u}_1|V|\tilde{u}_1\rangle=1\ , \tag{5.18}$$

$$\langle\tilde{u}_1|V|\tilde{u}_{j+1}\rangle=\langle\tilde{u}_{j+1}|V|\tilde{u}_1\rangle=B_{1j}\ , \tag{5.19}$$

$$\langle\tilde{u}_{i+1}|V|\tilde{u}_{j+1}\rangle=\sum_k B_{ik}B_{kj}\ . \tag{5.20}$$

We now consider the basis $\{|\tilde{u}_j\rangle\}$ with $1\le j\le N+1$. Because in the internal sum of Eq. (5.20) the term $k=N+1$ may also contribute, we find, similar to above,

$$\langle\tilde{u}_{i+1}|H-E|\tilde{u}_{j+1}\rangle=-(\underline{B}_N-\underline{B}_N^2-\underline{Q}_N\beta_N^2)_{ij}\ ,$$
$$1\le i,j\le N\ . \tag{5.21}$$

Within the complex-Kohn method[18] we now obtain

$$T^K_{N+1}=a_1^2\left[\langle\tilde{u}_1|V|\tilde{u}_1\rangle-\sum_{i,j=1}^{N}\langle\tilde{u}_1|V|\tilde{u}_{i+1}\rangle(\langle\tilde{u}_{i+1}|H-E|\tilde{u}_{j+1}\rangle)^{-1}\langle\tilde{u}_{j+1}|V|\tilde{u}_1\rangle\right]$$
$$=a_1^2[\underline{1}+\underline{B}_N(\underline{B}_N-\underline{B}_N^2-\underline{Q}_N\beta_N^2)^{-1}\underline{B}_N]_{11}=T^N_N\ . \tag{5.22}$$

Hence the complex Kohn method[18] gives exactly the same result as the NVP, provided one uses one Schwinger-Lanczos basis function more in the KVP as compared to the NVP.

The original version of the KVP (Refs. 16 and 17) starts from a variationally stable expression for the tangent of the phase shift, i.e., for the K matrix. This variationally stable expression involves the Hamiltonian H, but no Green's functions. We shall call this original version of the KVP the real KVP because it uses only real trial functions. The final working equation of the real KVP is virtually identically to the first part of Eq. (5.22) except for the fact that T should be replaced by $-\pi^{-1}K$ and that the functions $|\tilde{u}_j\rangle$ should be real functions subject to certain real boundary conditions.

To make contact with the real KVP we replace the free Green's function G_0 by the principle value Green's function $G_0^{(P)}$,

$$G_0^{(P)}(r,r')=-2k^{-1}\sin kr_<\cos kr_>\ , \tag{5.23}$$

and find that the recursion now yields real functions $|g_i\rangle$, $|u_i\rangle$ or $|\tilde{u}_i\rangle$. The matrix $\underline{B}$ now becomes real symmetric. If we consistently replace G_0 by $G_0^{(P)}$ and T by $-\pi^{-1}K$, then virtually all equations of this paper remain unchanged. In particular, we have

$$K^K_{N+1}=K^N_N\ , \tag{5.24}$$

i.e., the K matrix of the KVP is identical to the K matrix of the NVP (except for a different size of the Schwinger-Lanczos basis). It is well known that the (real) KVP is plagued by anomalies, i.e., by spurious singularities of the K matrix. Using the Schwinger-Lanczos basis, however, we find, that there are no anomalies because the NVP is anomaly-free. Hence we arrive at the quite remarkable result that the Schwinger-Lanczos basis leads to an anomaly-free KVP even when using real functions. This emphasizes again the unique features of the Schwinger-Lanczos basis set. In the conventional application of the KVP one has used basis functions which are energy independent except for the first two functions (which are required to asymptotically behave like $\sin kr$ and $\cos kr$). Within the Schwinger-Lanczos basis set, however, all functions are naturally energy dependent. This leads to a matrix representation of $H-E$, which is free of spurious singularities.

VI. THE MULTICHANNEL PROBLEM

The computation of diagonal T-matrix elements is sufficient for the study of the elastic scattering off spherical symmetric potentials. For most other problems, such as, e.g., the electron-molecule scattering problem, off-diagonal T-matrix elements are needed. Our present approach can be fairly straightforwardly generalized to evaluate off-diagonal T-matrix elements. We may write

$$\langle\phi_f|T|\phi_i\rangle=\langle\phi_f|V(V-VG_0V)^{-1}V|\phi_i\rangle$$
$$=\sum_{k,l}\langle\phi_f|V|g_k\rangle(\underline{1}-\underline{B})^{-1}_{kl}\langle g_l|V|\phi_i\rangle$$
$$=\sum_{k}\langle\phi_f|V|g_k\rangle(\underline{1}-\underline{B})^{-1}_{kl}\langle\phi_i|V|\phi_i\rangle^{1/2}\ . \tag{6.1}$$

Here $|\phi_f\rangle$ and $|\phi_i\rangle$ denote the final and initial free wave. The Lanczos vectors $|g_k\rangle$ as well as the tridiagonal matrix $\underline{B}$ are defined as above, where the initial free wave $|\phi_i\rangle$ is adopted to determine the starting vector $|g_1\rangle$ [cf. Eq. (3.18)]. In a true multichannel case (i.e., a rearrangement collision) one must replace the potential V and the free Green's function G_0 by the appropriate channel potential and channel Green's function, respectively. We will ignore this subtlety and consider only one arrangement channel as it is the case for, e.g., electronically elastic electron-molecule scattering.

The procedure of Eq. (6.1) has several shortcomings. Most important is the asymmetric treatment with respect to the initial and final state. As a consequence the T matrix becomes asymmetric; symmetry is reached only when the Lanczos algorithm is fully converged. The Lanczos iterative scheme is now more complicated. The matrix elements $\langle\phi_f|V|g_k\rangle$ have to be computed and stored when generating the Schwinger-Lanczos vectors $|g_k\rangle$. Furthermore, the whole first column of the matrix $(\underline{1}-\underline{B})^{-1}$ is now needed rather than only the (1,1) element. Thus the T matrix can no longer be represented by a simple continued fraction and our considerations on the convergence of the Lanczos method (see Sec. II and the Appendix) no longer apply. It has been shown elsewhere[10] that the computation of the off-diagonal T-matrix elements by Eq. (6.1) requires about twice as many iterations as the computation of diagonal matrix elements when the same degree of convergence is desired. Note, finally, that the Lanczos procedure is to be performed separately for each possible initial state. Each set of Lanczos iterations yields one column of the T matrix.

There is more efficient way to tackle the multichannel problem than the one offered by Eq. (6.1). One can compute all desired matrix elements of the T matrix at once by adopting the band-Lanczos algorithm.[5,10,19] Let us denote by $|\phi_1\rangle,\ldots,|\phi_n\rangle$ the states with respect to which we want to know the T-matrix elements. These functions may, e.g., represent the free partial waves $|E,l,m\rangle$ for various angular momenta (l,m). The band- (or block-) Lanczos algorithm takes all these n states into account simultaneously because it employs a set of n starting vectors which span the same space as the physical states $|\phi_j\rangle$.

To arrive at an appropriate set of starting vectors, which have to be orthonormal with respect to the poten-

tial, we generalize the Gram-Schmidt orthogonalization procedure[5] simply by replacing there the scalar product $\langle\ |\ \rangle$ by $\langle\ |V|\ \rangle$. Doing so one computes an upper triangular matrix $\underline{R}$ and a set of starting vectors $|g_1\rangle,\ldots,|g_n\rangle$ such that

$$|\phi_j\rangle=\sum_{i=1}^{n}|g_i\rangle R_{ij},\quad j=1,\ldots,n \tag{6.2}$$

and

$$\langle g_i|V|g_j\rangle=\delta_{ij},\quad i,j=1,\ldots,n \tag{6.3}$$

hold. The use of the Gram-Schmidt procedure is not crucial; one may adopt any other orthogonalization procedure. It is only important that the starting vectors are orthonormal with respect to the potential [cf. Eq. (6.3)] and that they span the same space as the physical states $|\phi_j\rangle$, i.e., that Eq. (6.2) holds for some matrix $\underline{R}$. Note that Eqs. (6.2) and (6.3) are a generalization of Eq. (3.18) to many dimensions.

Generalizing now the band-Lanczos recursion[5,10] to nonsymmetric operators and nonorthogonal states, similar to the way in which we have generalized above the simple Lanczos recursion, we arrive at the recursion

$$B_{i,i+n}|g_{i+n}\rangle=G_0V|g_i\rangle-\sum_{j=\max(i-n,1)}^{i+n-1}B_{ij}|g_j\rangle\ , \tag{6.4}$$

where

$$B_{ij}=\langle g_i|VG_0V|g_j\rangle\ , \tag{6.5a}$$

$$B_{ij}=B_{ji}\ , \tag{6.5b}$$

$$B_{ij}=0\ \text{ if }\ |i-j|\geq n+1\ . \tag{6.5c}$$

Hence the matrix $\underline{B}$ is a banded complex symmetric matrix with n subdiagonals on each side of the diagonal. The explicit working equations read in a computer adapted language

$$|r\rangle:=G_0V|g_i\rangle-\sum_{j=\max(i-n,1)}^{i-1}B_{ij}|g_j\rangle\ . \tag{6.6a}$$

For $j=i,\ldots,i+n-1$ do

$$B_{ij}:=\langle r|V|g_j\rangle\ , \tag{6.6b}$$

$$|r\rangle:=|r\rangle-B_{ij}|g_j\rangle\ , \tag{6.6c}$$

end do

$$B_{i,i+n}:=\langle r|V|r\rangle^{1/2}\ , \tag{6.6d}$$

$$|g_{i+n}\rangle:=|r\rangle B_{i,i+n}^{-1}\ . \tag{6.6e}$$

Note that one needs to keep only the last $2n$ Lanczos vectors and one scratch vector in the fast memory.

We now turn to the Schwinger expression for the T matrix and adopt the band-Lanczos basis $|g_i\rangle$ to invert the Schwinger operator $V-VG_0V$. We find

$$\begin{aligned}T_{ij}&=\langle\phi_i|V(V-VG_0V)^{-1}V|\phi_j\rangle\\&=\sum_{k,l}\langle\phi_i|V|g_k\rangle(\langle g_k|V-VG_0V|g_l\rangle)^{-1}\langle g_l|V|\phi_j\rangle\\&=\sum_{k,l=1}^{n}R_{ki}(\underline{1}-\underline{B})^{-1}_{kl}R_{lj}\ .\end{aligned} \tag{6.7}$$

Rather than to invert $(\underline{1}-\underline{B})$ one may solve n sets of linear equations

$$\sum_{k=1}^{N}(\delta_{lk}-B_{lk})X_{kj}=R_{lj},\quad j=1,\ldots,n \tag{6.8}$$

and find $T_{ij}=\sum R_{ki}X_{kj}$. An alternative way is to evaluate the inverse of the banded matrix $\underline{1}-\underline{B}$ by a matrix continued fraction, e.g., like in Ref. 19. The use of Eq. (6.7) has several advantages over the use of Eq. (6.1). Equation (6.7) is faster convergent with respect to the number of Lanczos iterations[10] and it treats diagonal and off-diagonal T-matrix elements on the same footing. The only disadvantage of the band-Lanczos method is its requirement of a larger working space in the computer memory.

VII. CONCLUSION

In this paper we have combined the Schwinger variational principle with the Lanczos algorithm. We have given arguments that one should not use the Schwinger operator $V-VG_0V$ as the Lanczos operator, but rather $1-V^{1/2}G_0V^{1/2}$ (or, equivalently, $V^{1/2}G_0V^{1/2}$). The inconvenient operator $V^{1/2}$ can easily be removed from all working equations. We have called the resulting approach the generalized Lanczos algorithm. The generalized Lanczos basis is nonorthogonal and it tridiagonalizes the Schwinger operator $V-VG_0V$. The combination of the Schwinger variational principle with the Lanczos algorithm has recently been investigated and applied by Dunezky and Wyatt.[7] These authors, however, have used the more complicated biorthonormal Lanczos algorithm. The need for the biorthogonal form arises quite naturally because they used the nonsymmetric operator G_0V to derive the Lanczos recursion. In the present investigation we start from the symmetric operator $V^{1/2}G_0V^{1/2}$. The usual Lanczos algorithm for this operator boils down straighforwardly to a recursion involving the operator G_0V. Rather than keeping separately left- and right-hand Lanczos vectors we use a single set of nonorthonormal vectors. Hence the new method is numerically less demanding compared to the one of Duneczky and Wyatt.

The Schwinger-Lanczos algorithm derived in this paper could be shown to be equivalent to the very efficient MCFV approach of Horáček and Sasakawa.[8,9] Hence the present investigation also provides a new and vivid interpretation of the MCFV approach. We believe that the use of the present approach rather than the MCFV is of advantage numerically. However, no detailed comparison on the numerical performance was performed.

We have further applied the Schwinger-Lanczos basis to solve other variational principles such as the $\tilde{C}$ functional, the Newton variational principle, or the Kohn variational principle. Within this basis set it was found that the Schwinger variational principle and the $\tilde{C}$ functional give the same result for the T matrix. Within the same basis set the result of the Kohn variational principle (provided one uses one basis function more) is found to be

identical to the one of the Newton variational principle. This equivalence remains valid when turning from the T matrix and complex basis functions to the K matrix and real basis functions. Hence the (real) Kohn variational principle, which is known to be plagued by anomalities, turns out to be anomaly-free when the Schwinger-Lanczos basis is used.

The generalization of the Schwinger-Lanczos approach to the multichannel case leads quite naturally to the adoption of a generalized band-Lanczos algorithm. This provides a new result which is different from the multichannel MCFV treatment. Our findings on the performance of the Schwinger-Lanczos basis when applied to other variational principles (see Sec. V) carries over to the multichannel case as well. The multichannel Kohn variational principle is thus anomaly-free when the Schwinger–band-Lanczos basis is used.

Finally, two technical remarks are in order. First, one may replace the free wave and the free Green's function by a distorted wave and a distorted-wave Green's function.[20–22] The potential V then denotes the residual interaction. Such a modification will speed up the convergence. Second, in an actual calculation one may wish to replace the T matrix by the K matrix and the free Green's function by the principle value Green's function, similar to what we have done at the end of Sec. V. By doing so all appearing quantities become real, which considerably simplifies the numerical treatment. In our analytic investigations, however, we preferred to work with the T matrix because the K matrix may be unbounded.

ACKNOWLEDGMENTS

Financial support by the Deutsche Forschungsgemeinschaft through the Sonderforschungsbereich No. 91 is gratefully acknowledged.

APPENDIX

In this appendix we investigate the convergence properties of the Lanczos algorithm used to compute the expectation value of an operator inverse, i.e., $\langle\phi|A^{-1}|\phi\rangle$. We start our investigation by assuming that the infinite tridiagonal Lanczos matrix $\underline{B}$ is known. It is then clear (see Sec. II) that the method converges, if the continued fraction

$$(\underline{B}^{-1})_{11}=\cfrac{1}{\alpha_1-\cfrac{\beta_1^2}{\alpha_2-\cfrac{\beta_2^2}{\alpha_3\cdots}}} \tag{A1}$$

converges. Terminating this continued fraction after the Mth term rather than after the Nth term ($M>N$) obviously has the effect of replacing α_N in Eq. (2.14) by

$$\alpha_N-\cfrac{\beta_N^2}{\alpha_{N+1}-\cdots\cfrac{\beta_{M-1}^2}{\alpha_M}} . \tag{A2}$$

Hence $(\underline{B}_N^{-1})_{11}$ will be close to $(\underline{B}_M^{-1})_{11}$ if $(\underline{B}_N^{-1})_{11}$ depends only weakly on α_N. A brief calculation yields

$$\frac{d}{d\alpha_N}(\underline{B}_N^{-1})_{11}=\left[\frac{\beta_1\beta_2\cdots\beta_{N-1}}{\det\underline{B}_N}\right]^2 . \tag{A3}$$

The determinants of the various $\underline{B}_N$ are connected by the recursion

$$\det\underline{B}_{N+1}=\alpha_{N+1}\underline{B}_N-\beta_N^2\det\underline{B}_{N-1} . \tag{A4}$$

Another estimate for the truncation error can be obtained by the partitioning technique[23]

$$(\underline{B}_{N+1}^{-1})_{1,1}=(\underline{B}_N-\underline{Q}_N\beta_N^2/\alpha_{N+1})_{11}^{-1} , \tag{A5}$$

where $\underline{Q}_N$ denotes an $N\times N$ matrix, the matrix elements of which are all vanishing except for the (N,N) element, which is unity. Developing the right-hand side into a Taylor series one obtains

$$(\underline{B}_{N+1}^{-1})_{11}=(\underline{B}_N^{-1})_{11}+(\underline{B}_N^{-1})_{1N}\frac{\beta_N^2}{\alpha_{N+1}}(\underline{B}_N^{-1})_{N1}+O(\beta_N^4) \tag{A6a}$$

$$=(\underline{B}_N^{-1})_{11}+\frac{(\beta_1\cdots\beta_N)^2}{\alpha_{N+1}(\det\underline{B}_N)^2}+O(\beta_N^4) . \tag{A6b}$$

It is clear that the continued fraction converges quickly if the β's are small compared to the α's. Thus, in order to proceed, let us assume that the β_j satisfy the inequality

$$|\beta_j|^2\le c^2|\alpha_j\alpha_{j+1}| \tag{A7}$$

for all j and for some $0<c<\frac{1}{2}$. One can then show that

$$|\det\underline{B}_N|\ge|\alpha_1\alpha_2\cdots\alpha_N|\prod_{j=1}^{N}\left[1-2c\cos\frac{\pi j}{N+1}\right]$$
$$\ge|\alpha_1\alpha_2\cdots\alpha_N|(1-4c^2)^{N/2} , \tag{A8}$$

hence

$$\left|\frac{d}{d\alpha_N}(\underline{B}_N^{-1})_{11}\right|\le\frac{1}{c^2|\alpha_1\alpha_N|}\left[\frac{c^2}{1-4c^2}\right]^N , \tag{A9}$$

and finally

$$|(\underline{B}_N^{-1})_{11}-(\underline{B}_\infty^{-1})_{11}|\le\frac{2}{|\alpha_1|}\left[\frac{c^2}{1-4c^2}\right]^N . \tag{A10}$$

Hence for $c^2<\frac{1}{5}$ we have proved the convergence. It is clear that the convergence is also guaranteed when the relation (A7) holds for $j\ge N_0$ only. The above equation is easily generalized to

$$|(\underline{B}_N^{-1})_{11}-(\underline{B}_\infty^{-1})_{11}|$$
$$\le\left|\frac{\beta_1\beta_2\cdots\beta_{N_0-1}}{\det\underline{B}_{N_0}}\right|^2 2|\alpha_{N_0}|c^2\left[\frac{c^2}{1-4c^2}\right]^{N-N_0} , \tag{A11}$$

which holds when the conditions $N>N_0$ and $|\beta_j|^2\le c^2|\alpha_j\alpha_{j+1}|$ for $j\ge N_0$ are satisfied.

We have shown that the Lanczos method converges if

the off-diagonal matrix elements of the Lanczos matrix are sufficiently small as compared to the diagonal ones. However, more important is to relate the convergence properties directly to the operator A because the Lanczos matrix $\underline{B}$ is known only after the Lanczos recursions are performed. The speed of convergence of the Lanczos method has recently been investigated[10] by using Eqs. (2.11) and (2.12). Generalizing this approach from matrices to operators we introduce the quantity q defined by

$$q=\inf_{\lambda}\sup_{i}\left|\frac{\lambda_i-\lambda}{\lambda}\right|, \tag{A12}$$

where the λ_i denote the eigenvalues of the operator A. (More precisely, they denote those eigenvalues, the eigenvectors of which have a nonvanishing overlap with the vector $|\phi\rangle$.) The quotient q has a simple geometrical meaning. Suppose there is a circle in the complex plane which encloses all the eigenvalues λ_i. q denotes the radius of this circle divided by the distance of its center from zero. One has to search for that circle which makes q smallest. If q is smaller than unity, then the error of the Lanczos method vanishes at least like q^N (Ref. 10), i.e.,

$$|\langle\phi|A^{-1}-A_N^{-1}|\phi\rangle|\leq\text{const}\times q^N . \tag{A13}$$

If A is a Hermitian, bounded, and definite operator, then one can easily show[10] that

$$q=\left|\frac{\lambda_{\max}-\lambda_{\min}}{\lambda_{\max}+\lambda_{\min}}\right| \tag{A14}$$

holds, where $\lambda_{\max}$ and $\lambda_{\min}$ denote the smallest and largest eigenvalue of A, respectively. Hence we can conclude that the Lanczos method converges for any Hermitian, bounded, and definite operator. The convergence is the faster, the smaller is the value of q. For more general operators it is difficult to prove the convergence and to estimate the speed of convergence. However, there is another class of operators for which we can prove the convergence. Assume that there is a constant $\lambda\neq 0$ such that $A-\lambda 1$ is a Hilbert-Schmidt operator,[12] i.e., the Hilbert-Schmidt norm

$$\|A-\lambda 1\|_{\text{HS}}^2=\text{tr}[(A-\lambda 1)^{\dagger}(A-\lambda 1)] \tag{A15}$$

is finite. Using the Lanczos basis to evaluate the trace we arrive at

$$\sum_{j=1}^{\infty}|\alpha_j-\lambda|^2+2|\beta_j|^2\leq\|A-\lambda 1\|_{\text{HS}}^2<\infty , \tag{A16}$$

i.e., $\beta_j\to 0$ and $\alpha_j\to\lambda$ as $j\to\infty$, which, together with Eq. (A11), proves the convergence.

In closing this section we want to summarize our results. The Lanczos method converges faster the better the eigenvalues of A are clustered around some common nonzero value. The clustering is very pronounced if $A-\lambda 1$ is a Hilbert-Schmidt operator. In this case almost all eigenvalues of A lie in a circle of radius, say $|\lambda/3|$ centered at λ. The finite number N_0 of eigenvalues not belonging to this circle does not harm the convergence. The convergence will be rapid, at least after we have performed more than N_0 iterations.

The Lanczos algorithm is known to be numerically unstable[5] due to roundoff errors. In all our considerations above we have ignored this fact because we have implicitly assumed exact arithmetic. The Lanczos instability can be cured by reorthogonalizing the Lanczos vectors, which, however, is expensive computationally. Fortunately it turns out that the Lanczos instabilities are irrelevant when computing the expectation value of an operator inverse.[6,10] The instability, however, slows down the convergence because—due to the instability—the Lanczos algorithm computes multiple copies of certain eigenvectors.

[1]J. Schwinger, Phys. Rev. **72**, 742 (1947).
[2]D. K. Watson and V. McKoy, Phys. Rev. A **20**, 1474 (1979).
[3]R. R. Lucchese, K. Takatsuka, and V. McKoy, Phys. Rep. **131**, 147 (1986).
[4]C. Lanczos, J. Res. Natl. Bur. Stand. **45**, 255 (1950).
[5]B. N. Parlett, *The Symmetric Eigenvalue Problem* (Prentice-Hall, Englewood Cliffs, NJ, 1980).
[6]J. K. Cullum and R. A. Willoughby, *Lanczos Algorithm for Large Symmetric Eigenvalue Computations* (Birkhäuser, Boston, 1985), Vol. 1.
[7]C. Duneczky and R. Wyatt, J. Chem. Phys. **87**, 4519 (1987); **89**, 1448 (1988); J. Phys. B **21**, 3727 (1988).
[8]J. Horáček and T. Sasakawa, Phys. Rev. A **28**, 2151 (1983); Phys. Rev. A **30**, 2274 (1984).
[9]J. Horáček and T. Sasakawa, Phys. Rev. C **32**, 70 (1985).
[10]H.-D. Meyer and S. Pal, J. Chem. Phys. **91**, 6195 (1989).
[11]B. Simon, *Quantum Mechanics for Hamiltonians Defined as Quadratic Forms* (Princeton University Press, Princeton, 1971).
[12]M. Reed and B. Simon, *Method of Modern Mathematical Physics, Vol. 1, Functional Analysis* (Academic, New York, 1972).
[13]L. S. Cederbaum and K. Schönhammer, Phys. Rev. A **12**, 2257 (1975).
[14]M. T. Lee, K. Takatasuka, and V. McKoy, J. Phys. B **14**, 4115 (1981).
[15]R. G. Newton, *Scattering Theory of Particles and Waves* (McGraw-Hill, New York, 1966), p. 320.
[16]W. Kohn, Phys. Rev. **74**, 1763 (1948).
[17]R. K. Nesbet, *Variational Methods in Electron-Molecule Scattering Theory* (Plenum, New York, 1980).
[18]W. H. Miller and B. M. D. D. Jansen op de Haar, J. Chem. Phys. **86**, 6213 (1987).
[19]W. Yang and W. H. Miller, J. Chem. Phys. **91**, 3504 (1989).
[20]T. N. Rescigno and A. E. Orel, Phys. Rev. A **23**, 1134 (1981); **24**, 1267 (1981).
[21]B. I. Schneider and L. A. Collins, Phys. Rev. A **24**, 1264 (1981).
[22]H.-D. Meyer, Phys. Rev. A **34**, 1797 (1986).
[23]P.-O. Löwdin, J. Chem. Phys. **19**, 1396 (1951).