Variational procedure and generalized Lanczos recursion for small-amplitude classical oscillations

E. V. Tsiper
Department of Physics, SUNY at Stony Brook, Stony Brook, NY 11794, USA

(Submitted 9 November 1999)
Pis'ma Zh. Éksp. Teor. Fiz. **70**, No. 11, 740–743 (10 December 1999)

A variational procedure is developed which yields the lowest frequencies of small-amplitude oscillations of classical Hamiltonian systems. The genuine Lanczos recursion is generalized to treat related non-Hermitian eigenvalue problems. © *1999 American Institute of Physics.* [S0021-3640(99)00723-9]

PACS numbers: 45.20.Jj, 02.60.−x

The normal modes ξ and frequencies ω of small oscillations of a classical system about equilibrium are determined by the secular equation[1]

$$\omega^2 M\xi = K\xi, \tag{1}$$

where M and K are $N\times N$ symmetric positive-definite matrices of the mass coefficients and spring constants, respectively. In many applications the number N of degrees of freedom is large, while only a few lowest frequencies are of interest.[2] Equation (1) represents a problem more complex than a regular symmetric eigenvalue problem, unless M or K is diagonal.

Equation (1) can be transformed into Hamiltonian form by introducing the canonical momentum $\eta = \omega M\xi$:

$$K\xi = \omega\eta, \quad T\eta = \omega\xi, \tag{2}$$

where $T = M^{-1}$. Thus the frequencies of the normal modes are the eigenvalues of a $2N\times 2N$ matrix

$$\begin{pmatrix} 0 & T \\ K & 0 \end{pmatrix}. \tag{3}$$

The spectrum of this matrix consists of pairs $\pm\omega$, since $(\xi, -\eta)$ is also a solution of (2) that corresponds to $-\omega$. The lowest frequency $\omega_{\min}$ is the lowest *positive* eigenvalue of the matrix (3).

Although the eigenvalues of the matrix (3) are always real, the matrix itself is non-Hermitian, unless $K = T$. Therefore, its diagonalization poses a formidable task. The major problem is that no general minimum principle exists that yields eigenvalues of

arbitrary diagonalizable non-Hermitian matrices. This precludes the formulation of a variational procedure similar to the Rayleigh–Ritz procedure for Hermitian matrices. If $K=T$, the matrix (3) is Hermitian, and its positive eigenvalues coincide with those of K and T.

As is known from quantum mechanics, the lowest eigenvalue $\epsilon_{\min}$ of a Hermitian matrix H can be obtained from the minimum principle

$$\epsilon_{\min}=\min_{\{\psi\}}\frac{(\psi H\psi)}{(\psi\psi)}. \tag{4}$$

The minimum is to be sought over all vectors ψ. The Ritz variational procedure is an approximation when the set $\{\psi\}$ in (4) is restricted to some subspace $\mathcal{K}$ of dimension $n<N$. The best approximation to $\epsilon_{\min}$ in the sense of (4) is obtained as the lowest eigenvalue of the $n\times n$ Rayleigh matrix $\tilde{H}$, obtained by projection of H onto $\mathcal{K}$.

The special paired structure of the matrix (3) makes it possible to generalize (4) so as to yield $\omega_{\min}$. In fact,

$$\omega_{\min}=\min_{\{\xi,\eta\}}\frac{(\xi K\xi)+(\eta T\eta)}{2|(\xi\eta)|}. \tag{5}$$

The minimum is to be sought over all possible phase space configurations $\{\xi,\eta\}$. Before providing the proof to this equation, let me point out some of its features.

First, it states that $\omega_{\min}$ is the minimum arithmetic mean of $(\xi K\xi)$ and $(\eta T\eta)$ over all pairs of vectors ξ, η with scalar product $(\xi\eta)=1$. Since K and T are both positive-definite, the right-hand side is strictly positive and so is $\omega_{\min}$. Second, Eq. (5) is symmetric in K and T, according to the nature of the problem. When $K=T$ the minimum is achieved at $\xi=\eta$, and (5) becomes the same as (4).

Note that the functional in (5) has no maximum, since the denominator can be made arbitrarily small. A global minimum, however, always exists. This is not obvious, since the set of all pairs of vectors with $(\xi\eta)=1$ is not compact. Indeed, say, any vector orthogonal to η can be added to ξ, making $|\xi|$ arbitrarily large. However, the functional in (5) grows indefinitely in this case, so that the global minimum is achieved at finite $|\xi|$ and $|\eta|$.

Variation of (5) with respect to ξ and η yields Eq. (2). Thus the solutions of (2) are the stationary points of (5). The global minimum (5), therefore, indeed gives $\omega_{\min}$. The singularity in the denominator poses no problem, since it corresponds to infinitely large values of the functional, while near the minimum it is analytic.

The minimum principle (5) can, in fact, be obtained from the Thouless minimum principle,[3] derived for non-Hermitian matrices that appear in the random phase approximation (RPA). Equation (5) is transformed into the Thouless minimum principle by the following substitution: $A=(K+T)/2$, $B=(K-T)/2$, $x=(\xi+\eta)/2$, and $y=(\xi-\eta)/2$.

A variational procedure similar to the Rayleigh–Ritz procedure can be formulated if the coordinates ξ and momenta η in (5) are restricted to some subspaces $\mathcal{U}$ and $\mathcal{V}$ of dimension n, respectively.

Let $\{\xi_i\}$ and $\{\eta_i\}$ be two sets of vectors that span $\mathcal{U}$ and $\mathcal{V}$, such that $(\xi_i \eta_j)=\delta_{ij}$. Expanding $\xi=\Sigma u_i \xi_i$, $\eta=\Sigma v_i \eta_i$ and varying (5) with respect to u_i and v_i, we find the latter to obey a $2n\times 2n$ eigenvalue equation

$$\begin{pmatrix} 0 & \tilde{T} \\ \tilde{K} & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}=\tilde{\omega}\begin{pmatrix} u \\ v \end{pmatrix}, \tag{6}$$

with $\tilde{K}_{ij}=(\xi_i K \xi_j)$ and $\tilde{T}_{ij}=(\eta_i T \eta_j)$. Equation (6) generalizes the Hermitian Rayleigh–Ritz eigenvalue equation for $\tilde{H}$. It has $2n$ solutions $\pm\tilde{\omega}$, the lowest positive one of which gives the best approximation to $\omega_{\min}$ in the sense of Eq. (5).

The Krylov subspace[2] for the matrix (3) can be constructed by operating with it many times on an arbitrary vector (ξ_1, η_1):

$$\begin{pmatrix} \xi_1 \\ \eta_1 \end{pmatrix}, \quad \begin{pmatrix} T\eta_1 \\ K\xi_1 \end{pmatrix}, \quad \begin{pmatrix} TK\xi_1 \\ KT\eta_1 \end{pmatrix}, \quad \ldots . \tag{7}$$

The subspace that spans the first n vectors of this sequence has the property of approximating an invariant subspace of (3). Thus it is natural to expand the approximation to an eigenvector of (3) as a linear combination of these vectors. In other words, the natural choice for the subspaces $\mathcal{U}$ and $\mathcal{V}$ for the variational procedure described above are the subspaces $\mathcal{U}_n$ and $\mathcal{V}_n$ that span the upper and lower components of first n vectors of (7).

In order to implement the variational procedure, it is necessary to construct a biorthogonal basis $\{\xi_i, \eta_i\}$, $i=1,\ldots, n$ in $\mathcal{U}_n$ and $\mathcal{V}_n$ and compute matrix elements of $\tilde{K}$ and $\tilde{T}$. Both tasks can be performed simultaneously using the following recursion:

$$\xi_{i+1}=\beta_{i+1}^{-1}(T\eta_i-\alpha_i\xi_i-\beta_i\xi_{i-1}) \tag{8a}$$

$$\eta_{i+1}=\delta_{i+1}^{-1}(K\xi_i-\gamma_i\eta_i-\delta_i\eta_{i-1}). \tag{8b}$$

The four coefficients α_i, β_i, γ_i, and δ_i are to be chosen at each step i so as to make ξ_{i+1} orthogonal to η_i and η_{i-1} and to make η_{i+1} orthogonal to ξ_i and ξ_{i-1}. This appears to be enough to ensure global biorthogonality $(\xi_i \eta_j)=\delta_{ij}$.

Indeed, assume biorthogonality to hold up to step i. Multiplying (8a) by η_j, $j<i-1$, we have $(\eta_j\xi_{i+1})\propto(\eta_j T\eta_i)=(\eta_i T\eta_j)=0$ due to the Hermiticity of T and the fact that $T\eta_j$ is a linear combination of all ξ_k with $k\leqslant j+1<i$. Thus, the biorthogonality also holds for the step $i+1$.

Multiplying (8a) by η_{i-1}, η_i, and η_{i+1} and using biorthogonality, we get $(\xi_i\eta_i)=1$, $\tilde{K}_{ii}=\alpha_i$, and $\tilde{K}_{i,i-1}=\tilde{K}_{i-1,i}=\beta_i$. Similarly, $\tilde{T}_{ii}=\gamma_i$ and $\tilde{T}_{i,i-1}=\tilde{T}_{i-1,i}=\delta_i$. All other matrix elements of $\tilde{K}$ and $\tilde{T}$ vanish.

The recursion (8) is a straightforward generalization of the Hermitian Lanczos recursion[2,4]

$$\psi_{i+1}=\beta_{i+1}^{-1}(H\psi_i-\alpha_i\psi_i-\beta_i\psi_{i-1}) \tag{9}$$

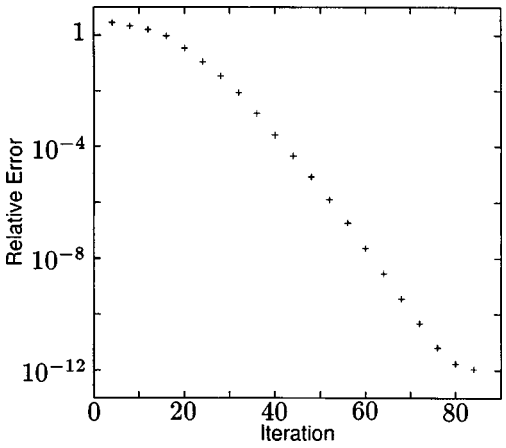

FIG. 1. Convergence of the generalized Lanczos algorithm for a random matrix of the form (3) and size $2N = 200000$.

applicable to any Hermitian matrix H. When $K = T$ and $\xi_1 = \eta_1$, both Equations (8) coincide with each other and with Eq. (9), up to notation.

As in the case of the Hermitian Lanczos algorithm, the several lowest frequencies can be found by projecting the ξ and η components of converged eigenvectors out of the $\mathcal{V}_n$ and $\mathcal{U}_n$ subspaces, respectively.

The method was tested on a set of large sparse random matrices of the form (3). Symmetric matrices T and K were generated having an average of 40 randomly distributed and randomly positioned matrix elements in each row. Both K and T were shifted by an appropriate constant to ensure positive-definiteness. Figure 1 demonstrates the convergence results for a matrix of size $2N = 200000$.

For smaller matrices up to $2N = 2000$, where it was possible to obtain all eigenvalues with regular methods, the present method has converged to the true lowest frequency in all instances.

In conclusion, we have proposed a method that generalizes the Rayleigh–Ritz variational procedure and Lanczos recursion to the case of non-Hermitian matrices of the form (3) which determine the normal modes and frequencies of small-amplitude oscillations of Hamiltonian systems.

Equations (2) have numerous applications beyond purely mechanical problems. The Schrödinger equation in a nonorthogonal basis represents a generalized symmetric eigenvalue problem similar to (1). The RPA and other time-dependent techniques in nuclear physics and quantum chemistry lead to equations similar to (2).[3,5] Last, eigenvectors of so-called Hamiltonian matrices, of which (3) is a special case, solve the nonlinear algebraic Riccati equation which appears in the theory of stability and optimal control.[6]

I would like to acknowledge numerous enlightening discussions with Vladimir Chernyak during my appointment at the University of Rochester.

[1] L. D. Landau and E. M. Lifshitz, *Mechanics*, Butterworth–Heinemann, Oxford (1996).
[2] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice Hall, London (1980).
[3] D. J. Thouless, Nucl. Phys. **22**, 78 (1961).
[4] C. Lanczos, J. Res. Natl. Bur. Stand. **45**, 255 (1950).
[5] J.-P. Blaizot and G. Ripka, *Quantum Theory of Finite Systems*, MIT Press, Cambridge, Mass. (1986).
[6] S. Bittanti, A. J. Laub, and J. C. Willems (eds.), *The Riccati Equation*, Springer-Verlag, Berlin (1991).

Published in English in the original Russian journal. Edited by Steve Torstveit.