THE CONVERGENCE OF A NEW METHOD FOR CALCULATING LOWER BOUNDS TO EIGENVALUES

F. GOERISCH and J. ALBRECHT
Institut für Mathematik, Technische Universität Clausthal
Erzstraße 1, D 3392 Clausthal-Zellerfeld, West Germany

The relationship between an inclusion theorem due to N. J. Lehmann [3] and one recently proposed [2] is investigated here. The theorem due to N. J. Lehmann yields better bounds to the eigenvalues, whereas the new theorem is in general considerably easier to apply. It is shown that sequences of bounds to eigenvalues can be obtained with the use of the new theorem, and that these sequences converge to the bounds provided by Lehmann's theorem. This fact is illustrated by means of numerical results for the following eigenvalue problem:

$$\Delta^2\phi = -\lambda\Delta\phi \text{ in } \Omega, \quad \phi = \frac{\partial\phi}{\partial n} = 0 \text{ on } \partial\Omega, \quad \Omega := \{(x,y)\in\mathbb{R}^2 : |x|<\tfrac{\pi}{2},\ |y|<\tfrac{\pi}{2}\},$$

which occurs in the calculation of buckling stresses of clamped plates under compression.

§1 The two inclusion theorems are first stated, in a version which deviates somewhat from that presented in the original papers [2], [3], but which is especially well suited for practical applications. The following assumptions and definitions are required for this purpose:

Assumptions

A1 D is a real vector space. M and N are symmetric bilinear forms on D; $M(f,f) > 0$ for all $f \in D$, $f \neq 0$.

A2 There exist sequences $(\lambda_i)_{i\in\mathbb{N}}$ and $(\phi_i)_{i\in\mathbb{N}}$ such that $\lambda_i \in \mathbb{R}$, $\phi_i \in D$, $M(\phi_i,\phi_k) = \delta_{ik}$ for $i,k \in \mathbb{N}$,
$M(f,\phi_i) = \lambda_i N(f,\phi_i)$ for all $f \in D$, $i \in \mathbb{N}$,

$$N(f,f) = \sum_{i=1}^{\infty} \lambda_i (N(f,\phi_i))^2 \text{ for all } f \in D.$$

A3 X is a real vector space; $T: D \to X$ is a linear operator; b is a symmetric bilinear form on X. $b(f,f) \geq 0$ for all $f \in X$ and $b(Tf,Tg) = M(f,g)$ for all $f,g \in D$.

A4 $\rho \in \mathbb{R}$, $\rho > 0$; $n \in \mathbb{N}$, $v_i \in D$ for $i=1,\ldots,n$.

Definitions

D1 Matrices A_0 and A_1 are defined by

$A_o := (M(v_i, v_k))_{i,k=1,\dots,n}$, $A_1 := (N(v_i, v_k))_{i,k=1,\dots,n}$.

D2 If A is a symmetric matrix of order n, with the property that $A_o - 2\rho A_1 + \rho^2 A$ is positive definite, $\mu_i(A)$ denotes the i-th smallest eigenvalue of the eigenvalue problem $(A_o - \rho A_1)z = \mu(A_o - 2\rho A_1 + \rho^2 A)z$.

The two inclusion theorems, whose relationship is to be investigated, yield inclusion intervals for the eigenvalues of the eigenvalue problem

$$M(f,\phi) = \lambda N(f,\phi) \text{ for all } f \in D. \qquad (1)$$

The theorems are as follows:

Theorem 1 (N. J. Lehmann [3])

Let $u_i \in D$ be such that $M(f,u_i) = N(f,v_i)$ for all $f \in D$, $i=1,\dots,n$; let the matrix A_2 be defined by $A_2 := (M(u_i,u_k))_{i,k=1,\dots,n}$, and let $A_o - 2\rho A_1 + \rho^2 A_2$ be positive definite. Moreover, suppose that $q \in \mathbb{N}$, $q \leq n$, $\mu_q(A_2) < 0$.

The interval $[\rho - \rho(1-\mu_q(A_2))^{-1}, \rho)$ then contains at least q eigenvalues[1)] of the eigenvalue problem (1).

Theorem 2 ([2])

Let $w_i \in X$ be such that $b(Tf,w_i) = N(f,v_i)$ for all $f \in D$, $i=1,\dots,n$; let the matrix $\tilde{A}_2$ be defined by $\tilde{A}_2 := (b(w_i,w_k))_{i,k=1,\dots,n}$, and let $A_o - 2\rho A_1 + \rho^2 \tilde{A}_2$ be positive definite. Moreover, suppose that $q \in \mathbb{N}$, $q \leq n$, $\mu_q(\tilde{A}_2) < 0$.

The interval $[\rho - \rho(1-\mu_q(\tilde{A}_2))^{-1}, \rho)$ then contains at least q eigenvalues[1)] of the eigenvalue problem (1).

If the assumptions of theorem 1 are satisfied, and if w_i is defined by $w_i := Tu_i$ for $i=1,\dots,n$, the assumptions of theorem 2 are also fulfilled because of $\tilde{A}_2 = A_2$. Thus, theorem 1 is an immediate consequence of theorem 2.

The importance of these theorems is due to the fact that they provide a means of calculating accurate lower bounds to the eigenvalues of problem (1). If the eigenvalues of (1) are arranged in a non-decreasing order,

$$\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots ,$$

if ρ is a lower bound to the eigenvalue λ_{p+q} $(p,q \in \mathbb{N})$ and if, for example, the assumptions of theorem 2 are satisfied, then $\rho - \rho(1-\mu_q(\tilde{A}_2))^{-1}$ is a lower bound to λ_p . For an appropriate choice of the quantities involved, this bound to λ_p is very accurate, even if ρ is only a comparatively rough lower bound to λ_{p+q} .

---

1) Eigenvalues are always counted according to their multiplicity.

It is often difficult, or even impossible, to explicitly give the elements u_i required in theorem 1; by means of theorem 2, in contrast, inclusion intervals for the eigenvalues can be determined with comparative ease - provided that X, b and T have been appropriately chosen (compare §3). However, the results thus obtained cannot be better than those which would be provided by theorem 1, as is now shown:

Lemma 1

Let the assumptions of theorem 1 and 2 be fulfilled.
Then $\rho-\rho(1-\mu_q(\tilde{A}_2))^{-1} \leq \rho-\rho(1-\mu_q(A_2))^{-1}$.

Proof: Since $b(Tu_i,w_k) = N(u_i,v_k) = M(u_i,u_k)$ for $i,k=1,\ldots,n$, it follows that $b(w_i-Tu_i,w_k-Tu_k) = b(w_i,w_k) - M(u_i,u_k)$; hence, the matrix $\tilde{A}_2-A_2$ is positive semidefinite. With the use of the comparison theorem, one obtains $\mu_q(A_2) \leq \mu_q(\tilde{A}_2)$. The assertion can now be immediately deduced.

§2 On the basis of theorem 2, a sequence of inclusion intervals $[\tau_m,\rho)$ will now be constructed in such a manner that $(\tau_m)_{m\in\mathbb{N}}$ converges, and the interval $[\lim_{m\to\infty} \tau_m,\rho)$ coincides with the corresponding inclusion interval from theorem 1. For this purpose, the following additional assumptions and definitions are required.

Assumptions

A5 $\hat{w}_i \in X$ for $i=1,\ldots,n$ and $w_i^* \in X$ for $i\in\mathbb{N}$,
$b(Tf,\hat{w}_i) = N(f,v_i)$ for all $f\in D$, $i=1,\ldots,n$,
$b(Tf,w_i^*) = 0$ for all $f\in D$, $i\in\mathbb{N}$.
The matrix $(b(w_i^*,w_k^*))_{i,k=1,\ldots,m}$ is regular for all $m\in\mathbb{N}$.

A6 $X_o := \{g\in X\colon b(Tf,g) = 0 \text{ for all } f\in D\}$; for all $g\in X_o$ and all $\varepsilon\in\mathbb{R}$ with $\varepsilon > 0$ there exist numbers $m\in\mathbb{N}$, $c_1,\ldots,c_m\in\mathbb{R}$ such that
$b(g-\sum_{i=1}^{m} c_i w_i^*, g-\sum_{i=1}^{m} c_i w_i^*) \leq \varepsilon$.

Remark: If $b(f,f) > 0$ holds for all $f\in X$ with $f \neq 0$, that is, if $(X,b(.,.))$ is a pre-Hilbert space, the assumption A6 states precisely that the subspace spanned by $\{w_i^*\colon i\in\mathbb{N}\}$ is dense in X_o.

Definitions

D3 $\hat{A}_2 := (b(\hat{w}_i,\hat{w}_k))_{i,k=1,\ldots,n}$;
$F_m := (-b(\hat{w}_i,w_k^*))_{i=1,\ldots,n;k=1,\ldots,m}$, $G_m := (b(w_i^*,w_k^*))_{i,k=1,\ldots,m}$,
$A_{2,m} := \hat{A}_2-F_m G_m^{-1} F_m'$ for all $m\in\mathbb{N}$.

The inclusion intervals $[\tau_m,\rho)$ can now be given:

Theorem 3

Let $m,q \in \mathbb{N}$ with $q \leq n$; let the matrix $A_o - 2\rho A_1 + \rho^2 A_{2,m}$ be positive definite, and let $\mu_q(A_{2,m}) < 0$.
If τ_m is defined by $\tau_m := \rho - \rho(1-\mu_q(A_{2,m}))^{-1}$, the interval $[\tau_m, \rho)$ contains at least q eigenvalues of the eigenvalue problem (1).
Proof: Let $F_m G_m^{-1} = (d_{ik})_{i=1,\ldots,n;k=1,\ldots,m}$. The assertion follows immediately from theorem 2, if the w_i occurring there are defined by

$$w_i := \hat{w}_i + \sum_{k=1}^{m} d_{ik} w_k^* \text{ for } i=1,\ldots,n.$$

The following result concerning the convergence of the sequence $(\tau_m)_{m \in \mathbb{N}}$ is now obtained:

Theorem 4

Let the assumptions of theorem 1 be satisfied. If τ_m is defined by $\tau_m := \rho - \rho(1-\mu_q(A_{2,m}))^{-1}$ for $m \in \mathbb{N}$, then $\lim_{m\to\infty} \tau_m = \rho - \rho(1-\mu_q(A_2))^{-1}$.

Proof: Let $F_m G_m^{-1} = (d_{ik}^{(m)})_{i=1,\ldots,n;k=1,\ldots,m}$ for $m \in \mathbb{N}$. Then

$$b(Tu_i - \hat{w}_i - \sum_{k=1}^{m} d_{ik}^{(m)} w_k^*, w_j^*) = 0 \tag{2}$$

for $i=1,\ldots,n, j=1,\ldots,m$, $m \in \mathbb{N}$. Let $\varepsilon \in \mathbb{R}$ with $\varepsilon > 0$. Since $Tu_i - \hat{w}_i \in X_o$ for $i=1,\ldots,n$, there exist numbers $l \in \mathbb{N}$, and $c_{ik} \in \mathbb{R}$ for $i=1,\ldots,n$, $k=1,\ldots,l$ such that

$$b(Tu_i - \hat{w}_i - \sum_{k=1}^{l} c_{ik} w_k^*,\ Tu_i - \hat{w}_i - \sum_{k=1}^{l} c_{ik} w_k^*) \leq \varepsilon$$

for $i=1,\ldots,n$. With the use of (2), it can be shown that

$$b(Tu_i - \hat{w}_i - \sum_{k=1}^{m} d_{ik}^{(m)} w_k^*,\ Tu_i - \hat{w}_i - \sum_{k=1}^{m} d_{ik}^{(m)} w_k^*) \leq \varepsilon$$

for $i=1,\ldots,n$ and all $m \in \mathbb{N}$ with $m \geq l$. By means of the Cauchy-Schwarz inequality, it follows that

$$|b(Tu_i - \hat{w}_i - \sum_{k=1}^{m} d_{ik}^{(m)} w_k^*,\ Tu_j - \hat{w}_j - \sum_{k=1}^{m} d_{jk}^{(m)} w_k^*)| \leq \varepsilon$$

for $i,j=1,\ldots,n$ and all $m \in \mathbb{N}$ with $m \geq l$. Hence,

$$\lim_{m\to\infty} b(Tu_i - \hat{w}_i - \sum_{k=1}^{m} d_{ik}^{(m)} w_k^*,\ Tu_j - \hat{w}_j - \sum_{k=1}^{m} d_{jk}^{(m)} w_k^*) = 0$$

for $i,j=1,\ldots,n$. From the equation

$$(b(Tu_i - \hat{w}_i - \sum_{k=1}^{m} d_{ik}^{(m)} w_k^*,\ Tu_j - \hat{w}_j - \sum_{k=1}^{m} d_{jk}^{(m)} w_k^*))_{i,j=1,\ldots,n} = A_{2,m} - A_2$$

it follows that $A_{2,m} - A_2$ is positive semidefinite for $m \in \mathbb{N}$, and that $\lim_{m\to\infty} A_{2,m} = A_2$. This gives $\lim_{m\to\infty} \mu_q(A_{2,m}) = \mu_q(A_2)$, from which the assertion follows immediately.

Remark: The sequence $(\tau_m)_{m\in\mathbb{N}}$ is non-decreasing.

§3 The practical application of the results presented is now illustrated with the use of an example. - The quantities D, M, N, X, T, b occurring in the assumptions A1 and A3 are defined in the following manner ($\overset{\circ}{W}{}_2^{(2)}(\Omega)$ denotes the Sobolev space defined in [4]):

$$D := \{f \in \overset{\circ}{W}{}_2^{(2)}(\Omega): f(x,y) = f(y,x) = f(-x,y) \text{ for } (x,y) \in \Omega\},$$

$$M(f,g) := \int_\Omega \Delta f \Delta g dxdy, \quad N(f,g) := \int_\Omega \left(\frac{\partial f}{\partial x}\frac{\partial g}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial g}{\partial y}\right)dxdy \text{ for } f,g \in D,$$

$$X := \{f \in L_2(\Omega): f(x,y) = f(y,x) = f(-x,y) \text{ for } (x,y) \in \Omega\},$$

$$Tf := -\Delta f \text{ for } f \in D, \quad b(f,g) := \int_\Omega fg dxdy \text{ for } f,g \in X,$$

where $\Omega := \{(x,y) \in \mathbb{R}^2: |x| < \frac{\pi}{2}, |y| < \frac{\pi}{2}\}$.
In this case, the eigenvalue problem (1) is the weak form of the eigenvalue problem

$$\left.\begin{aligned} &\Delta^2\phi = -\lambda\Delta\phi \text{ in } \Omega, \ \phi = \frac{\partial\phi}{\partial n} = 0 \text{ on } \partial\Omega, \\ &\phi(x,y) = \phi(y,x) = \phi(-x,y) \text{ for } (x,y) \in \Omega. \end{aligned}\right\} \quad (3)$$

For specifying the quantities occurring in A4 and A5, a number $r \in \mathbb{N}$ is first chosen; ρ, n, v_i, $\hat{w}_i$, w_i^* are then defined as follows:

$$\rho := 10, \quad n := \tfrac{1}{2}r(r+1),$$

$$v_i(x,y) := \cos^{s+1}(x)\cos^{t+1}(y) + \cos^{s+1}(y)\cos^{t+1}(x) \text{ for } i=1,\dots,n,$$

where $s,t \in \mathbb{N}$ are determined by $r \geq s \geq t$, $i=\frac{1}{2}s(s-1)+t$,

$$\hat{w}_i := v_i \text{ for } i=1,\dots,n,$$

$$w_i^*(x,y) := \cosh(ix)\cos(iy) + \cosh(iy)\cos(ix) \text{ for } i \in \mathbb{N}.$$

The assumptions A1, A3, A4 and A5 are obviously fulfilled, and the proofs of A2 and A6 proceed in analogy with the corresponding proofs in [4], p. 472 and [1], respectively.
By means of the Cauchy-Schwarz inequality, it follows from the comparison theorem that the eigenvalues of the problem

$$-\Delta\phi = \lambda\phi \text{ in } \Omega, \ \phi = 0 \text{ on } \partial\Omega,$$

$$\phi(x,y) = \phi(y,x) = \phi(-x,y) \text{ for } (x,y) \in \Omega,$$

are lower bounds to the corresponding eigenvalues of (1); hence $\rho=10$ is a lower bound for the second eigenvalue of (1).
The first six terms of the sequence $(\tau_m)_{m\in\mathbb{N}}$ calculated for q=1 with the use of theorem 3 are compiled in the first six rows of table 1 for various values of n. Since at least one eigenvalue of (1) is contained in each of the intervals $[\tau_m,\rho)$, the numbers τ_m are lower bounds to the lowest eigenvalue of (1). An upper bound Λ to this eigenvalue, which has been determined with the use of the functions $v_1,\dots,v_n$ by

means of the Rayleigh-Ritz method, is given in the last row of table 1, for the respective value of n.

By virtue of theorem 4, the sequences $(\tau_m)_{m\in\mathbb{N}}$ converge to the corresponding bounds which would result from theorem 1. In order to apply theorem 1, however, it would be necessary to determine the exact solution u_i of the boundary value problem

$$\Delta^2 u_i = -\Delta v_i \text{ in } \Omega,\ u_i = \frac{\partial u_i}{\partial n} = 0 \text{ on } \partial\Omega,$$

which is not an easy task.

	n = 1	n = 10	n = 21
τ_1	5.049	5.057	5.057
τ_2	5.250 0	5.265 8	5.265 9
τ_3	5.283 52	5.302 42	5.302 46
τ_4	5.284 556	5.303 564	5.303 602
τ_5	5.284 582 0	5.303 587 3	5.303 625 3
τ_6	5.284 582 21	5.303 587 37	5.303 625 40
Λ	5.333 333 34	5.303 662 26	5.303 626 22

Table 1 Bounds to the lowest eigenvalue of (3)

The method based on theorem 3 has also been applied with great success to many other eigenvalue problems involving partial differential equations.

The authors gratefully acknowledge the support of this work by the Deutsche Forschungsgemeinschaft.

References

[1] Colautti, M.P.: Su un teorema di completezza connesso al metodo di Weinstein per il calcolo degli autovalori. Atti Accad. Sci. Torino, Cl. Sci. Fis. Mat. Nat. 97 (1962/63), 171 -191

[2] Goerisch, F. and H. Haunhorst: Eigenwertschranken für Eigenwertaufgaben mit partiellen Differentialgleichungen. Z. Angew. Math. Mech. 65 (1985), 129 - 135

[3] Lehmann, N.J.: Optimale Eigenwerteinschließungen. Numer. Math. 5 (1963), 246 - 272

[4] Rektorys, K.: Variational methods in mathematics, science and engineering. Dordrecht-Boston: D. Reidel Publishing Company 1977