PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

On the Eigenvalue Problem Tu-$\lambda $Su = 0 with Unbounded and Nonsymmetric Operators T and S

W. V. Petryshyn

Phil. Trans. R. Soc. Lond. A 1968 **262**, 413-458
doi: 10.1098/rsta.1968.0001

ON THE EIGENVALUE PROBLEM $Tu-\lambda Su=0$ WITH UNBOUNDED AND NONSYMMETRIC OPERATORS T AND S†

By W. V. PETRYSHYN

Department of Mathematics and Institute for Computer Research, University of Chicago

(*Communicated by G. Temple, F.R.S.—Received* 14 *December* 1966—*Revised* 3 *July* 1967)

CONTENTS

In this paper we study both the theoretical problem of the existence and the practical problem of the approximate calculation of eigenvalues and eigenvectors of (i) $Tu-\lambda Su = 0$, where T and S are some linear (in general unbounded and nonhermitian) operators in a Hilbert space. After a short discussion of a class of K-symmetric operators, in section 2 the author proves the existence of eigenvalues and eigenvectors of (i) under various conditions on T and S and investigates conditions under which the set of eigenvectors of (i) is complete. Section 3 indicates briefly the applicability and the unifying property of the generalized method of moments to the approximate solution of (i). Section 4 presents and thoroughly studies a very general new iterative method for the approximate solution of (i). The advantage of this method is that it does not require the practically inconvenient preliminary reduction of (i) to an equivalent problem with bounded operators and that under certain rather general conditions the convergence is monotonic. Furthermore, by specializing operators and parameters, our iterative method contains as a special case, almost every known iterative method for the calculation of eigenvalues (mostly proved previously only for symmetric matrices and bounded operators). Finally the applicability and the numerical effectiveness of the iterative method is illustrated by calculating the smallest eigenvalue for a selfadjoint and non-selfadjoint eigenvalue problems arising in the problems of elastic stability.

Introduction

The problem of proving the existence of eigenvalues and eigenvectors for (i) $Tu-\lambda Su = 0$, where T and S are certain linear operators in a Hilbert space, and the problem of solving the eigenvalue problem (i) approximately have been studied by a number of authors.‡

† The results presented in this paper were derived by the author in 1962 during his temporary membership at the Courant Institute of Mathematical Sciences, New York University, and were issued under the above title in January 1963, as an AEC Research and Development Report NYO-10,425. These results are published in view of the interest in them expressed by a number of pure and applied mathematicians in several countries. The paper is presented as it appeared in its original form except for a shortened version of § 1 and the illustrative examples at the end of § 4.

‡ The articles cited in this paper are only those to which a direct reference is made. For a more complete list of references concerned with problem (i) see [4, 6, 10, 16, 25, 37].

[Because of the large number of citations made in this paper the system of referencing that is customary in these pages has not been followed.]

However, the investigation of both of these problems for (i) was limited mostly to selfadjoint bounded and unbounded differential and abstract operators. Thus, when T and S are selfadjoint operators, the existence problem was considered in [4, 5, 8, 10, 14, 25] in case the operators are positive definite and in [12, 18, 36] in case S is the identity and the spectrum of T contains at most eigenvalues of finite multiplicity. The investigation of the approximation problem, both for bounded and unbounded operators, led to the development of a group of *direct methods* of Ritz, Galerkin, moments, and others [4, 5, 10, 14, 15, 16, 25, 27] and to a group of *iterative methods* of gradient type [1, 2, 3, 6, 11, 13, 16, 17, 19, 20, 22, 29, 30, 33, 37]. Unfortunately, as was pointed out in [16], the study of the gradient methods was almost [33] entirely restricted to selfadjoint and positive definite finite matrices and bounded operators.

The purpose of this paper is to study both of the above mentioned problems (i.e. existence and approximation) connected with the eigenvalue problem (i), where T and S are linear, unbounded K-symmetric operators studied in [23, 26, 39] which, as is known [23, 24, 40, 41] include certain non-selfadjoint differential operators of even and odd order. Our investigation allows us not only to extend to this larger class of eigenvalue problems (i) the results of the above authors and to present these seemingly different methods in a more unified manner but also affords the introduction of a general iterative method for the direct solution of the eigenvalue problem (i) without the practically inconvenient preliminary reduction of (i) to an equivalent eigenvalue problem with bounded operators.

After stating in § 1 certain properties of K-symmetric and H_0-bounded operators, in § 2 we consider the problem of existence of eigenvalues and eigenvectors of (i) and their properties under various conditions on T and S. Accordingly, in §§ 2·1 and 2·2 we extend to problem (i) the corresponding results in [8, 25] while in § 2·3 we consider conditions under which the results derived in [14, 36, 18, 12] are also valid for (i). Using the properties of H_0-bounded operators we prove stronger assertions concerning the Fourier expansion and the completeness of eigenvectors than those obtained in [18, 12].

In § 3 we indicate briefly the applicability and the unifying property of the generalized method of moments for the approximate solution of the eigenvalue problem (i); the section contains slight generalization of the corresponding results of Mikhlin [25] and Polsky [27]. In fact, in § 3·1 we prove two lemmas which generalize the results of Polsky [27] while in § 3·2, following Mikhlin, we formulate the generalized method of moments and prove its convergence. In § 3·3, we obtain as special cases *the ordinary Ritz method*, *the generalized Ritz method*, *the Galerkin method*, and *the method of moments*.

In § 4 we investigate a general iterative method for the solution of (i) which at the same time unifies and extends to the eigenvalue problem (i) with K-symmetric operators the results in [1, 3, 6, 11, 16, 17, 20, 29, 33] obtained there mostly for positive definite symmetric matrices and bounded operators. Thus, in §§ 4·1 and 4·2 we formulate the method while in § 4·3 we derive the basic convergence theorems. In § 4·4 we discuss conditions under which we obtain the convergence to the smallest eigenvalue and an error estimate for the approximate eigenvectors. Section 4·5 deals with the problem of constructing various methods and their special cases. Here we discuss *the method of constant factor* [6], a relatively new method called here *the method with relative minimal norms* [20, 29, 30] and its special cases [2], *the accelerated method with relative minimal norms* and its special cases [9, 1], *the generalized method* *of steepest descent* [15, 3, 33], the method introduced by Hestenes & Karush [11] which we call *the modified method of steepest descent*, and others. Finally, in § 4·6 we illustrate the applicability and the numerical effectiveness of the iterative method (4·8) to (4·10) by calculating (by means of the method with relative minimal norms) the smallest eigenvalue λ_1 for a selfadjoint and non-selfadjoint eigenvalue problems arising in the problems of elastic stability [35].

1. On a class of K-p.d. and non-K-p.d. operators

In this section we define a class of linear (in general unbounded and non-Hermitian) operators in a complex Hilbert space H and summarize some of their properties obtained in [26, 39] which we will use in the subsequent sections on eigenvalue problems. Especially those properties are mentioned which are useful in the application of the theory and the methods discussed in this paper to integral, integro-differential, and differential eigenvalue problems which need not be Hermitian. For the detailed proofs of the assertions made in this section see [26, 39].

1·1. *Certain properties of K-p.d. and non-K-p.d. operators*

Let H be a complex and separable Hilbert space. An operator T defined on a dense domain D_T in H will be called *K-positive definite* (K-p.d.)† if there exists a closeable‡ operator K with $D_K \supseteq D_T$ mapping D_T onto a dense subset KD_T of H and two constants $\alpha_1 > 0$ and $\alpha_2 > 0$ such that

$$(Tu, Ku) \geqslant \alpha_1 \|u\|^2 \quad (u \in D_T), \tag{1·1}$$

$$\|Ku\|^2 \leqslant \alpha_2 (Tu, Ku) \quad (u \in D_T). \tag{1·2}$$

Let us first note that the class of K-p.d. operators, as defined above, contains among others, for example, the following operators: (*a*) Positive definite operators T; in this case we choose the operator K to be either the identity I or, if T is also selfadjoint, to be any root of T. (*b*) Closeable and densely invertible§ operators T when we take K to be T. (*c*) The operators T of the form $T = -S^{2j+1}$ or $T = S^{2j+2}$ when for some i, $0 \leqslant i < j$, the operator $S^{2(j+i+1)}$ is positive definite; in this case we take $K = S^{2i+1}$ or $K = S^{2i+2}$, provided, of course, that K so defined is closeable and such that KD_T is dense in H. This class, in particular, contains ordinary differential operators of an odd and even order and weakly elliptic partial differential operators of an odd or even order which in general are non-selfadjoint [24, 40, 41]. (*d*) A subclass of bounded symmetrizable operators investigated by many authors [38].

Lemma 1·1. *If T is K-p.d., then*

(*a*) *T is invertible*;

(*b*) *T is K-symmetric*; *i.e.* $(Tu, Kv) = (Ku, Tv) \quad (u, v \in D_T)$;

(*c*) $|(Tu, Kv)|^2 \leqslant (Tu, Ku)\,(Tv, Kv) \quad (u, v \in D_T)$;

(*d*) *T is closeable.*

† The class of K-p.d. operators with K closed was introduced by Martyniuk [23]. A more general theory was developed in [26, 39].

‡ Let us recall that K is said to be *closeable* if whenever u_n is a sequence in D_K and f an element in H such that $u_n \to 0$ and $Ku_n \to f$, as $n \to \infty$, then $f = 0$. Let us add that, as was observed by Rellich [31], in applications it is more convenient to work with closeable than with closed operators.

§ We will call an operator T *invertible* if T has a bounded inverse, *densely invertible* if T is invertible and the range R_T is dense in H, and *continuously invertible* if T is densely invertible and $R_T = H$.

Let us observe that if the space H is real then the entire discussion in this article concerning the eigenvalue problems remains valid provided T is assumed to be also K-symmetric.

Let f be an element in H and $F(u)$ the functional

$$F(u) = (Tu, Ku) - (Ku, f) - (f, Ku) \tag{1·3}$$

defined on D_T. It was shown in [26, 39] that the problem of solving the equation

$$Tu = f \tag{1·3a}$$

is equivalent to the problem of minimizing the functional $F(u)$ and that to solve the latter it is, in general, necessary to extend somewhat the set D_T on which $F(u)$ is defined and with it also the operator T. It is known [7, 25, 26, 39] that the variational problem is solvable and that T possesses a closed and continuously invertible K-p.d. extension. Indeed, this follows from the following arguments.

Let $D[T]$ denote the set D_T with the new metric

$$[u, v] = (Tu, Kv), \quad |u|^2 = [u, u] \quad (u, v \in D_T). \tag{1·4}$$

Clearly, $D[T]$ satisfies all the axioms of a Hilbert space except possible that it is incomplete. Furthermore, in view of (1·1) and (1·2), we have the inequalities

$$|u| \geqslant \gamma_1 \|u\|, \quad \gamma_1 = \alpha_1^{\frac{1}{2}} > 0 \quad (u \in D_T) \tag{1·5}$$

and

$$\|Ku\| \leqslant \gamma_2 |u|, \quad \gamma_2 = \alpha_2^{\frac{1}{2}} > 0 \quad (u \in D_T). \tag{1·6}$$

Let H_0 denote the completion of $D[T]$ in the metric (1·4).

Lemma 1·2. (*a*) *$D[T]$ is dense in H_0.*

(*b*) *H_0 is a subset of H in the sense of uniquely identifying the elements from H_0 with certain elements from H.*

(*c*) *K can be extended to a bounded operator K_0 mapping all of H_0 to H, such that $K \subset K_0 \subset \bar{K}$, where $\bar{K}$ is the closure† of K in H.*

(*d*) *The inequalities* (1·5) *and* (1·6) *are valid for all u in H_0.*

Having constructed the auxiliary space H_0, it is now easy to solve the variational problem. In fact, by lemma 1·2 (*c*), (f, Ku) is a bounded conjugate linear functional of u in H_0 and hence the Fréchet–Riesz theorem implies that to every fixed element f in H there exists a unique element $w \in H_0$ such that for all u in H_0

$$(f, Ku) = [w, u]. \tag{1·7}$$

Consequently, the functional

$$F(u) = [u, u] - [u, w] - [w, u] = |u-w|^2 - |w|^2, \tag{1·8}$$

which by definition (1·3) is valid for all u in D_T, can be also extended to the entire space H_0. Considered in H_0, $F(u)$ attains its minimum d at $u = w$ with

$$d = \inf_{u \in H_0} F(u) = F(w) = -|w|^2. \tag{1·9}$$

We shall formulate the above result in the following lemma.

† The operator $\bar{K}$ is called a *closure* or a *trivial closed extension* of K in D_K if it is defined on the set $D_{\bar{K}} (\supset D_K)$ consisting of all elements u in H for which there exists a sequence $\{u_n\}$ in D_K and an element f in H such that $\|u_n - u\| \to 0$ and $\|Ku_n - f\| \to 0$ as $n \to \infty$; in this case we set $\lim_n Ku_n = \bar{K}u = f$.

Lemma 1·3. *If T is K-p.d., then $d = \inf_{u\in H_0} F(u)$ is attained at $u = w$, where w is uniquely determined by* (1·7). *Furthermore, the value of d is given by* (1·9).

Let us note that, in general, the element w which minimizes the functional $F(u)$ is not necessarily in D_T so that the equation (1·3*a*) may not have a solution unless T could be somewhat extended. Theorem 1·1 below shows that for K-p.d. T such an extension is always possible.

Theorem 1·1. *If T is K-p.d., then T can be extended to a closed K_0-p.d. operator T_0 such that $T_0 \supseteq T$, T_0 is continuously invertible, and D_{T_0} consists of all elements realizing the infimum of $F(u)$ in H_0 as f ranges through the entire space H.*

The operator T_0 will be called a *solvable generalized Friedrichs extension* (s.g.F.e.) of T. For $K = I$, theorem 1·1 furnishes a selfadjoint extension of a symmetric positive definite operator constructed by Friedrichs [7, 25].

Remark 1. T may still have other K-p.d. extensions. But among these extensions there is only one, the operator T_0 we have just constructed, whose domain is contained in H_0 and which in some sense is maximal [26]; i.e. if T' is an arbitrary K-p.d. extension of T such that $D_{T'} \subset T_0$, then $T_0 \supseteq T'$.

Theorem 1·2 below can be used in one of the often applied techniques in the investigation of complicated operator equations which consists in comparing these equations with much simpler operator equations, the properties of which are well known. The 'closeness', defined in some sense, of the two operator equations implies the community of various important properties such as existence and uniqueness of solutions, the applicability of various approximate methods and their convergence, etc.

Theorem 1·2. *If T is K-p.d. and K is closed with $D_K = D_T$, then there exists a constant $\theta_1 > 0$ such that*

$$\|Tu\| \leqslant \theta_1 \|Ku\| \quad (u \in D_T). \tag{1·10}$$

Furthermore, T is closed, $R_T = R_K = H$, and

$$\frac{\gamma_1}{\theta_1 \cdot \gamma_2}\|u\| \leqslant \frac{1}{\theta_1 \cdot \gamma_2}\|u\| \leqslant \|Ku\| \leqslant \gamma_2 |u| \leqslant \gamma_2^2 \|Tu\| \leqslant \theta_1 \cdot \gamma_2^2 \|Ku\| \quad (u \in D_T). \tag{1·11}$$

Corollary 1·1. *If T is K-p.d. and K is closed with $D_K = D_T$, then the operators T and K form an acute angle.*†

Finally in this section we state a theorem concerning s.g. Friedrichs extensions for a much more general class of nonsymmetric and non-K-symmetric operators of which we shall make use in § 4.

Theorem 1·3. *Let T be K-p.d. and L be an operator with $D_L = D_T$. If there exist $\eta_1 > 0$ and $\eta_2 > 0$ such that*

$$|(Lu, Ku)| \geqslant \eta_1 |u|^2 \qquad (u \in D_T), \tag{1·12}$$

$$|(Lu, Kv)| \leqslant \eta_2 |u|\,|v| \qquad (v, u \in D_T), \tag{1·13}$$

† Following Sobolevsky [34] we say that two densely defined operators P and R form an *acute angle* if $D_P = D_R$, $(Pu, Ru) \geqslant \delta \|Pu\|\,\|Ru\|$ for all u in D_P and some $\delta > 0$, and they vanish only at zero element. Later on we will also use this definition in a slightly more extended sense in which the above inequality is replaced by the more general one: $|(Pu, Ru)| \geqslant \delta \|Pu\|\,\|Ru\|$ for all u in D_P.

then L has a s.g.F.e. L_0 such that L_0 is closed, $L_0 \supset L$, L_0 is continuously invertible, and $L_0 = T_0 W_0$, where W_0 is a certain extension of the operator $T_0^{-1} L$ in H_0.

As was shown in [39], by specializing our operators we derive from theorem 1·3 certain results of Lax & Milgram [21] under less restrictive conditions.

1·2. H_0-*bounded operators*

In later discussion we shall also have the opportunity to deal with closeable operators S such that $D_S \supseteq H_0$. We call such operators H_0-*bounded.*† If S is H_0-bounded, then by $\underline{S}$ we denote the operator from H_0 to H defined by $\underline{S}u = Su$. This definition is explained by the lemma.

Lemma 1·4. *If S is H_0-bounded, then $\underline{S}$ is bounded, i.e., $\|Su\| \leqslant \eta_5 |u|$ for some constant $\eta_5 > 0$.*

Lemma 1·5. *If S is H_0-bounded, then ST_0^{-1} is bounded in H.*

An operator S is called H_0-compact if S is H_0-bounded and $\underline{S}$ is compact, i.e. if for every bounded set Q_0 in H_0 the image $\underline{S}Q_0$ is compact in H. Such operators are of importance in proving the existence of eigenvectors and the applicability of various approximate methods. The three theorems below offer us various possibilities and conditions under which the operator $T_0^{-1}S$ is compact in H_0. This, as is known, is useful in the application to differential equations and eigenvalue problems.

Theorem 1·4. *The assertion (a): S is H_0-compact $\Rightarrow$ (b): $T_0^{-1}S$ is compact in H_0 and (c): ST_0^{-1} is compact in H. Furthermore, $(b) \Rightarrow (c)$. If in addition we assume that K is closed and $D_K = D_T$, then $(a) \Leftrightarrow (b) \Leftrightarrow (c)$.*

Theorem 1·5. *If S is H_0-bounded and KT_0^{-1} is compact in H, then $T_0^{-1}S$ is compact in H_0.*

Our last theorem in this section offers us the possibility of verifying the compactness of $L_0^{-1}S$ in H_0, where $D_S \supset D_T$ and L is a complicated operator satisfying the conditions of theorem 1·3 and $L_0 = T_0 W_0$, by verifying instead the compactness of the simpler operator $T_0^{-1}S$.

Theorem 1·6. *$L_0^{-1}S$ is compact in H_0 if and only if $T_0^{-1}S$ is compact in H_0.*

Remark 2. In §§ 2, 3, and 4 whenever we consider the unbounded K-p.d. operators it will be assumed that, if necessary these operators have been already extended.

2. The eigenvalue problem $Tu - \lambda Su = 0$

In this section an attempt is made to present a theoretical basis for the general discussion in the next two sections of the generalized method of moments and of an iterative method for the approximate solution of the eigenvalue problem (2·1) below. Here we derive some of the general properties of the eigenvalues and eigenelements of (2·1) and also consider the problem of their existence under various conditions. Since some theorems in this section represent generalizations of the corresponding theorems for the selfadjoint eigenvalue problems considered in [25] which are proved by similar arguments, the details of some

† If K is closed and $D_K = D_T$, then in view of theorem 1·2, the concepts of H_0-boundedness and H_0-compactness of S reduce to the concepts of T-boundedness and T-compactness introduced by Gokhberg & Krein [9].

parts of their proofs could have been omitted. However, for the convenience of the reader the proofs are given in full.

The results in this section represent an extension to the eigenvalue problem (2·1) of the corresponding results for the selfadjoint differential and abstract eigenvalue problems obtained in [5, 8, 10, 12, 14, 18, 25, 36]. At the same time we also obtain some additional new results.

2·1. *General remarks*

Let H be a complex Hilbert space and let us consider now the linear eigenvalue problem in H

$$Tu - \lambda Su = 0, \tag{2·1}$$

where T is K-p.d., S is an operator with $D_S \supseteq D_T$, and λ is a complex number. A value of λ for which (2·1) has a nontrivial solution u in D_T will be called an *eigenvalue* of (2·1) and u its corresponding *eigenelement*. The set of all eigenvalues of (2·1) will be denoted by $p\sigma$(2·1) and called the *point spectrum* or *discrete spectrum* of (2·1). The number of linearly independent eigenelements u belonging to the same eigenvalue λ will be called the *multiplicity* of λ.

Let us assume for the present that S is K-*real*; i.e. (Su, Ku) is real for all u in D_T. It is easy to see that S is K-symmetric on D_T.

Lemma 2·1. *If T is K-p.d. and S is K-symmetric, then*

(*a*) *All eigenvalues λ of* (2·1) *are real.*

(*b*) *The eigenvectors w and w' belonging to distinct eigenvalues λ and λ' are orthogonal in the sense that $(Tw, Kw') = 0$ and $(Sw, Kw') = 0$.*

(*c*) *The set $p\sigma$(2·1) contains at most countably many eigenvalues.*

Proof. (*a*) Let λ be an eigenvalue of (2·1) and w its corresponding eigenelement. Then $Tw = \lambda Sw$ and hence $(Tw, Kw) = \lambda(Sw, Kw)$. This and (1·1) imply that $(Sw, Kw) \neq 0$ and therefore that

$$\lambda = \frac{(Tw, Kw)}{(Sw, Kw)}. \tag{2·2}$$

The relation (2·2) shows that the eigenvalue λ is real and that it can be expressed in terms of the corresponding eigenelement. Formula (2·2) will be very useful in our later discussion.

(*b*) Let w be an eigenelement of (2·1) with eigenvalue λ and w' an eigenelement with eigenvalue $\lambda' \neq \lambda$. Then, in view of lemma 1·1,

$$0 = (Tw, Kw') - (Kw, Tw') = \lambda(Sw, Kw') - \lambda'(Kw, Sw')$$
$$= (\lambda - \lambda')(Sw, Kw')$$

and

$$(Tw, Kw') = \lambda(Sw, Kw') = 0.$$

(*c*) If to each eigenvalue of (2·1) we order its corresponding eigenelements normalized in the metric (1·4), then the totality of these eigenelements must constitute an orthonormal set in the separable Hilbert space H_0. Consequently, the set $p\sigma$(2·1) can at most be denumerably infinite.

We see from lemma 2·1 that the eigenvalue problem (2·1) has at most countably many eigenvalues. The problem whether there exist any eigenvalues of (2·1) at all is, in general, a difficult one and will be discussed below for some classes of eigenvalue problems (2·1) in which the operators T and S satisfy some additional conditions.

2·2. *K-p.d. eigenvalue problem*

Let us assume that there exist constants $\beta_1 > 0$, $\beta_2 > 0$, and $\beta_3 > 0$ such that

$$(Su, Ku) \geqslant \beta_1 \|u\|^2 \quad (u \in D_T) \tag{2·3}$$

and either
$$\|Ku\|^2 \leqslant \beta_2 (Su, Ku) \quad (u \in D_T) \tag{2·4}$$

or S is closeable and
$$\|Su\|^2 \leqslant \beta_3 (Su, Ku) \quad (u \in D_T). \tag{2·4'}$$

The eigenvalue problem (2·1) in which the operator T is K-p.d. and S has the above properties will be called K-p.d. Let H_1' be the completion of D_T in the metric

$$[u, v]_1 = (Su, Kv), \quad |u|_1^2 = [u, u]_1. \tag{2·5}$$

It is not hard to see that under our conditions the space H_1' is contained in H in the sense of identifying elements from H_1' with certain elements from H. Furthermore, the inequalities (2·3), (2·4) and (2·4') remain valid for all u in H_1' provided, of course, that for u in H_1' we replace (Su, Ku) by $[u, u]_1$ and K and S are also used to denote their extensions to all of H_1' in case of (2·3) to (2·4) or (2·3) to (2·4'), respectively. Since (2·1) is K-p.d., the functional $E(u)$ defined by

$$E(u) = \frac{(Tu, Ku)}{(Su, Ku)} \tag{2·6}$$

is positive on D_T and therefore has the infimum (inf.) which we denote by λ_1. Thus, there exists a *minimizing sequence* $\{u_i\}$ in D_T such that

$$\inf_{u \in D_T} E(u) = \lim_i E(u_i) = \lambda_1. \tag{2·7}$$

However, from the existence of such a minimizing sequence of elements we cannot conclude without imposing further conditions that there exists an element w_1 in D_T or even in H_0 for which $E(w_1)$ is actually equal to λ_1. Before we discuss this question let us first observe that the comparison of formulas (2·2) and (2·6) shows that the eigenvalues λ of (2·1) can be looked for among the values of the functional $E(u)$. In fact, we have the following useful lemma.

LEMMA 2·2. *The inf. λ_1 of $E(u)$ is an eigenvalue of* (2·1) *if and only if there exists an element $w_1 \neq 0$ in D_T such that*

$$\frac{(Tw_1, Kw_1)}{(Sw_1, Kw_1)} = \lambda_1.$$

Moreover, if such a w_1 exists then λ_1 is the smallest eigenvalue of (2·1).

The proof of this lemma is omitted for, as will be seen below, it will be a special case of lemma 2·4. Let us note that lemma 2·2 allows us to replace the problem of finding the least eigenvalue of (2·1) by the problem of finding an element in H realizing the inf. of $E(u)$ provided, of course, that such an element exists. To prove its existence we assume the validity of the following condition:

(α): *Every set of elements in $H_{01} \equiv H_0 \cap H_1'$ bounded in the H_0-norm is compact in the H_1'-norm.*

LEMMA 2·3. *If condition* (α) *is satisfied, then H_0 is a subset of H_1' in the sense of identifying uniquely the elements from H_0 with certain elements from H_1'. In this sense $D_T \subseteq H_0 \subseteq H_1'$.*

Proof. Let h_0 be an arbitrary element in H_0. Since D_T is dense in H_0 there exists a sequence $\{u_n\}$ in D_T such that $|u_n - h_0| \to 0$ as $n \to \infty$. Thus, $\{u_n\}$ is a Cauchy sequence in H_0 and, in particular, it is bounded in H_0-norm. Hence, by condition (α), $\{u_n\}$ contains a subsequence $\{u_{n_i}\}$ which is convergent in the H_1'-metric. Since H_1' is complete there exists an element h_1 in H_1' such that $|u_{n_i}-h_1|_1 \to 0$ as $i \to \infty$. Let us note that in view of lemma 1·2 and the properties of H_1' it follows easily that every subsequence of $\{u_n\}$, convergent in the H_1'-metric, converges to the same element h_1 in H_1' and that any two equivalent sequences in H_0-norm produce equivalent subsequences convergent in the H_1'-metric. We can thus assign to each element h_0 of H_0 a unique element h_1 of H_1', the identification obviously being linear and such that $h_0 = h_1$ for elements h_0 in D_T. Moreover, this identification correspondence is also one-to-one for if $h_1 = 0$, then $h_0 = 0$. To see this note first that, by (2·3), $\|u_{n_i}\| \leqslant \beta^{-\frac{1}{2}} |u_{n_i}|_1 \to 0$ as $i \to \infty$ and, by (1·6), $\|Ku_{n_i} - Ku_{n_j}\| \leqslant \gamma_2 |u_{n_i} - u_{n_j}| \to 0$ as i and j increase indefinitely. This and the closeability of K imply that for any u in D_T

$$[h_0, u] = \lim_n [u_n, u] = \lim [u_{n_i}, u] = \lim_i (Tu_{n_i}, Ku) = \lim_i (Ku_{n_i}, Tu) = 0.$$

Consequently, we must conclude that $h_0 = 0$, as was to be shown.

Let us note in passing that lemma 2·3 is also valid under the weaker condition

$$(Tu, Ku) \geqslant \tilde{\alpha}(Su, Ku) \quad (u \in D_T)$$

for some constant $\tilde{\alpha} > 0$. However, the stronger condition (α) is imposed for it is essential in the existence proof.

Let us write $E(u)$ in the form

$$E(u) = \frac{[u, u]}{[u, u]_1} \tag{2·6'}$$

and observe that, in view of lemma 2·3, the right side in (2·6′) which is defined for all u in D_T has also meaning for all u in H_0. We may, therefore, use it to extend the functional $E(u)$ to all of H_0. Furthermore, since D_T is dense in H_0 with respect to both the H_0 and H_1' metric, it is not hard to show that

$$\inf_{u \in D_T} E(u) = \inf_{u \in H_0} E(u) = \lambda_1. \tag{2·7}$$

An element $w \neq 0$ in H_0 will be called a *generalized eigenelement* of (2·1) belonging to λ if it satisfies, for every u in H_0, the identity

$$[w, u] = \lambda [w, u]_1; \tag{2·8}$$

the number λ satisfying (2·8) will be called a *generalized eigenvalue* of (2·1). It is evident that every ordinary eigenelement w belonging to the eigenvalue λ of (2·1) is also a generalized eigenelement of (2·1) but the converse need not be true.

Lemma 2·4. *If there exists an element* $w_1 \neq 0$ *in* H_0 *such that*

$$E(w_1) = \frac{[w_1, w_1]}{[w_1, w_1]_1} = \lambda_1, \tag{2·9}$$

then w_1 *is a generalized eigenelement of* (2·1) *belonging to* λ_1 *and* λ_1 *is the least generalized eigenvalue of* (2·1).

Proof. Suppose there exists $w_1 \neq 0$ in H_0 such that (2·9) is valid. Let us define the bilinear functional

$$Q(v, u) = [v, u] - \lambda_1 [v, u]_1$$

for all u and v in H_0. By (2·7), $Q(u,u) \geqslant 0$ for all u in H_0 with $Q(w_1, w_1) = 0$. If θ is any real number, u any element in H_0, and $\gamma = \theta Q(w_1, u)$, then using the fact that $Q(w_1, w_1) = 0$ we get

$$0 \leqslant Q(w_1+\gamma u, w_1+\gamma u) = \gamma Q(u, w_1) + \bar{\gamma} Q(w_1, u) + |\gamma|^2 Q(u,u)$$

which, after substituting the value for γ, reduces to the inequality

$$\theta\,|Q(w_1,u)|^2\{2+\theta Q(u,u)\} \geqslant 0.$$

Since the left member of this inequality changes sign with θ small unless $Q(w_1, u) = 0$, we conclude that $Q(w_1, u) = [w_1, u] - \lambda_1 [w_1, u]_1 = 0$ for all u in H_0. Clearly λ_1 is the least generalized eigenvalue of (2·1). This establishes the validity of lemma 2·4.

Let us note that if $w_1 \in D_T$ and u is chosen to be any element in D_T, then from the last identity we derive the equality $(Tw_1 - \lambda_1 Sw_1, Ku) = 0$ valid for all u in D_T. Since the set KD_T is dense in H, $Tw_1 - \lambda_1 Sw_1 = 0$; i.e. λ_1 is an eigenvalue of (2·1) and w_1 its corresponding eigenelement. That λ_1 is the least eigenvalue follows from (2·6) and the definition of λ_1. Thus, lemma 2·2 is also proved.

Theorem 2·1.† (*a*) *If condition* (α) *is satisfied, then the problem* (2·1) *possesses a generalized eigenelement* $w_1 \neq 0$ *in* H_0 *belonging to* λ_1 *which is the least generalized eigenvalue of* (2·1).

(*b*) *If, in addition, we assume that* S *is either* H_0*-bounded or is such that for some* $\eta_6 > 0$

$$\|Su\| \leqslant \eta_6 |u| \quad (u \in D_T), \tag{2·10}$$

then w_1 *is an ordinary eigenelement of* (2·1) *corresponding to* λ_1 *which is the least eigenvalue of* (2·1).

Proof. Let $\{u_n\colon (u_n \in D_T, n = 1, 2, \ldots)\}$ be a minimizing sequence of elements for $E(u)$ which is normalized in the H_1'-metric. Then $|u_n|_1 = 1$ and $\lim_n E(u_n) = \lim_n |u_n|^2 = \lambda_1$. This implies that the sequence $\{u_n\}$ is bounded in the H_0-norm. Hence, by condition (α), it contains a subsequence which converges in the H_1'-metric. For simplicity, this subsequence will also be denoted by $\{u_n\}$. Let us note that, in view of (2·7), for any u in H_0 we evidently have $|u|^2 \geqslant \lambda_1 |u|_1^2$; hence

$$|u_n - u_m|^2 - \lambda_1 |u_n - u_m|_1^2 \geqslant 0 \tag{2·11}$$

and

$$|u_n + u_m|^2 - \lambda_1 |u_n + u_m|_1^2 \geqslant 0. \tag{2·12}$$

It is not hard to see that the sum of the two left members is

$$2\{[\,|u_n|^2 - \lambda_1 |u_n|_1^2] + [\,|u_m|^2 - \lambda_1 |u_m|_1^2]\}$$

and approaches zero as n and m increase indefinitely. Hence the left members in (2·11) and (2·12) approach zero as n and m increase indefinitely. In particular

$$\{|u_n - u_m|^2 - |u_n - u_m|_1^2\} \to 0 \quad \text{as} \quad n, m \to \infty. \tag{2·13}$$

Since H_1' is complete and $\{u_n\}$ is a convergent sequence in H_1' there exists $w_1 \in H_1'$ such that $|u_n - w_1|_1 \to 0$, as $n \to \infty$. Then clearly $|w_1|_1 = 1$ and $|u_n - u_m|_1 \to 0$ as $n, m \to \infty$. This and (2·13) imply that $|u_n - u_m| \to 0$ as $n, m \to \infty$. Hence the sequence $\{u_n\}$ converges also in the

† Theorem 2·1 offers an extension of the results obtained by Hilbert & Courant [5] and Friedrichs [8] for selfadjoint (s.a.) differential eigenvalue problems and of Mikhlin [25] for abstract s.a. and positive definite eigenvalue problems in H. The proof of theorem 2·1 (*a*) follows essentially the arguments of the above authors.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

H_0-metric. In view of lemma 2·3, the sequence $\{u_n\}$ converges in H_0 to the same element w_1 in H_0. Thus,

$$\lambda_1 = E(w_1) = \frac{[w_1, w_1]}{[w_1, w_1]_1} = \lim_n E(u_n), \tag{2·14}$$

i.e. w_1 *solves the minimum problem*. Lemma 2·4 implies that w_1 is a generalized eigenelement of (2·1) corresponding to λ_1 which is clearly the least generalized eigenvalue of (2·1).

To prove theorem 2·1 (*b*) it is sufficient, in view of lemma 2·2, to show that $w_1 \in D_T$. To do this let us first observe that since, by theorem 2·1 (*a*), w_1 is a generalized eigenelement we have for all u in H_0 and in particular for all u in D_T the identity

$$[w_1, u] = \lambda_1 [w_1, u]_1 \quad (u \in D_T), \tag{2·15}$$

which, on account of the first hypothesis on S and lemma 2·3, can be written in the form

$$[w_1, u] = \lambda_1 (Sw_1, Ku) \quad (u \in D_T). \tag{2·16}$$

Let us recall that by theorem 1·1 the domain D_T of T consists† of all elements in H_0 realizing the minimum of the functional $F(u) = [u, u] - (Ku, f) - (f, Ku)$, where f ranges through all of H. Therefore, if in particular we put $f = \lambda_1 Sw_1$ and use the relation (2·16) we obtain in this case for the value of $F(u)$ the quantity

$$F(u) = [u, u] - [u, w_1] - [w_1, u] = |u - w_1|^2 - |w_1|^2. \tag{2·17}$$

The relation (2·17) shows that $F(u)$ attains its minimum at $u = w_1$. This shows that $w_1 \in D_T$ and therefore, by lemma 2·2, w_1 is an ordinary eigenelement belonging to the smallest eigenvalue λ_1 of (2·1).

If instead of the first we use the second condition (2·10) on S, then as was shown in § 1 the operator S can be extended to closure to $\bar{S}$ defined on all of H_0; hence, to prove theorem 2·1 (*b*), all we need is to replace in (2·16) the operator S by $\bar{S}$. This completes the proof of theorem 2·1.

To determine the succeeding generalized or ordinary eigenvalues of (2·1), lemma 2·1 (*b*), which is also valid for the generalized eigenelements, suggests looking for the second generalized or ordinary eigenelement w_2 of (2·1) among the elements of the set

$$H_0^1 \equiv \{u;\, u \in H_0,\, [w_1, u]_1 = 0\}.$$

It follows from (2·8) that the condition $[w_1, u]_1 = 0$ is equivalent to the condition $[w_1, u] = 0$. Furthermore, if such an element w_2 exists, then

$$\lambda_2 = \inf_{u \in H_0^1} E(u) = E(w_2) = \frac{[w_2, w_2]}{[w_2, w_2]_1}.$$

In general the generalized eigenelements w_{n+1} belonging to λ_{n+1}, if they exist, are such that

$$\lambda_{n+1} = E(w_{n+1}) = \inf_{u \in H_0^n} E(u), \tag{2·18}$$

where $H_0^n \equiv \{u \colon u \in H_0,\, [w_i, u]_1 = 0 \quad (i = 1, 2, \ldots, n)\}$ provided, of course, that the existence of the generalized eigenelements $w_1, w_2, \ldots\ w_n$ belonging to the increasing set of generalized eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ has been already established.

† We recall that, as was stated in remark 2, in §§ 2 and 3 the operator T is understood to be an extension constructed by theorem 1·1 and, hence, it is continuously invertible.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

THEOREM 2·2. *If* (2·1) *is K-p.d. and condition* (α) *is satisfied, then* (2·1) *has countably many generalized eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$, *determined by the variational principle* (2·18) *such that* $\lambda_n \to \infty$ *as* $n \to \infty$; *the corresponding generalized eigenelements* $w_1, w_2, \ldots, w_n, \ldots$ *form a complete set in each of the spaces* H_0, H_1', *and* H.

Proof. We proceed to establish the existence of the generalized eigenelement w_{n+1} belonging to $\lambda_{n+1} = \inf_{u \in H^n} E(u)$ assuming that the existence of the generalized eigenelements $w_1, w_2, \ldots, w_n$ belonging to the corresponding increasing set of the first n generalized eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ is already established. Theorem 2·1 (a) justifies such an assumption.

Let $\{u_i\colon u_i \in D_n = D_T \cap H_0^n \ (i = 1, 2, 3, \ldots)\}$ be a minimizing sequence of elements for $E(u)$ on H_0^n which is normalized in the H_1'-metric. Repeating the same argument as in the proof of the previous theorem we come to the conclusion that there exists $w_{n+1} \in H_0$ such that

$$|u_i - w_{n+1}|_1 \to 0, \quad |u_i - w_{n+1}| \to 0 \quad \text{as} \quad i \to \infty. \tag{2·19}$$

Evidently, $|u_i|_1 \to |w_{n+1}|_1$ and $E(u_i) = |u_i|^2 \to |w_{n+1}|^2$, as $i \to \infty$. Hence $|w_{n+1}|_1 = 1$ and $E(w_{n+1}) = \lambda_{n+1}$. Furthermore, since $u_i \in D_n$, it follows from (2·19) that $w_{n+1} \in H_0^n$. In fact, by (2·19), $[w_{n+1}, w_j]_1 = \lim_i [u_i, w_j]_1 = 0$ and $[w_{n+1}, w_j] = \lim_i [u_i, w_j] = \lambda_j \lim [u_i, w_j]_1 = 0$, $j = 1, 2, \ldots, n$. Thus w_{n+1} solves the minimum problem. To prove that w_{n+1} is a generalized eigenelement of (2·1) belonging to λ_{n+1} we have to show that

$$[w_{n+1}, u] = \lambda_{n+1}[w_{n+1}, u]_1 \tag{2·20}$$

for all u in H_0. Using the same arguments as in the proof of lemma 2·4 and the property of λ_{n+1} we easily derive the validity of (2·20) for all v in H_0^n. Let u be an arbitrary element in H_0 and $w = u - \sum_{j=1}^{n} [u, w_j]_1 w_j$. It is obvious that $w \in H_0^n$ and for this w the relation (2·20) is valid. Since

$$\left.\begin{aligned} [w, w_{n+1}]_1 &= [u, w_{n+1}]_1 - \sum_{j=1}^{n-1} [u, w_j]_1 [w_j, w_{n+1}]_1 = [u, w_{n+1}]_1 \\ \text{and} \quad [w, w_{n+1}] &= [u, w_{n+1}] - \sum_{j=1}^{n} [u, w_j]_1 [w_j, w_{n+1}] = [u, w_{n+1}], \end{aligned}\right\} \tag{2·21}$$

it follows that the identity (2·20) is valid for all u in H_0. Hence w_{n+1} is a generalized eigenelement of (2·1) belonging to the generalized eigenvalue λ_{n+1}, which is clearly the smallest generalized eigenvalue following λ_n. It is obvious that, since the set $H_{01} \subset H_1'$ is separable and infinite dimensional and $\{w_n\}$ is an orthonormal set in H_{01}, the set of corresponding generalized eigenvalues $\{\lambda_n\}$ is countable and has infinity as its only limit point. In fact, if $\lambda_n \to \lambda_0 < \infty$, then $|w_n|^2 = \lambda_n |w_n|_1^2 = \lambda_n \leqslant \lambda_0$, where $\{w_n\}$ is the H_1'-orthonormal sequence of generalized eigenelements belonging to $\{\lambda_n\}$. Thus, $\{w_n\}$ is bounded in the H_0-norm and therefore, by condition (α), is compact in H_1'. But this is impossible since $\{w_n\}$ forms an orthonormal sequence in H_1'.

To prove the second part of theorem 2·2 note that, in view of lemma 2·3, it is sufficient to show that every element in H_{01} can be approximated arbitrarily closely (in the sense of the norms in H_0, H_1', and H) by some linear combination (possibly different in each case) of $\{w_i\}$.

Let u be any element in H_{01} and $u_n = \sum_{i=1}^{n} [u, w_i]_1 w_i$. If we put $v_n = u - u_n$, then $[v_n, w_j]_1 = 0$ and $[v_n, w_j] = 0$ for $j = 1, 2, \ldots, n$. Hence $v_n \in H_0^n$ and therefore, by (2·18),

$$\lambda_{n+1} \leqslant \frac{|v_n|^2}{|v_n|_1^2} \quad \text{or} \quad \frac{|v_n|^2}{\lambda_{n+1}} \geqslant |v_n|_1^2. \tag{2·22}$$

Since $|v_n|^2 = |u-u_n|^2 = |u|^2 - |u_n|^2 < |u|^2$ and $\lambda_{n+1} \to \infty$ as $n \to \infty$, (2·22) yields

$$|u-u_n|_1 \to 0 \quad \text{as} \quad n \to \infty. \tag{2·23}$$

This and (2·3) also show that

$$\|u-u_n\| \to 0 \quad \text{as} \quad n \to \infty. \tag{2·24}$$

Furthermore, if $m > n$, then the inequality $0 \leqslant |u_m - u_n|^2 = |u_m|^2 - |u_n|^2 < |u|^2$ shows that the sequence $\{|u_n|^2\}$ is monotonically increasing and bounded. Hence $|u_m - u_n| \to 0$, as $n, m \to \infty$. Since H_0 is complete, lemma 2·3 and the relations (2·23) and (2·24) imply that

$$|u-u_n| \to 0 \quad \text{as} \quad n \to \infty \tag{2·25}$$

Thus, from (2·23), (2·24), and (2·25) we see that the set of linear combinations of $\{w_i\}$ is dense in H_{01} in the sense of all three norms. This completes the proof of theorem 2·2.

THEOREM 2·3. *If in addition to the hypothesis assumed in theorem* 2·2 *we also assume that S satisfies the conditions of theorem* 2·1 (*b*), *then the problem* (2·1) *has countably many eigenvalues λ_n such that $\lambda_n \to \infty$, as $n \to \infty$, and the corresponding eigenelements w_n form a complete set in each of the spaces H_0, H_1' and H.*

Proof. Theorem 2·3 follows from theorem 2·2 and the fact that under the additional conditions on S all elements w_n belong to D_T. The proof of the latter assertion is exactly the same as the proof of theorem 2·1 (*b*). Lemma 2·2 completes then the proof of theorem 2·3.

Before we go on in our discussion let us remark that in view of the importance of condition (α) it is interesting to know whether it can be replaced by any other condition. To that end we prove

LEMMA 2·5. *Let S satisfy conditions* (2·3) *and* (2·4). *If S is also H_1'-bounded or is closeable and such that*

$$\|Su\| \leqslant \eta_5 |u|_1 \quad (u \in D_T), \tag{2·26}$$

then condition (α) *is satisfied if and only if $T^{-1}S_0$ is compact in H_0, where S_0 is either S or $\bar{S}$.*

Proof. Assume that condition (α) is satisfied.† Let Q_0 be an arbitrarily bounded set in H_0; i.e., there exists an $M_0 > 0$ such that

$$|u| \leqslant M_0 \quad (u \in Q_0). \tag{2·27}$$

By condition (α), Q_0 is compact in the H_1'-metric. Hence we can extract from it a sequence $\{u_n\}$ such that

$$|u_m - u_n|_1 \to 0 \quad \text{as} \quad m, n \to \infty. \tag{2·28}$$

Using (1·4) and lemma 1·1 (*c*) with $T = S_0$ we get

$$\begin{aligned} |T^{-1}S_0(u_m-u_n)|^4 &= (S_0(u_m-u_n), KT^{-1}S_0(u_m-u_n))^2 \\ &\leqslant |u_m-u_n|_1^2 \, (S_0 T^{-1}S_0(u_m-u_n), KT^{-1}S_0(u_m-u_n)). \end{aligned} \tag{2·29}$$

In view of lemma 1·5, lemma 1·4 or (2·26), the inequality (2·27), and (2·28) we derive from (2·29) the relation

$$|T^{-1}S_0(u_m-u_n)|^4 \leqslant 4M_0^2 \eta_5 \|S_0 T^{-1}\| \,.\, \|KT^{-1}\| \,.\, |u_m-u_n|_1^2 \to 0$$

as $m, n \to \infty$. Hence the operator $T^{-1}S_0$ is compact in H_0.

† This part of lemma 2·5 remains also valid if S satisfies the weaker conditions specified in theorem 2·1 (*b*).

To prove the converse assume that $T^{-1}S_0$ is compact in H_0 and Q_{01} is an arbitrary set of elements in H_{01} which is bounded in the H_0-norm; i.e. there exists $M_{01} > 0$ such that

$$|u| \leqslant M_{01} \quad (u \in Q_{01}). \tag{2·30}$$

Since $T^{-1}S_0$ is compact in H_0 we can extract a sequence $\{u_n\}$ from Q_{01} such that

$$|T^{-1}S_0(u_m-u_n) \to 0 \quad \text{as} \quad m, n \to \infty. \tag{2·31}$$

Let us also observe that our conditions on S imply that the formula

$$[u, v]_1 = (S_0 u, Kv) \tag{2·32}$$

is valid for all $u, v \in H_1'$. In fact, it is valid by definition for all $u, v \in D_T$ and by the continuity of S_0, as an operator from H_1' to H, it remains valid for all $u, v \in H_1'$. Therefore, by (2·30) and (2·31),

$$|u_m-u_n|_1^2 = (S_0(u_m-u_n), K(u_m-u_n)) = [T^{-1}S_0(u_m-u_n), u_m-u_n]$$
$$\leqslant 2M_{01}|T^{-1}S_0(u_m-u_n)| \to 0 \quad \text{as} \quad n \to \infty. \tag{2·33}$$

Hence, $\{u_n\}$ is a convergent sequence in the H_1'-metric; i.e. Q_{01} is compact in H_1'-metric and, consequently, condition (α) is satisfied.

Let us note that if $D_K \supseteq D_S$ and the inequalities (2·3) and (2·4) are valid for all $u \in D_S$, then S is K-p.d. and the completion of D_S in the metric (2·5) is a Hilbert space $H_1 \subset H$. Using theorem 1·1 we can extend S to the K-p.d. s.g. Friedrichs extension. Let us denote this extension by S. In this case we can also approach the problem of this section in the following way. Since S^{-1} exists and is defined on all of H we can apply it to equation (2·1) and obtain an equivalent equation

$$S^{-1}Tu - \lambda u = 0 \quad (u \in D_T \subset H_1), \tag{2·1a}$$

in which $A = S^{-1}T$, considered as an operator in H_1, is symmetric and positive. In fact, $[Au, v]_1 = (Tu, Kv) = (Ku, Tv) = (Ku, SAv) = [u, Av]_1$ for all u and v in D_T. Thus, we can introduce the space H_A as the completion of D_T in the metric $[u, v]_A = [Au, v]_1$. Since $[u, v]_A = [Au, v]_1 = (Tu, Kv)$, we see that $H_A = H_0$. The positive operator A can be extended to a selfadjoint operator in H_1. Let A also denote this extention and let us regard the eigenvalues and eigenelements of (2·1) to be the eigenvalues and the corresponding eigenelements of the extended operator A. Since $H_A = H_0$, our condition (α) is unchanged; i.e. every set bounded in the H_A-metric is compact in the H_1-metric. Thus we can extend to the eigenvalue problem (2·1a) with extended A all the results obtained in [25] for the selfadjoint and positive definite eigenvalue problem.

Let us note that the inconvenience of such an approach to the problem (2·1), even in case S is K-p.d., is that we do not know, in general, under what conditions will the eigenvectors of A satisfy the equation (2·1) in the usual sense and furthermore it requires first the extension of S and then the computation of its inverse S^{-1} which generally presents a practical difficulty.

2·3. *K-real eigenvalue problem*

In this section we consider the problem of the existence of eigenvalues and eigenelements of (2·1) under conditions different from those assumed in the previous section. Our basic assumption here is the condition:

(β): *The operator $G(\lambda) \equiv T - \lambda S$ is continuously invertible for all complex numbers λ except possibly for the eigenvalues λ of* (2·1).

Let us observe in passing that, for example, condition (β) will hold if the spectrum $\sigma(2{\cdot}1)$ of (2·1), i.e. the set of all λ for which $G(\lambda)$ is not densely invertible, contains only eigenvalues and S is a bounded operator in H. Indeed, if this is the case, then for any λ in the resolvent set $\rho(2{\cdot}1)$ of (2·1) the range space R_G of $G(\lambda)$ is dense in H and the inverse G^{-1} exists and is bounded on R_G. Consequently, to obtain condition (β) it is thus sufficient to show that $G(\lambda)$ is closed for each $\lambda \in \rho(2{\cdot}1)$. Let $\{u_n\}$ be a sequence in $D_G = D_T$ and f in H such that $\|u_n - u\| \to 0$ and $\|Gu_n - f\| \to 0$, as $n \to \infty$. Then, since S is bounded and $u_n \to u$,

$$Tu_n = Gu_n + \lambda Su_n \to f + \lambda Su, \quad \text{as} \quad n \to \infty.$$

Hence $u \in D_T = D_G$ and $Tu = f + \lambda Su$; i.e. $Gu = f$, G is closed, and $R_G = H$.

It will be seen below that condition (β) allows us to reduce the investigation of the eigenvalue problem (2·1) to the class of symmetric eigenvalue problems investigated by Heuser [12] and Kharazov [18].† To that end we consider the space $D[T]$ introduced in § 1·1. Although $D[T]$ is, in general, not complete it satisfies all the conditions imposed on the space considered in [12, 18]. Let us also define the operator $N = T^{-1}S$ which maps $D[T]$ into itself. It is not hard to see that the definition (1·4) of the metric in $D[T]$ and the K-symmetry of S imply that N is a symmetric operator in $D[T]$. In fact,

$$[Nu, v] = (Su, Kv) = (Ku, TNv) = [u, Nv] \quad (u, v \in D[T]).$$

By means of the operator N the problem (2·1) can be written in $D[T]$ in the form

$$u - \lambda Nu = 0 \tag{2·1 a}$$

or

$$Nu - \mu u = 0, \quad \mu = 1/\lambda, \quad \lambda \neq 0. \tag{2·1 b}$$

It is obvious that if $\lambda \neq 0$ is an eigenvalue of (2·1) with multiplicity m and $\phi_1, \ldots, \phi_m$ are linearly independent eigenelements belonging to λ then, since $|u| = 0$ if and only if $\|u\| = 0$, $\phi_1, \ldots, \phi_m$ are also linearly independent eigenelements in $D[T]$ belonging to the eigenvalue $\mu = 1/\lambda$ of N. The converse is also true provided that the eigenelements belong to $D[T]$. Furthermore, from the relation

$$N - \mu = T^{-1}(S - \mu T), \quad \mu = 1/\lambda, \tag{2·34}$$

we conclude that $\lambda \neq 0 \notin p\sigma(2{\cdot}1)$ if and only if $\mu \neq 0 \notin p\sigma(N)$. The relation (2·34) and condition (β) also imply that if $\mu \neq 0$ is not an eigenvalue of N then the operator $N - \mu$ maps $D[T]$ onto itself; i.e., the operator N satisfies the property (E) imposed by Heuser [12]. Also, the quantity $[Nu, u] = (Su, Ku)$ cannot be equal to zero for all u in $D[T]$ unless $Su = 0$ for all u in D_T for otherwise from the identity

$$\begin{aligned}(S(u+v), K(u+v)) - (S(u-v), K(u-v)) + \mathrm{i}(S(u+\mathrm{i}v), K(u+\mathrm{i}v)) \\ - \mathrm{i}(S(u-\mathrm{i}v), K(u-\mathrm{i}v)) = 4(Su, Kv) \quad (u, v \in D_T),\end{aligned}$$

we would have to conclude that $(Su, Kv) = 0$ for all u and v in D_T. Since KD_T is dense in H, this would imply that $Su = 0$ for all u in D_T.

† Both papers are essentially an extension of the results obtained by Wielandt [36]. In his study Kharazov used the arguments applied by Kamke [14] to the differential and selfadjoint eigenvalue problem (2·1). The distinct and interesting feature of this argument is that the existence proof does not involve the use of the variational principle for the functional $E(u)$.

Thus, the operator N in $D[T]$ satisfies all the conditions imposed by Heuser. Consider also the space $H_0 \supseteq D[T]$ and let $\bar{N}$ denote the closure† of N in $D[T]$ with $D[T] \subset D_{\bar{N}} \subset H_0$. Then using the results of [12] it is now not hard to derive the following two theorems.

THEOREM 2·4. (*a*) *The operator* $\bar{N}$ *is the unique selfadjoint extension of* N.

(*b*) *The spectrum* $\sigma(N)$ *of* N *is a nonempty closed set of real numbers such that* $\sigma(N) = \sigma(\bar{N})$, $p\sigma(N)-\{0\} = p\sigma(\bar{N})-\{0\}$, $c\sigma(N) = c\sigma(\bar{N})$, *and* $r\sigma(N)-\{0\} = r\sigma(\bar{N}) = \phi$, *where* $c\sigma(N)$ *and* $r\sigma(N)$ *denote, respectively, the continuous and the residual spectrum of* N *and* ϕ *is the empty set.*‡

(*c*) *The set* $c\sigma(N)-\{0\}$ *consists exactly of all nonzero limit points of the eigenvalues of* N *which are not eigenvalues themselves.*

THEOREM 2·5. *The operator* N *has at least one nonzero eigenvalue.*

COROLLARY 2·1. *The problem* (2·1) *has at least one nonzero eigenvalue.*

We see from theorems 2·4 and 2·5 that the passage from the eigenvalue problem (2·1) in H to the eigenvalue problem (2·1*b*) in $D[T] \subset H_0$ changes in general the nature of the spectrum $\sigma(2{\cdot}1)$ of (2·1) from the pure point spectrum of (2·1) to the mixed spectrum of N. Furthermore, it was shown in [12] that the set of eigenelements in $D[T]$ belonging to the corresponding nonzero eigenvalues of N or (2·1) is not, in general, large enough to permit the expansion of each element of the form $Nu = T^{-1}Su$ into a Fourier series with respect to these eigenelements.

However, suppose that in addition to condition (β) it is known that

(γ) *The spectrum* $\sigma(N)$ *of* N, *considered as an operator in* $D[T]$, *contains only eigenvalues of finite multiplicity with zero as its sole possible limit point.*

Then from the results obtained in [18] for the operator N in $D[T]$ and from the above discussion it is not hard to derive for the problem (2·1) the following additional theorem.

THEOREM 2·6. (*a*) *The set of eigenvalues* $|\lambda_1| \leqslant |\lambda_2| \leqslant \ldots \leqslant |\lambda_n| \leqslant \ldots$ *with the corresponding set of eigenelements* $w_1, w_2, \ldots, w_n, \ldots$ *normed in the sense of the metric* (1·4) *possesses the following extremal property*: *On the set of elements* $u \in D[T]$ *satisfying the condition* $[u, w_i] = 0$ *or* $[u, w_i]_1 = 0$ *for* $i = 1, 2, \ldots, n-1$, *the absolute value of the functional*

$$\phi(u) = \frac{(Su, Ku)}{(Tu, Ku)} \tag{2·35}$$

attains its maximum equal to $|\mu_n| = 1/|\lambda_n|$, $n = 1, 2, 3, \ldots$, *at* $u = w_n$. *Furthermore, the eigenvalues found by this variational method exhaust entirely the set* $p\sigma(2{\cdot}1)$.

(*b*) *For any* u *in* $D[T]$ *we have the representation* (*generalization of the Hilbert–Schmidt theorem*)

$$T^{-1}Su = \sum_{i=1}^{\infty} \frac{(Tu, Kw_i)}{\lambda_i} w_i = \sum_{i=1}^{\infty} (Su, Kw_i)\, w_i \tag{2·36}$$

which converges in the H_0- *and* H-*metric to* $T^{-1}Su$.

† The operator $\bar{K}$ is called a *closure* or a *trivial closed extension* of K in D_K if it is defined on the set $D_{\bar{K}}(\supset D_K)$ consisting of all elements u in H for which there exists a sequence $\{u_n\}$ in D_K and an element f in H such that $\|u_n - u\| \to 0$ and $\|Ku_n - f\| \to 0$ as $n \to \infty$; in this case we set $\lim_n Ku_n = \bar{K}u = f$.

‡ Let us recall that $\mu \in c\sigma(N) \Leftrightarrow R_{N-\mu}$ is dense in H_0 and $(N-\mu)^{-1}$ is unbounded on $R_{N-\mu}$ while $\mu \in r\sigma(N) \Leftrightarrow R_{N-\mu}$ is not dense in H_0.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

(*c*) *For $\lambda \notin p\sigma(2\cdot1)$ the unique solution u in D_T of the equation $Tu-\lambda Su = f$, for any f in H, is representable in the form*

$$u = T^{-1}f+\lambda \sum_{i=1}^{\infty} \frac{1}{\lambda_i-\lambda}(f, Kw_i)\, w_i \tag{2·37}$$

converging in the H_0-metric.

Let us observe that using the properties of the H_0-bounded operators established in § 1 we can prove a stronger assertion than that contained in theorem 2·6 (*b*) and, at the same time, derive the necessary and sufficient condition for the set of eigenelements, determined by theorem 2·6 (*a*), to be complete both in H_0 and in H.

Lemma 2·6. *If S is H_0-bounded, then for every $u \in D[T]$ the series*

$$Su = \sum_{i=1}^{\infty} (Tu, Kw_i)\, Sw_i = \sum_{i=1}^{\infty} [u, w_i]\, Sw_i \tag{2·38}$$

converges in the H-metric to Su.

Proof. Since $[w_i, w_j] = \delta_{ij}$ we have for any $u \in D[T]$ the Bessel's inequality

$$\sum_{i=1}^{n} |[u, w_i]|^2 \leqslant |u|^2$$

valid for each n. This and the completeness of H_0 imply that there exists an element v in H_0 such that

$$v = \sum_{i=1}^{\infty} [u, w_i]\, w_i. \tag{2·39}$$

Since S is H_0-bounded, lemma 1·4 and (2·39) imply that

$$\left\| Sv-\sum_{i=1}^{n} [u, w_i]\, Sw_i \right\| \leqslant \eta_5 \left| v-\sum_{i=1}^{n} [u, w_i]\, w_i \right| \to 0 \quad \text{as} \quad n \to \infty;$$

i.e. the series

$$Sv = \sum_{i=1}^{\infty} [u,w_i]\, Sw_i \tag{2·40}$$

converges to Sv in H. Using the fact that T^{-1} is bounded in H and $T^{-1}Sw_i = \lambda_i^{-1} w_i$ we derive from (2·40) the series

$$T^{-1}Sv = \sum_{i=1}^{\infty} \frac{[u, w_i]}{\lambda_i} w_i. \tag{2·41}$$

Taking into account the proved formula (2·36) we get

$$T^{-1}(Sv-Su) = \sum_{i=1}^{\infty} \frac{[u, w_i]}{\lambda_i} w_i - T^{-1}Su = 0,$$

whence, in view of (2·40), we obtain the expansion (2·38).

Theorem 2·7. *If S is H_0-bounded, then the set of eigenelements $\{w_i\}$ belonging to the eigenvalues $\{\lambda_i\}$ determined by theorem 2·6 (a) is complete in H_0 if and only if $Su \neq 0$ whenever $u \in H_0$ and $u \neq 0$.*

Proof. Necessity. We have to show that the completeness of $\{w_i\}$ in H_0 implies that $Su \neq 0$ whenever $u \in H_0$ and $u \neq 0$. Assume the contrary; i.e. $\{w_i\}$ is complete and $Su^* = 0$ for some $u^* \neq 0$ in H_0. Then for every i we have

$$[w_i, u^*] = (Tw_i, Ku^*) = \lambda_i(Sw_i, Ku^*) = \lambda_i(Kw_i, Su^*) = 0.$$

This contradicts the assumption that $\{w_i\}$ is complete in H_0. Hence we must have $Su^* \neq 0$.

Sufficiency. Let us first observe that, since $D[T]$ is dense in H_0, to prove the completeness of $\{w_i\}$ in H_0 it is sufficient to show the validity of the expansion

$$u = \sum_{i=1}^{\infty} [u, w_i] w_i \tag{2·42}$$

for every u in $D[T]$. Suppose that $Sw \neq 0$ whenever $w \in H_0$ and $w \neq 0$. Since $\{w_i\}$ is an orthonormal sequence in $D[T] \subset H_0$, the sequence $\left\{\sum_{i=1}^{n} [u, w_i] w_i\right\}$ converges in H_0-norm to some element v in H_0; i.e.

$$v = \sum_{i=1}^{\infty} [u, w_i] w_i.$$

Hence, by lemma 1·4,
$$Sv = \sum_{i=1}^{\infty} [u, w_i] Sw_i. \tag{2·43}$$

Taking into account lemma 2·6 we obtain from (2·43) the equality

$$S(v-u) = \sum_{i=1}^{\infty} [u, w_i] Sw_i - Su = 0.$$

Our assumption then implies that $v = u$; i.e. $\{w_i\}$ is complete in H_0.

COROLLARY 2·2. *If S is H_0-bounded and $Su \neq 0$ whenever $u \in H_0$ and $u \neq 0$, then the set $\{w_i\}$ of eigenelements determined by theorem 2·6 (a) is complete in H.*

Proof. The proof follows from theorem 2·7, the inequality (1·1), and the denseness of H_0 in H.

3. GENERALIZED MOMENTS METHOD (G.M. METHOD)

In this section we indicate briefly the applicability and the unifying property of the g.m. method for the approximate solution

$$Tu - \lambda Su = 0, \tag{3·1}$$

where T is K-p.d. and S is such that $D_S \supset D_T$. This will allow us to extend the applicability of the g.m.-method to non-selfadjoint differential eigenvalue problems of an even and odd order.

3·1. *Some general lemmas*

To justify the g.m.-method we shall first discuss the relevant results concerning the eigenvalue problem

$$Nx - \mu x = 0, \tag{3·2}$$

where N is a compact operator in some Hilbert space X.

Let $\{X_n\}$ be a projectionally complete sequence of expanding finite dimensional subspaces of X [28], i.e. $\|f - P_n f\| \to 0$, as $n \to \infty$, where P_n is an orthogonal projection of X on X_n. By the *approximate eigenvalues and eigenelements* of N we mean the eigenvalues $\mu^{(n)}$ and the corresponding eigenelements $w^{(n)} \in X_n$ of the operator $N_n = P_n N$; i.e. $\mu^{(n)}$ and $w^{(n)}$ are such that

$$N_n w^{(n)} - \mu^{(n)} w^{(n)} = 0. \tag{3·3}$$

Let us first note that, since $\{X_n\}$ is projectionally complete in X and N is compact, the sequence $\{N_n\}$ converges uniformly to N [32]; i.e.,

$$\|N_n - N\| \to 0 \quad \text{as} \quad n \to \infty. \tag{3·4}$$

The following lemmas due to Polsky [27] summarize the relevant properties of the eigenvalue problems (3·2) and (3·3) for compact operators.

LEMMA 3·1. *The set of eigenvalues of N is identical with the set of limit points of all sequences of approximate eigenvalues.*

LEMMA 3·2. *Let μ be an eigenvalue of N and $\mu^{(n)}$ of N_n such that $\mu^{(n)} \to \mu$. From every sequence of normed approximate eigenelements $\{w^{(n)}\}$ belonging to $\mu^{(n)}$ we can extract a convergent subsequence so that its limit w is an eigenelement of N belonging to μ.*

REMARK 3. It was shown in [27] by an example that for an arbitrary compact operator it is not possible to obtain every eigenelement as a limit of approximate eigenelements. However, any element in an invariant subspace of N can be obtained as a limit of a sequence of elements belonging to the corresponding approximate invariant subspaces. This implies that, in particular, if the eigenvalues μ of N and their corresponding approximations $\mu^{(n)}$ of N_n are simple, then any eigenelement of N can be obtained as the limit of the corresponding approximate eigenelements of N_n.

It was already noted that lemma 3·1 in this generality was first proved by Polsky [27] by means of the theory of analytic functions. For the sake of completeness and in view of the difficulty in obtaining his paper we prove below two general lemmas concerning a bounded, not necessarily compact, operator N so that lemma 3·1 and theorem 3·1 (*a*) below will follow from them as corollaries.

LEMMA 3·3. *Let N_n be a sequence of bounded operators on X converging uniformly to a bounded operator N. If $\mu^{(n)} \in \sigma(N_n)$ and $\mu^{(n)} \to \mu^*$ as $n \to \infty$, then $\mu^* \in \sigma(N)$.*

Proof. Suppose, contrary to the assertion, that $\mu^* \notin \sigma(N)$. Then $\mu^* \in \rho(N)$ and, since $\rho(N)$ is open, there is a neighbourhood U of μ^* belonging to $\rho(N)$ and a constant $c > 0$ such that $\|(\mu - N)x\| \geq c\|x\|$ for all μ in U and x in X. Hence

$$\|(\mu - N_n)x\| \geq \|(\mu - N)x\| - \|(N_n - N)x\| \geq (c - \|N - N_n\|)\|x\| \quad (\mu \in U,\ x \in X).$$

Since $\|N_n - N\| \to 0$, as $n \to \infty$, there is an integer $n_0 > 0$ such that $\|N_n - N\| < \frac{1}{2}c$ for $n \geq n_0$. Therefore, $\|(\mu - N_n)x\| > \frac{1}{2}c\|x\|$ and $\|(\mu - N)^{-1}(N_n - N)x\| \leq \frac{1}{2}\|x\|$ for all μ in U and x in X so that for $n \geq n_0$ the equation $y = (\mu - N_n)x = (\mu - N)x + (N - N_n)x$ is uniquely solvable for each μ in U and each y in X. This is, however, impossible since $\mu_n \to \mu^*$, as $n \to \infty$. Thus, $\mu^* \in \sigma(N)$.

LEMMA 3·4. *If $\|N_n - N\| \to 0$ and μ^* is an isolated point of the spectrum $\sigma(N)$, then there exists $\mu^{(n)} \in \sigma(N_n)$ such that $\mu^{(n)} \to \mu^*$, as $n \to \infty$.*

Proof. Assume, contrary to the assertion, that there is a subsequence $\{N_{n_i}\}$ of $\{N_n\}$, $n_1, n_2, \ldots, \to \infty$, and some neighbourhood U of μ^* such that $\sigma(N_{n_i}) \cap U = \phi$. Since μ^* is an isolated point of $\sigma(N)$ we can take U so that it contains no other point of $\sigma(N)$ except μ^*. Let C be a circle with centre μ^* and radius $\eta > 0$ so that C lies in U. It is not hard to see that there is a constant $c_1 > 0$ such that

$$\|(\mu - N)x\| \geq c_1\|x\| \quad (\mu \in C,\ x \in X). \tag{3·5}$$

Furthermore, since for $\mu \in C$ and $x \in X$

$$\|(\mu - N_{n_i})x\| \geq \|(\mu - N)x\| - \|(N - N_{n_i})x\| \geq (c_1 - \|N - N_{n_i}\|)\|x\|$$

and $\|N-N_{n_i}\| \to 0$, there is an integer $n_0 > 0$ such that

$$\|(\mu-N_{n_i})x\| > \tfrac{1}{2}c_1\|x\| \quad (\mu \in C,\ x \in X). \tag{3·6}$$

This shows that the resolvents $R(\mu) \equiv (\mu-N)^{-1}$ and $R_{n_i}(\mu) = (\mu-N_{n_i})^{-1}$ are uniformly bounded for all $\mu \in C$ and $n_i \geqslant n_0$. From the identity

$$R_{n_i}(\mu)-R(\mu) = R(\mu)\,(N-N_{n_i})\,R_{n_i}(\mu) \quad (\mu \in C) \tag{3·7}$$

and the inequalities (3·5) and (3·6) we obtain for all $n_i \geqslant n_0$ and $\mu \in C$

$$\|R_{n_i}(\mu)-R(\mu)\| \leqslant 2/c_1^2\|N-N_{n_i}\| \to 0 \tag{3·8}$$

as $n_i \to \infty$. The relation (3·8) shows that the sequence $\{R_{n_i}(\mu)\}$ of analytic functions converges uniformly on C to the function $R(\mu)$. The properties of analytic functions imply that the resolvent $R(\mu)$, being a uniform limit on C of $\{R_{n_i}(\mu)\}$, will be also analytic in the interior of C. This, however, contradicts the fact that μ^* is an isolated point of $\sigma(N)$, i.e. that μ^* is an isolated singular point of $R(\mu)$ which, as is known, is either a pole of $R(\mu)$ (in which case it is an eigenvalue of N of some finite multiplicity) or an essential singularity of $R(\mu)$. This contradiction completes the proof of lemma 3·3.

Remark 4. It is easy to see now that lemma 3·1 is an immediate consequence of lemmas 3·3 and 3·4. Indeed, since N is compact in X, its spectrum $\sigma(N)$ consists at most of a countable number of eigenvalues of finite multiplicity which are themselves isolated points of $\sigma(N)$; furthermore, the sequence of operators $\{N_n\} \equiv \{P_n N\}$ converges uniformly to N. Thus, if N is compact, the hypotheses of lemmas 3·3 and 3·4 are satisfied; their conclusions yield the validity of lemma 3·1.

3·2. *G.m. method*

In this section we consider the problem of the approximate solution of the eigenvalue problem (3·1) by means of the g.m. method [26]. The essence of this method consists in the following. We choose a system of linearly independent elements $\phi_i \in D_T$, $i = 1, 2, \ldots$, which is complete in H_0, and then construct another system $K\phi_i$.† The approximate eigenelement $w_k^{(n)}$ of (3·1) is taken in the form

$$w_k^{(n)} = \sum_{i=1}^{n} a_{ki}^{(n)}\phi_i, \tag{3·9}$$

where the coefficients $a_{k1}^{(n)}, a_{k2}^{(n)}, \ldots, a_{kn}^{(n)}$ are determined from the condition that

$$(Tw_k^{(n)}-\lambda Sw_k^{(n)}, K\phi_j) = 0 \quad (j = 1, 2, \ldots, n). \tag{3·10}$$

The condition (3·10) leads to linear homogeneous algebraic system

$$\sum_{i=1}^{n}\{t_{ij}-\lambda s_{ij}\}\,a_{ki}^{(n)} = 0 \tag{3·11}$$

† It is easy to see that in this case the system $\{K\phi_i: (i = 1, 2, \ldots)\}$ is complete in H. In fact, if v is any element in H such that $(v, K\phi_i) = 0$, then by theorem 1·1 there is an element $w \in D_T$ such that

$$(v, K\phi_i) = (Tw, K\phi_i) = [w, \phi_i] = 0 \quad \text{for} \quad i = 1, 2, \ldots, n, \ldots.$$

Since $\{\phi_i\}$ is complete in H_i, $w = 0$ and, consequently, $v = Tw = 0$; i.e. $\{K\phi_i\}$ is complete in H. It is known [26] that the converse is also valid if $D_K = D_T$.

with λ as a parameter, where $t_{ij} = (T\phi_i, K\phi_j)$ and $s_{ij} = (S\phi_i, K\phi_j)$. The system (3·11) evidently admits of a nontrivial solution for the $a_{ki}^{(n)}$'s if and only if

$$\det |t_{ij} - \lambda s_{ij}|_{i,j=1}^{n} = 0. \qquad (3{\cdot}12)$$

The vanishing of this determinant provides us with an algebraic equation of the nth degree whose roots $\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_n^{(n)}$ will represent the approximations to the eigenvalues of (3·1). The coefficients $a_{ki}^{(n)}$ in (3·9) are then determined from the homogeneous system

$$\sum_{i=1}^{n} \{t_{ij} - \lambda_k^{(n)} s_{ij}\} a_{ki}^{(n)} = 0. \qquad (3{\cdot}13)$$

The following arguments offer the justification of the g.m. method.

If we define the operator N by setting $N = T^{-1}S$,† then the eigenvalues of N, being defined as those values of μ for which

$$Nu - \mu u = 0, \qquad (3{\cdot}14)$$

will be reciprocals $\mu = 1/\lambda$ of the eigenvalues λ of our original problem (3·1), as is seen at once by operating on both sides of (3·1) with T^{-1}. However, the eigenelements w of N, which by definition lie in H_0, will not in general satisfy (3·1) in the usual sense unless $w \in D_T$; for that it is sufficient that $w \in D_S$. Nevertheless, any eigenelement of N will be considered as an eigenelement of (3·1) (if necessary, in generalized sense). Considered in H_0, the condition (3·10) reduces to $[w_k^{(n)} - \lambda T^{-1}Sw_k^{(n)}, \phi_j] = 0$ $(j = 1, 2, \ldots, n)$ while (3·11) can be written in the form

$$\sum_{i=1}^{n} \{s_{ij} - \mu t_{ij}\} a_{ki}^{(n)} = 0, \qquad (3{\cdot}15)$$

where $s_{ij} = [N\phi_i, \phi_j]$ and $t_{ij} = [\phi_i, \phi_j]$, which is precisely the system one obtains when the method is applied to equation (3·14) in the space H_0.

Let $\{\psi_i\}$ denote the set of elements in D_T obtained by orthonormalizing the sequence $\{\phi_i\}$ in the H_0-metric. It was shown in [26] that the orthonormalization process leads to one and the same result; i.e. $w_k^{(n)}$ will not be changed. However, if we take $w_k^{(n)}$ in the form

$$w_k^{(n)} = \sum_{i=1}^{n} \mathrm{d}_{ki}^{(n)} \psi_i$$

the system (3·15) is then replaced by its equivalent and more convenient form

$$\sum_{i=1}^{n} \gamma_{ij} \mathrm{d}_{ki}^{(n)} - \mu \mathrm{d}_{kj}^{(n)} = 0 \quad (j = 1, 2, \ldots, n), \qquad (3{\cdot}16)$$

where $\gamma_{ij} = [N\psi_i, \psi_j] \quad (i, j = 1, 2, \ldots, n).$

Let H_n be a subspace of H_0 determined by $\psi_1, \psi_2, \ldots, \psi_n$ and P_n the orthogonal projection of H_0 onto H_n. Since $\{\psi_i\}$ is complete in H_0, the sequence $\{H_n\}$ is projectionally complete in H_0. It is easy to see that the equation

$$P_n N w_k^{(n)} - \mu w_k^{(n)} = 0 \qquad (3{\cdot}17)$$

† Let us note that the domain of definition of $T^{-1}S$ is equal to D_S and, therefore, if $D_S \subset H_0$, then by N we shall also denote the extension of $T^{-1}S$ to all of H_0 while in case $D_S \supset H_0$ we will use N to denote the restriction of $T^{-1}S$ to H_0.

is equivalent to the system (3·16); i.e. the applicability of the g.m. method to (3·14) is equivalent to determining $a_{k1}^{(n)}, \ldots, a_{kn}^{(n)}$ from equation (3·17). It is now easy to prove

THEOREM 3·1. *If $T^{-1}S$ can be extended to a compact operator N in H_0, then*

(*a*) *All eigenvalues of* (3·1) (*or of N*) *and only they can be obtained as limits of all possible sequences of approximate eigenvalues determined by the g.m. method* (3·11).

(*b*) *Let $\{\lambda^{(n)}\}$ be a sequence of approximate eigenvalues of* (3·1) *converging to an eigenvalue $\lambda \neq 0$ of* (3·1). *Then the sequence $\{w^{(n)}\}$ of the corresponding normalized approximate eigenelements contains a subsequence converging to an eigenelement w of N belonging to $\mu = 1/\lambda$.*

(*c*) *Assuming additionally that S is K-positive and the sequences $\{\lambda_k^{(n)}\}$ and $\{\lambda_k\}$ of approximate and exact eigenvalues are arranged in the order of increasing magnitude, each as many times as its multiplicity indicates, then $\{\lambda_k^{(n)}\}$ converges to λ_k, as $n \to \infty$.*

Proof. Since N is compact in H_0 and $\{H_n\}$ is projectionally complete in H_0, the sequence of operators $\{N_n\} \equiv \{P_n N\}$ converges uniformly to N in H_0 and $\sigma(N)$ consists at most of a countable number of eigenvalues of finite multiplicity which are themselves isolated points of $\sigma(N)$. In view of this and the discussion preceding theorem 3·1 and the fact that $\sigma(N) = \sigma(T^{-1}S)$, the assertion of theorem 3·1 (*a*) follows immediately from lemmas 3·3 and 3·4.

To prove theorem 3·1 (*b*) note that the sequence $\{w^{(n)}\}$, being bounded, is weakly compact. Let us use $\{w^{(n)}\}$ also to denote a subsequence which converges weakly to some element w in H_0. Evidently the sequence $\{\mu^{(n)}w^{(n)}\}$, where $\mu^{(n)} = 1/\lambda^{(n)}$, converges weakly to μw and, since N is compact, $\{Nw^{(n)}\}$ converges strongly to Nw. Hence, in view of this and (3·3), the equality

$$Nw - \mu^{(n)}w^{(n)} = Nw - N_n w^{(n)} = Nw - Nw^{(n)} + Nw^{(n)} - N_n w^{(n)}$$

and the uniform convergence of N_n to N in H_0 imply that

$$|Nw - \mu^{(n)}w^{(n)}| \leqslant |Nw - Nw^{(n)}| + |N - N_n| |w^n| \to 0$$

as $n \to \infty$; i.e. $\mu^{(n)}w^{(n)} \to Nw$ in the H_0-norm. Consequently, $Nw = \mu w$, as desired.

In order to prove (*c*) note that, since S is K-positive, the operator N is selfadjoint and positive in H_0. So if we characterize the eigenvalues of N as extreme values of the quotient $[Nu, u]/[u, u]$, using the recursive characterization, we have all the information we need to establish theorem 3·1 (*c*). In fact, we have for all normalized elements u orthogonal to the first $(k-1)$ eigenelements of N

$$\begin{aligned} \mu_k^{(n)} &\leqslant \max\,[P_n NP_n u, P_n u] = \max\{[Nu, u] + [P_n NP_n u, u] - [Nu, u]\} \\ &\leqslant \mu_k + \max|[P_n NP_n u, u] - [Nu, u]| \leqslant \mu_k + |P_n N_n P_n - N|. \end{aligned}$$

In the same way we obtain the inequality

$$\mu_k \leqslant \max\,[Nu, u] = \max\{[P_n NP_n u, u] + [Nu, u] - [P_n NP_n u, u]\} \leqslant \mu_k^{(n)} + |N - P_n NP_n|.$$

Thus, $|\mu_k^{(n)} - \mu_k| \leqslant |N - P_n NP_n| \to 0$, as desired. This completes the proof of theorem 3·1.

REMARK 5. Theorem 3·1 shows that the applicability of the g.m. method for the approximate calculation of the eigenvalues and eigenelements of the problem (3·1) depends essentially on the complete continuity of the operator $T^{-1}S$ in H_0. This underlines the usefulness of the discussion in § 1·2 which deals with various conditions under which the operator $T^{-1}S$ can be verified to be compact in H_0.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Furthermore, theorem 3·1 allows us to formulate the eigenvalue problem for non-selfadjoint differential and integrodifferential equations of an even and odd order. For further discussion of this topic as well as for the statement of theorem 3·1 (*a*) see [23, 24].

3·3. *Special cases*

In this section we show that, by specifying the operators K and S, the most important direct methods used in the approximate solution of the eigenvalue problem (3·1) can be deduced as special cases of the g.m. method. In each of these special cases theorem 3·1 remains valid provided, of course, that the operators satisfy the indicated conditions. This will always be assumed to be the case.

(i) *Ordinary Ritz method.* If $K = I$ and $S = I$, then T is selfadjoint and positive definite on D_T, H_0 is the completion of D_T in the metric $[u, v] = (Tu, v)$, and the g.m. method reduces in this case to the ordinary Ritz method

$$\sum_{i=1}^{n} \{(T\phi_i, \phi_j) - \lambda(\phi_i, \phi_j)\} a_{ki}^{(n)} = 0 \tag{3·11$_{\text{i}}$}$$

which has been extensively studied and used in various applications [4, 10, 25, 27].

(ii) *Generalized Ritz method.* In this case we take $S = I$ so that T is K-p.d. and the g.m. method becomes

$$\sum_{i=1}^{n} \{T\phi_i, K\phi_j) - \lambda(\phi_i, K\phi_j)\} a_{ki}^{(n)} = 0. \tag{3·11$_{\text{ii}}$}$$

In particular, when the operator T is selfadjoint and positive definite we may take K to be some root of T.

(iii) *Galerkin method.* If $K = I$, then T is selfadjoint and positive definite so that the g.m. method reduces to the well known [4, 25, 27] Galerkin method determined by the system

$$\sum_{i=1}^{n} \{(T\phi_i, \phi_j) - \lambda(S\phi_i, \phi_i)\} a_{ki}^{(n)} = 0. \tag{3·11$_{\text{iii}}$}$$

The space H_0 is here the same as in the case (i).

(iv) *Moments method.* If $K = T$, then T is continuously invertible, H_0 is the space D_T with the metric $[u, v] = (Tu, Tv)$, and the g.m. method reduces to the ordinary moments of moments

$$\sum_{i=1}^{n} \{(T\phi_i, T\phi_j) - \lambda(S\phi_i, T\phi_j)\} a_{ki}^{(n)} = 0. \tag{3·11$_{\text{iv}}$}$$

Since in this case $H_0 = D_T$, we see that $N = T^{-1}S$ and the eigenelements of N satisfy actually the equation (3·1) in the ordinary sense.

Let us point out at the end of this section that, in view of the wide freedom in the choice of K and the generality of S, the g.m. method can be applied to a much larger class of problems than any of the methods mentioned above which it unifies. Furthermore, when applied to the differential eigenvalue problems the g.m. method will give a better character of convergence than the methods of Garkin or Ritz.

4. A general iterative method

The purpose of this section is to present and investigate a new and fairly general iterative method for the approximate solution of the eigenvalue problem

$$Tu - \lambda Su = 0, \tag{4·1}$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

where T is K-p.d.† and S is K-p. relative to D_T with $D_S \supseteq D_T$; i.e.

$$(Su, Ku) > 0 \quad (u \in D_T,\ u \neq 0). \tag{4·2}$$

It will be shown that the method, which is developed here for the problem (4·1) with unbounded and nonsymmetric operators T and S, has also the property of unifying a number of existing iterative schemes which were developed by various authors [1, 6, 11, 13, 15, 17, 20, 22, 29] mostly for positive definite symmetric matrices and bounded operators. At the same time it extends the applicability of these special methods to our class of problems. It is hoped that this investigation will at least partly fill the gap indicated by Kantorovich [16] and at the same time form the basis for the possible discovery of new and more effective schemes when specified to a particular class of problems.

4·1. *The formulation of the method*

In this section we formulate the method and derive some of its properties. Let us assume at the outset† that there is given a well-known K-p.d. operator C and constants $M_1 \geqslant M_2 > 0$ such that $D_C = D_T$,

$$M_2(Cu, Ku) \leqslant (Tu, Ku) \leqslant M_1(Cu, Ku) \quad (u \in D_T), \tag{4·3}$$

and for every r in H of the form $r = Tu - \lambda Su$, $u \in D_T$, the equation

$$Ch = r \tag{4·4}$$

is uniquely and relatively easily solvable for h. It was observed before that if λ is an eigenvalue of (4·1) and $u \in D_T$ a corresponding eigenvector, then it follows from (4·1) that

$$\lambda(u) = \frac{(Tu, Ku)}{(Su, Ku)}. \tag{4·5}$$

This implies that the eigenvalue problem (4·1) is equivalent to the problem of finding solutions of the operator equation

$$Tu - \lambda(u)\, Su = 0 \quad (u \in D_T). \tag{4·6}$$

We present a method for finding the solutions of (4·1) which is based on this observation and which avoids the transformation of the problem to one with bounded operators.

To solve (4·6) we use the following iteration scheme. If $u_0 \neq 0$ is an arbitrary initial approximation belonging to D_T and u_i is the iterant obtained at the ith step of our process, then the succeeding iterant u_{i+1} is taken in the form

$$u_{i+1} = u_i - t_i C^{-1} r_i \quad (i = 0, 1, 2, \ldots) \tag{4·7}$$

or

$$u_{i+1} = u_i - t_i h_i \quad (i = 0, 1, 2, \ldots), \tag{4·8}$$

where r_i denotes the *residual* at u_i; i.e.

$$r_i \equiv Tu_i - \lambda^i Su_i, \quad \lambda^i \equiv \lambda(u_i) \quad (i = 0, 1, 2, \ldots) \tag{4·9}$$

† Since our concern in this section is to develop a practical procedure for the approximate calculation of eigenvalues and eigenvectors of (4·1) we assume, for the sake of simplicity, that the operators T, S, and C satisfy rather strong conditions. However, later we will indicate how these conditions may be somewhat relaxed. As a matter of fact, most of the results in the first few sections are also valid when T and C are only K-p. Furthermore, it will always be assumed in this section that the eigenvalue problem (4·1) possesses eigenvalues with corresponding eigenvectors which have the properties used below. If necessary, this can always be attained by making recourse to § 2.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

the element h_i is the solution of the equation

$$Ch_i = r_i \quad (i = 0, 1, 2, \ldots) \tag{4·10}$$

and t_i, $i = 0, 1, 2, \ldots$, are real numbers to be determined by some process. Observe that it is not necessary to construct the inverse operator C^{-1} since, as is the case in our method, it is sufficient to know the solutions of $Ch = r$ for special r in H. Let us note that whenever u_i is a nonzero iterant in D_T then u_{i+1} is also a nonzero iterant belonging to D_T so that (4·8) determines a well-defined sequence $\{u_i\}$ in D_T. In fact, if $u_i \neq 0$ and $u_i \in D_T$, then by (4·6), (4·9), and (4.10)

$$(Ch_i, Ku_i) = 0 \quad (i = 0, 1, 2, \ldots) \tag{4·11}$$

and, consequently, by (4·8)

$$(Cu_{i+1}, Ku_{i+1}) = (Cu_i, Ku_i) + t_i^2(Ch_i, Kh_i) \quad (i = 0, 1, 2, \ldots). \tag{4·12}$$

Since C is K-p.d. and $D_C = D_T$, the relations (4·12), (4·10), and (4·8) show that $u_{i+1} \neq 0$ and that $u_{i+1} \in D_T$ whenever $u_i \neq 0$ and $h_i \neq 0$. The last condition may always be assumed to be satisfied for, if $h_i = 0$ for any i, then u_i is an eigenvector of (4·1) and so a solution is obtained.

Remark 6. Before we proceed let us note that the method (4·7) to (4·10) is analogous to the scheme presented by the author [26] for the solution of nonhomogeneous operator equations in H. Furthermore, it is not hard to see that, in view of (4·3), the element h_i determined by (4·10) is equal in direction to the gradient of the functional $\lambda(u)$ at $u = u_i$ relative to the metric in the space H_2, where H_2 is the completion of D_T in the metric

$$[u, v]_2 = (Cu, Kv), \quad |u|_2^2 = [u, u]_2 \quad (u, v \in D_T). \tag{4·13}$$

Hence we see that under present conditions the method is analogous to some sort of a gradient method and as such it is an extension of the procedure developed in [11] for the solution of (4·1) with finite selfadjoint and positive definite matrices.

For the sake of completeness let us compute the gradient of $\lambda(u)$ in H_2. Since

$$\operatorname{grad}\lambda(u) = \{1/(Su, Ku)\}[\operatorname{grad}(Tu, Ku) - \lambda(u)\operatorname{grad}(Su, Ku)],$$

it is sufficient to find the grad (Tu, Ku) and grad (Su, Ku) in H_2. To compute grad (Tu, Ku) take fixed vector u and h in D_T and consider the function $f(t) = (T(u+th), K(u+th))$ for real t. A simple calculations shows that

$$\mathrm{d}f/\mathrm{d}t|_{t=0} = 2\mathscr{R}(Tu, Kh)$$

and
$$|\mathscr{R}(Tu, Kh)| \leqslant |(Tu, Kh)| = |[C^{-1}Tu, h]_2| \leqslant |C^{-1}Tu|_2\,|h|_2.$$

Hence, for a fixed $|h|_2$ the equality is attained for such h for which h is proportional to $C^{-1}Tu$. Thus, in H_2-metric, grad $(Tu, Ku) = 2C^{-1}Tu$. Similarly, grad $(Su, Ku) = 2C^{-1}Su$ (in H_2). Combining this with our formula for grad $\lambda(u)$ (in H_2) we get the formula:

$$\operatorname{grad}\lambda(u) = \{2/(Su, Ku)\}\,C^{-1}\{Tu - \lambda(u)\,Su\} \quad (\text{in } H_2);$$

i.e.
$$\operatorname{grad}\lambda(u) = \{2/(Su, Ku)\}\,h,$$

where h is determined by $Ch = Tu - \lambda(u)\,Su$.

We first show that under very general conditions on $\{t_i\}$ the sequence $\{\lambda^i\}$ converges. To find these conditions we consider the change

$$\Delta\lambda = \lambda(u) - \lambda(u - th) \quad (h \neq 0), \tag{4·14}$$

where for notational reasons we omitted 'i' on u, h, and t. In view of (4·8), a simple calculation shows that

$$\Delta\lambda = \frac{-2t\lambda(u)\,\mathscr{R}(Sh, Ku) + t^2\lambda(u)\,(Sh, Kh) + 2t\mathscr{R}(Tu, Kh) - t^2(Th, Kh)}{(Su, Ku) - 2t\mathscr{R}(Sh, Ku) + t^2(Sh, Kh)}$$

$$= \frac{t^2\lambda(u)\,(Sh, Kh) + 2t(Ch, Kh) - t^2(Th, Kh)}{(Su, Ku) - 2t\mathscr{R}(Sh, Ku) + t^2(Sh, Kh)}$$

$$= \frac{\mathrm{d}\{a\lambda(u)\,t^2 + 2t - et^2\}}{1 - 2bt + ct^2} = \mathrm{d}p(u, t), \tag{4·15}$$

where we define
$$\mathrm{d} \equiv \frac{(Ch, Kh)}{(Su, Ku)}, \quad p(u, t) \equiv \frac{2t - ast^2}{1 - 2bt + ct^2} \tag{4·16}$$

with
$$\left.\begin{aligned} a &\equiv \frac{(Sh, Kh)}{(Ch, Kh)}, \quad b \equiv \frac{\mathscr{R}(Sh, Ku)}{(Su, Ku)}, \\ c &\equiv \frac{(Sh, Kh)}{(Su, Ku)}, \quad e \equiv \frac{(Th, Kh)}{(Ch, Kh)}, \\ s &\equiv \lambda(h) - \lambda(u). \end{aligned}\right\} \tag{4·17}$$

Formula (4·15) shows that for a given u in D_T the change $\Delta\lambda$ depends on the behaviour of $p(u, t)$, as a function of t, when t varies through the entire real axis from $-\infty$ to $+\infty$, so that the mode and the rate of convergence of the sequence $\{\lambda^i\}$ will depend on $p(u, t)$.

4·2. *The graph of $p(u, t)$*

To graph p note that $p = 0$ when $t = 0$ and $t = 2/as$. Moreover, since

$$1 - 2bt + ct^2 = |u - th|_1^2/|u|_1^2 > 0 \quad \text{for all } t,$$

it is clear that p is a continuous function of t for all real t including $t = \pm\infty$ for, in the latter case,

$$\lim_{t\to-\infty} p(u, t) = \lim_{t\to\infty} p(u, t) = -as/c.$$

Let us add that $p = -as/c$ also for $t = as/2(abs - c)$. Furthermore, a simple calculation shows that

$$\frac{\mathrm{d}p}{\mathrm{d}t} = 2\,\frac{(abs - c)\,t^2 - ast + 1}{(1 - 2bt + ct^2)^2}. \tag{4·18}$$

Hence $\mathrm{d}p/\mathrm{d}t = 0$ for those values of t for which

$$(abs - c)\,t^2 - ast + 1 = 0. \tag{4·19}$$

Since $p = -as/c$ for $t = \pm\infty$, the equation $\mathrm{d}p/\mathrm{d}t = 0$ has at least one nonzero real root. In fact, it is not hard to see that it has two distinct roots $\tilde{\theta}$ and θ given by

$$\left.\begin{aligned} \tilde{\theta} &= \frac{sa + \surd\{(sa)^2 - 4(abs - c)\}}{2(abs - c)}, \\ \theta &= \frac{sa - \surd\{(sa)^2 - 4(abs - c)\}}{2(abs - c)}. \end{aligned}\right\} \tag{4·20}$$

Indeed, if θ' is a solution of (4·19), then

$$abs - c = \frac{as\theta' - 1}{\theta'^2}$$

and
$$\theta'^2[(sa)^2-4(abs-c)] = (as\theta')^2-4\theta'^2(abs-c)$$
$$= (as\theta')^2-4(as\theta'-1)$$
$$= (as\theta'-2)^2 \tag{4·21}$$

with
$$p(u,\theta') = \frac{(2-as\theta')\,\theta'}{(2-as\theta')\,(1-b\theta')} = \frac{\theta'}{1-b\theta'} \tag{4·22}$$

for the fact that the denominator of p is positive for all t implies that $2-as\theta' \neq 0$. This and (4·21) imply that the discriminant $(sa)^2-4(abs-c) > 0$. Hence the equation (4·19) has two distinct real roots given by (4·20) or equivalently by

$$\tilde{\theta} = \frac{2}{sa-\surd\{(sa)^2-4(abs-c)\}}, \quad \theta = \frac{2}{sa+\surd\{(sa)^2-4(abs-c)\}}. \tag{4·23}$$

The foregoing discussion shows that the graph of $p(u,t)$ depends on the algebraic sign of $(abs-c)$ and s. Hence the following five cases are sufficient to describe it.

Case 1. If $s = 0$, then

$$p = \frac{2t}{1-2bt+c^2}, \quad \lim_{t\to-\infty} p(u,t) = \lim_{t\to\infty} p(u,t) = 0,$$

$$\tilde{\theta} = -\left(\frac{1}{c}\right)^{\frac{1}{2}}, \quad \theta = \left(\frac{1}{c}\right)^{\frac{1}{2}}, \quad \min_t p = p(u,\tilde{\theta}) = \frac{-1}{\surd c+b} < 0$$

and
$$\max_t p = p(u,\theta) = \frac{1}{\surd c+b} > 0.$$

The graph of p in this case is given by figure 1.

Case 2*i*. If $s > 0$ and $abs-c > 0$, then it is not hard to see that in this case

$$0 < \theta < \gamma' < \gamma'' < \tilde{\theta},$$

where
$$\gamma' \equiv \frac{2}{as}, \quad \gamma'' \equiv \frac{as}{2(abs-c)},$$

$\tilde{\theta}$ and θ are given by (4·23); and

$$\min_t p = p(u,\tilde{\theta}) = \frac{\tilde{\theta}}{1-b\tilde{\theta}} < 0 \quad \text{and} \quad \max p = p(u,\theta) = \frac{\theta}{1-b\theta} > 0.$$

The graph of p is given by figure 2*i*.

Case 2*j*. If $s > 0$ and $abs-c < 0$, then $\tilde{\theta} < \gamma'' < 0 < \theta < \gamma'$,

$$\min_t p = p(u,\tilde{\theta}) = \frac{\tilde{\theta}}{1-b\tilde{\theta}} < 0 \quad \text{and} \quad \max_t p = p(u,\theta) = \frac{\theta}{1-b\theta} > 0.$$

The graph of $p(u,t)$ is given by figure 2*j*.

Case 3*i*. If $s < 0$ and $abs-c > 0$, then $\theta < \gamma'' < \gamma' < \tilde{\theta} < 0$,

$$\min_t p = p(u,\tilde{\theta}) = \frac{\tilde{\theta}}{1-b\tilde{\theta}} < 0, \quad \text{and} \quad \max_t p = p(u,\theta) = \frac{\theta}{1-b\theta} > 0$$

with figure 3*i* representing the graph of p.

Case 3*j*. If $s<0$ and $abs-c<0$, then $\gamma'<\tilde{\theta}<0<\gamma''<\theta$ with

$$\min_t p = p(u,\tilde{\theta}) = \frac{\tilde{\theta}}{1-b\tilde{\theta}} < 0 \quad \text{and} \quad \max p = \max p(u,\theta) = \frac{\theta}{1-b\theta} > 0,$$

and the graph given by figure 3*j*.

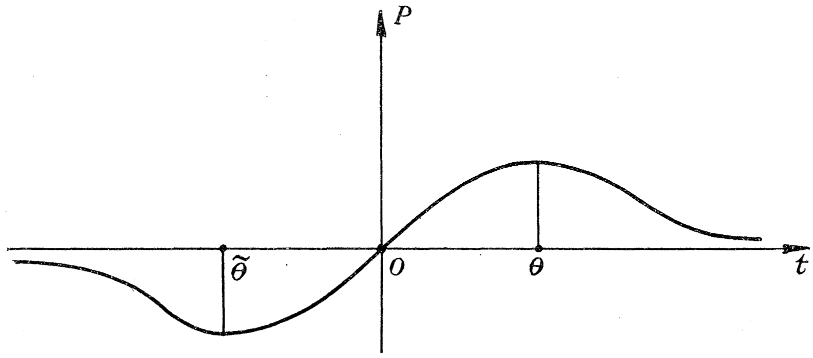

FIGURE 1

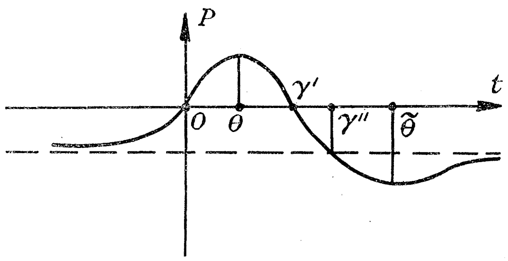

FIGURE 2*i*

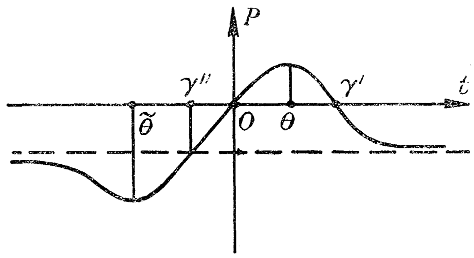

FIGURE 2*j*

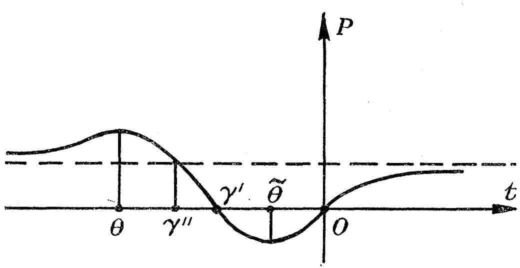

FIGURE 3*i*

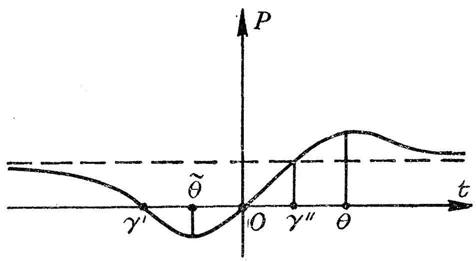

FIGURE 3*j*

REMARK 7. We see from the graphs that, for a given u in D_T, the function $p(u,t)$ attains its maximum value which is positive at the second root θ of equation (4·19). This remark will be useful when we discuss the choice of the optimum parameter.

4·3. *The convergence of the method*

Since $d>0$, it is seen from (4·15) and the graphs of $p(u,t)$ that, for a given u, $\Delta\lambda>0$ for all t which are such that $t\in(0,\infty)$ when $s=0$ (case 1), $t\in(0,\gamma')$ when $s>0$ (case 2), and $t\in C[\gamma',0]$ when $s<0$ (case 3), where $C[\gamma',0]$ denotes the complement of $[\gamma',0]$ in the set R of real numbers.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The above remark indicates that at each step of the process the range of t for which $\Delta\lambda > 0$ depends upon the iterant u. Let us first show that for a given eigenvalue problem with fixed operators T, S, K, and C we can determine an interval of variation of t for which $\Delta\lambda > 0$ independently of u.

LEMMA 4·1. *If T, S, K, and C satisfy the conditions of § 4·1, then $\Delta\lambda > 0$, independently of u in D_T, for any t in the interval*

$$0 < t < 2/M_1, \tag{4·24}$$

where M_1 is the constant determined by (4·3). If in addition we assume that $D_S = D_C$ and that there exist two constants $m_1 > 0$ and $l_1 > 0$ such that

$$(Tu, Ku) \geqslant m_1(Su, Ku) \quad (u \in D_T), \tag{4·25}$$

and

$$(Su, Ku) \geqslant l_1(Cu, Ku) \quad (u \in D_T), \tag{4·26}$$

then $\Delta\lambda > 0$, independently of u, for any t in the interval

$$0 < t < \frac{2}{M_1 - m_1 l_1}. \tag{4·27}$$

Proof. We have seen above that, when $s \leqslant 0$, $\Delta\lambda > 0$ for any $t > 0$ and that, when $s > 0$, $\Delta\lambda > 0$ for t in $0 < t < \gamma'$. Thus, to prove the first part of lemma 4·1, it is sufficient to estimate the quantity γ' in the latter case. Since, in virtue of (4·16) and (4·17),

$$0 < \frac{2}{\gamma'} = as = \frac{(Th, Kh)}{(Ch, Kh)} - \frac{(Sh, Kh)}{(Ch, Kh)} \frac{(Tu, Ku)}{(Su, Ku)} < \frac{(Th, Kh)}{(Ch, Kh)} \tag{4·28}$$

we obtain from (4·28) and (4·3) the inequality $0 < 2/\gamma' < M_1$ or $\gamma' > 2/M_1$ valid for all u in D_T. This is precisely (4·24).

To prove the second part of lemma 4·1 note that when our operators satisfy the additional conditions (4·25) and (4·26) then from (4·28) we obtain the inequality $0 < 2/\gamma' \leqslant M_1 - l_1 m_1$ or $\gamma' \geqslant 2/(M_1 - l_1 m_1)$, i.e. (4·27).

REMARK 8. The additional conditions (4·25) and (4·26) will be satisfied, for example, when we deal with finite matrices or bounded operators in H. In this case lemma 4·1 generalizes the corresponding results for symmetric and positive definite matrices [6, 13].

LEMMA 4·2. *If at each step of the process (4·8) to (4·10) the sequence $\{t_i\}$, $t_i \neq 0$, is so chosen that*

$$p_i \equiv p(u_i, t_i) \geqslant \delta > 0 \quad (h_i \neq 0, i = 0, 1, 2, \ldots) \tag{4·29}$$

then (*a*) *The sequence $\{\lambda^i\}$ converges monotonically to some real number λ^*.*†

(*b*) *The series $\sum_{i=0}^{\infty} d_i$ converges.*

(*c*) *The sequence $\{u^i\} \equiv \left\{\frac{u_i}{|u_i|_1}\right\}$ is bounded in the H_2-norm and*

$$|C^{-1}(Tu^i - \lambda^i S u^i)|_2 \to 0 \quad \text{as} \quad i \to \infty. \tag{4·30}$$

Proof. Let $t_i \neq 0$ be so chosen that (4·29) is satisfied. Since $d_i > 0$, the definition (4·14) and formula (4·15) show that

$$\lambda^i - \lambda^{i+1} = d_i p_i \geqslant d_i \delta > 0. \tag{4·31}$$

† Since $d_i > 0$, it follows from (4·15) that for the validity of lemma 4·2 (*a*) it is sufficient to assume that $p_i > 0$ for $i = 0, 1, 2, \ldots$.

Hence $\{\lambda^i\}$ is a monotonically decreasing sequence of positive numbers and, therefore, converges to some number λ^*.

To prove lemma 4·2 (b) note that by the recurring relation (4·15)

$$\lambda^{i+1} = \lambda^i - d_i p_i = \lambda^{i-1} - d_{i-1} p_{i-1} - d_i p_i = \ldots = \lambda^0 - \sum_{j=0}^{i} d_j p_j \tag{4·32}$$

and consequently the convergence of $\{\lambda^i\}$ implies the convergence of the series $\sum_{j=0}^{\infty} d_j p_j$. In virtue of condition (4·29), this implies the validity of lemma 4·2 (b).

Consider the sequence $\{u^i\}$. Clearly it is bounded in the H_1'-norm since $|u^i|_1 = 1$ $(i = 0, 1, 2, \ldots)$. Moreover, by lemma 4·2 (a) and (4·3), $\{u^i\}$ is bounded in the H_2-metric for

$$|u^i|_2^2 = (Cu^i, Ku^i) = \frac{(Cu_i, Ku_i)}{(Su_i, Ku_i)} = \frac{(Cu_i, Ku_i)}{(Tu_i, Ku_i)} \cdot \lambda^i \leqslant \frac{\lambda^0}{M_2}.$$

We see that in this case the upper bound of $\{u^i\}$ depends on the initial element u_0. To complete the proof note that (4·30) follows from (b). In fact, from (b) we derive

$$d_i = \frac{(Ch_i, Kh_i)}{(Su_i, Ku_i)} = |C^{-1}(Tu^i - \lambda^i Su^i)|_2^2 \to 0 \quad \text{as} \quad i \to \infty.$$

Let us remark that when u_0 is chosen arbitrarily the sequence $\{\lambda^i\}$ may converge to zero, as $i \to \infty$. However, since by (4·32) the limit $\lambda^* = 0$ if and only if

$$\lambda_0 = \sum_{i=0}^{\infty} d_i p_i \tag{4·33}$$

it is rather unlikely that (4·33) may be satisfied in practice.

In order to replace $\{u^i\}$ by $\{u_i\}$ in the above lemma we need, in general, some additional hypothesis on $\{t_i\}$.

Lemma 4·3. *If in addition to* (4·29) *we assume that for some* $\delta_2 > 0$

$$t_i^2 < \delta_2 \quad (i = 0, 1, 2, \ldots), \tag{4·34}$$

then (a) *There exists a constant* $c > 0$ *such that* $\lim_i (Cu_i, Ku_i) = \lim_i |u_i|_2^2 = c$.

(b) $0 < |u_0|_2^2 < |u_1|_2^2 < \ldots < |u_i|_2^2 < \ldots \leqslant c$.

(c) $|C^{-1}(Tu_i - \lambda^i Su_i)|_2 \to 0,$ *as* $i \to \infty$.

Proof. Using (4·12) recursively we obtain the relation

$$|u_{i+1}|_2^2 = |u_i|_2^2 (1 + t_i^2 g_i) = |u_{i-1}|_2^2 (1 + t_{i-1}^2 g_{i-1})(1 + t_i^2 g_i) = \ldots = |u_0|_2^2 \prod_{j=0}^{i} (1 + t_j^2 g_j), \tag{4·35$_0$}$$

where $g_j = |h_j|_2^2 / |u_j|_2^2$. It follows from it that $\{|u_i|_2^2\}$ forms a monotonically increasing sequence which is bounded below by $|u_0|_2^2$. It is known that the product $\prod_{j=0}^{i} (1 + t_j^2 g_j)$ converges, as $i \to \infty$, if and only if the series $\sum_{j=0}^{\infty} t_j^2 g_j$ converges. In view of (4·34), the series $\sum_{j=0}^{\infty} t_j^2 g_j$ converges if $\sum_{=0}^{\infty} g_j$ converges. Since, by (4·3), lemma 4·2 (a), and the assumption $\lambda^* > 0$,

$$g_j = \frac{d_j}{\lambda^i} \cdot \frac{(Tu_j, Ku_j)}{(Cu_j, Ku_j)} \leqslant \frac{M_1}{\lambda^*} d_j \quad (j = 0, 1, 2, \ldots), \tag{4·35}$$

the convergence of $\sum_{j=0}^{\infty} g_j$ follows from (4·35) and lemma 4·2 (*b*). Thus, there exists a constant $c > 0$ such that

$$\prod_{j=0}^{\infty} (1+t_j^2 g_j) = \lim_i |u_{i+1}|_2^2 = c.$$

Evidently, c is the least upper bound of the monotonically increasing sequence $\{|u_{i+1}|_2^2\}$. This implies lemma 4·3 (*b*). The assertion (*c*) follows immediately from (*b*) and the fact that $\lim_i g_i = 0$. Indeed, by lemma 4·3 (*b*),

$$\lim |C^{-1}(Tu_i - \lambda^i Su_i)|_2^2 = \lim_i |h_i|_2^2 = \lim_i \{g_i |u_i|\}_2^2 \leqslant c \lim_i g_i = 0.$$

This completes the proof of lemma 4·3.

Suppose there exists a constant $l_2 > 0$ such that

$$(Cu, Ku) \geqslant l_2(Su, Ku) \quad (u \in D_T). \tag{4·36}$$

The inequalities (4·3) and (4·36) imply that the operators $C^{-1}T$ and $C^{-1}S$, considered on D_T, are bounded in the H_2-norm, where C is understood in the extended sense. Let W_0 and N_0 denote the unique extensions of $C^{-1}T$, $C^{-1}S$ to all of H_2, respectively.

Theorem 4·1. *Let T, S, K, and C satisfy conditions* (4·3) *and* (4·36) *and let $\{t_i\}$ be such that* (4·29) *and* (4·34) *are fulfilled. Then for any $u_0 \neq 0$ in D_T the sequence $\{\lambda^i\}$, where $\{u_i\}$ is determined by* (4·8)–(4·10), *converges to a number $\lambda^* \geqslant M_2 l_2 > 0$ and $|C^{-1}(Tu_i - \lambda^i Su_i)|_2 \to 0$, as $i \to \infty$. If the sequence $\{u_i\}$ does not converge weakly to zero (in H_2), then there is $u^* \neq 0$ in H_2 so that*

$$W_0 u^* - \lambda^* N_0 u^* = 0. \tag{4·37}$$

If, in addition, $D_S \supseteq H_2$, then λ^* is an eigenvalue of (4·1) *and u^* its eigenvector.*

Proof. The first part of theorem 4·1 follows from lemma 4·3 and (4·36) since, by (4·36) and (4·3),

$$\lambda^i = \frac{|u_i|^2}{|u_i|_2^2} \frac{|u_i|_2^2}{|u_i|_1^2} \geqslant M_2 l_2 > 0.$$

To prove the second part note that since by lemma 4·3 (*b*) the sequence $\{u_i\}$ is bounded (in H_2) we can extract from it a weakly convergent subsequence $\{u_k\}$ (in H_2). By hypothesis, the weak limit u^* (in H_2) of $\{u_k\}$ is not zero. Hence, in view of the boundedness of $C^{-1}T$ and $C^{-1}S$ (in H_2), the sequence $\{z_k\} \equiv \{C^{-1}Tu_k - \lambda^k C^{-1}Su_k\}$ converges weakly to $W_0 u^* - \lambda^* N_0 u^*$ (in H_2). But, by the first part of theorem 4·1, $\{z_k\}$ converges also strongly to zero (in H_2). Hence the equation (4·37) is satisfied.

To show the last assertion of our theorem observe that, in virtue of (4·3), T satisfies all the conditions of theorem 1·3. Hence it has a K-p.d. s.g.F.e. T_0 given by $T_0 = CW_0$. Furthermore, the condition $D_S \supseteq H_2$ implies that $N_0 = C^{-1}S$ and that (4·37) can be written in the form $W_0 u^* = \lambda^* C^{-1} S u^*$. Applying C to both sides of this equation and using the fact that $CW_0 = T_0$ we obtain the equality

$$T_0 u^* - \lambda^* S u^* = 0,$$

i.e. λ^* is an eigenvalue of (4·1) and u^* is its corresponding eigenvector.

Remark 9. Let us remark that so far no use has been made of the freedom in choosing the operator K. Thus, if K is chosen to be some closed operator with $D_K = D_T$, then by theorem 1·2 and the inequality (4·3) T and C are continuously invertible operators for which the inequality (1·12) is valid. Hence, in this case, $D[T] = H_0 = H_2$ so that the condition $D_S \supseteq H_2$ is always satisfied. Of course, other choices of K are possible so that $D_S \supseteq D_K \supseteq H_2$. However, we will not dwell on it now.

Let us remark that if $C^{-1}S$ is also compact in H_2, then $|C^{-1}Su_k - C^{-1}Su^*|_2 \to 0$ and $|T^{-1}Cz_k|_2 \to 0$, as $k \to \infty$. Consequently, the boundedness of $T^{-1}C$ and $T^{-1}S$ (in H_2) and lemma 4·3 imply that

$$|u_k - u^*|_2 \leqslant |T^{-1}Cz_k|_2 + |\lambda^k - \lambda^*| \, |T^{-1}Su_k|_2 + \lambda^* |T^{-1}C|_2 \, |C^{-1}Su_k - C^{-1}Su^*|_2 \to 0, \quad (4·38)$$

as $k \to \infty$; i.e. in this case $\{\lambda^i\}$ converges to an eigenvalue λ^* of (4·1) and $\{u_k\}$ converges strongly (in H_2) to the eigenvector u^* corresponding to λ^*.

At the end of this section let us point out that in general we cannot assert that λ^* is equal to the smallest eigenvalue

$$\lambda_1 \left(= \inf_{u \in D_T} \frac{(Tu, Ku)}{(Su, Ku)} \right)$$

of (4·1). However, if we assume that λ_1 is an isolated eigenvalue of (4·1) of finite multiplicity, then for some special choices of the operator C with some additional condition on $\{t_i\}$ we shall show that $\lambda^* = \lambda_1$ and that $\{u_i\}$ converges to some vector in the eigenspace belonging to λ_1.

4·4. *Two simple choices of C*

In this section we show that if C is taken to be either T or S, then the assertions of the last paragraph indeed take place.

Let us denote by H_0^1 the space of eigenvectors belonging to λ_1 and by ${}^\perp H_0^1$ the orthogonal complement of H_0^1 in H_0. Assume that $u_0 \notin {}^\perp H_0^1$ and that $\{t_i\}$ is bounded from below by some constant $\delta_1 > 0$; i.e.

$$u_0 = \xi_0 w + v_0 \quad (w \in H_0^1, \ |w| = 1, \ \xi_0 > 0, \ v_0 \in D_T \cap {}^\perp H_0^1) \quad (4·39)$$

and

$$t_i \geqslant \delta_1 > 0 \quad (i = 0, 1, 2, \ldots). \quad (4·40)$$

Let us define the closeness of two elements u and z in H_2 by

$$\sin^2 [u; z]_2 = 1 - \left[\frac{u}{|u|_2}, \frac{z}{|z|_2} \right]_2^2. \quad (4·41)$$

It is obvious that $\sin^2 [u; z]_2 = 1$ when u is orthogonal to z (in H_2) and that $\sin^2 [u; z]_2 = 0$ when u and z have the same direction. Furthermore, it is not hard to show that when $u = \xi w + v$, where $|w|_2 = 1$ and $[w, v]_2 = 0$, then

$$\sin^2 [u; w]_2 = |v|_2^2 / |u|_2^2. \quad (4·42)$$

Accordingly we shall say that a sequence of elements $\{u_i\}$ converges 'in direction' to an element w in H_2-metric if

$$\lim_{i \to \infty} \sin^2 [u_i; w]_2 = \lim_{i \to \infty} \left(1 - \frac{[u_i, w]_2^2}{|u_i|_2^2 \, |w|_2^2} \right) = 0. \quad (4·43)$$

Lemma 4·4. *If C is either T or S and the initial approximation u_0 is of the form* (4·39), *then the sequence $\{u_n\}$ $(n = 0, 1, 2, \ldots)$, determined by* (4·8) *to* (4·10) *is of the form*

$$u_n = \xi_n w + v_n \quad (n = 0, 1, 2, \ldots), \quad (4·44)$$

where $w \in H_0^1$, $|w| = 1$, $\xi_n > 0$, and $v_n \in D_T \cap {}^\perp H_0^1$.

Proof. In view of mathematical induction and our assumption (4·39) it is sufficient to prove (4·44) for $n = i+1$, assuming its validity for $n = i$, i.e. assume that $u_i = \xi_i w + v_i$, where $|w| = 1$, $w \in H_0^1$, $\xi > 0$, and $v_i \in D_T \cap {}^\perp H_0^1$. Using this assumption and the formulas (4·8) to (4·10) we obtain in a straightforward manner that when $C = T$, then

$$u_{i+1} = u_i - t_i h_i = \xi_i \left[1 + t_i \left(\frac{\lambda^i}{\lambda_1} - 1\right)\right] w + (1-t_i) v_i + t_i \lambda^i T^{-1} S v_i$$

$$= \xi_{i+1} w + v_{i+1}, \tag{4·45}$$

where $\xi_{i+1} = \xi_i \left[1 + t_i \left(\frac{\lambda^i}{\lambda_1} - 1\right)\right] > 0$ and $v_{i+1} = (1-t_i) v_i + t_i \lambda^i T^{-1} S v_i$,

and when $C = S$, then

$$u_{i+1} = u_i - t_i h_i = \xi_i [1 + t_i(\lambda^i - \lambda_1)] w + (1 + t_i \lambda^i) v_i - t_i S^{-1} T v_i$$

$$= \tilde{\xi}_{i+1} w + \tilde{v}_{i+1}, \tag{4·46}$$

where $\tilde{\xi}_{i+1} = \xi_i [1 + t_i(\lambda^i - \lambda_1)] > 0$ and $\tilde{v}_{i+1} = (1 + t_i \lambda^i) v_i - t_i S^{-1} T v_i$.

It follows from (4·45) and (4·46) that lemma 4·4 will be proved if we show that

$$[v_{i+1}, u] = [\tilde{v}_{i+1}, u] = 0$$

for any $u \in H_0^1$. But this is obvious since $v_i \in {}^\perp H_0^1$ and hence

$$[v_{i+1}, u] = (1-t_i)\,[v_i, u] + t_i \lambda^i (S v_i, K u) = 0$$

and
$$[\tilde{v}_{i+1}, u] = \lambda_1 [\tilde{v}_{i+1}, u]_1 = \lambda_1 (1 + t_i \lambda^i)\,[v_i, u]_1 - t_i (T v_i, K u) = 0.$$

THEOREM 4·2. *If $C = T$, λ_1 is an isolated eigenvalue of* (4·1) *of finite multiplicity, $\{t_i\}$ satisfies conditions* (4·29), (4·34), *and* (4·40), *u_0 is of the form* (4·39), *and $\{u_i\}$ is determined by* (4·8) *to* (4·10), *then*

(*a*) *$\{\lambda^i\}$ converges to $\lambda^* = \lambda_1$;*

(*b*) *$\{u_i\}$ converges in direction to w in the H_0-metric with the error estimate*

$$|u_i - \xi_i w| \leqslant \left(\frac{\lambda_2(\lambda^i - \lambda_1)}{\lambda_2 - \lambda_1}\right)^{\frac{1}{2}} |u_i|_1 \quad (\xi_i \neq 0). \tag{4·47}$$

Proof. To prove theorem 4·2 (*a*) we use lemma 4·4 and (4·45) to get

$$(T u_{i+1}, K w) = \left[1 + t_i \left(\frac{\lambda^i}{\lambda_1} - 1\right)\right] (T \xi_i w, K w) = \left[1 + t_i \left(\frac{\lambda^i}{\lambda_1} - 1\right)\right] (T u_i, K w).$$

Using the above relation recursively we obtain the formula

$$(T w_{i+1}, K w) = \prod_{1=j}^{i} (1 + \tau_j)\,(T u_0, K w), \tag{4·48}$$

where $\tau_j = t_j \left(\frac{\lambda^j}{\lambda_1} - 1\right) \geqslant 0$ since $\lambda^i \geqslant \lambda_1$ for all i and $(T u_0, K w) > 0$. Formula (4·48) shows that $\{(T u_{i+1}, K w)\}$ is a monotonically increasing sequence of positive numbers. Since $(T u_{i+1}, K w) \leqslant |u_{i+1}|$, lemma 4·3 (*b*) implies that the sequence $\{(T u_{i+1}, K w)\}$ is bounded. Consequently, the product $\prod_{j=0}^{i} (1 + \tau_j)$ converges, as $i \to \infty$. But, for $\tau_j \geqslant 0$, the convergence of

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

$\prod_{j=0}^{\infty}(1+\tau_j)$ is equivalent to the convergence of the series $\sum_{j=0}^{\infty}\tau_j$. The convergence of the latter imples that $\tau_j \to 0$, as $j\to\infty$. Since $\lambda_j \geqslant \lambda_1$ and $t_i \geqslant \delta_1 > 0$ we get

$$\left(\frac{\lambda^j}{\lambda_1}-1\right) \leqslant \frac{1}{\delta_1}\tau_j \to 0 \quad \text{as} \quad j\to\infty,$$

showing that $\lambda^j \to \lambda_1$, as $j\to\infty$.

The proof of theorem 4·2 (*b*) follows from (4·44) and the fact that

$$\begin{aligned}(\lambda^i-\lambda_1)\,(Su_i, Ku_i) &= (Tu_i, Ku_i)-\lambda_1(Su_i, Ku_i)\\ &= (Tv_i, Kv_i)-\lambda_1(Sv_i, Kv_i).\end{aligned} \tag{4·49}$$

Since λ_1 is an isolated eigenvalue of (4·1) and, by definition, the next eigenvalue

$$\lambda_2 = \inf.\left\{\frac{(Tu, Ku)}{(Su, Ku)}:\quad u\in D_T \cap {}^{\perp}H_0^1\right\} \quad \text{and} \quad v_i \in D_T\cap {}^{\perp}H_0^1$$

we obtain from (4·49) the inequality $(\lambda^i-\lambda_1)\,(Su_i, Ku_i) \geqslant (\lambda_2-\lambda_1)\,(Sv_i, Kv_i)$; i.e.

$$|v_i|_1^2 \leqslant \frac{\lambda^i-\lambda_1}{\lambda_2-\lambda_1}|u_i|_1^2. \tag{4·50}$$

To obtain the estimate in the H_0-metric note that $|v_i|^2/\lambda_2 \geqslant |v_i|_1^2$, $-(\lambda_1/\lambda_2)\,|v_i|^2 \leqslant -\lambda_1|v_i|_1^2$, and $|u_i|_1^2 \leqslant (1/\lambda_1)\,|u_i|^2$. This and (4·49) show that

$$\frac{\lambda^i-\lambda_1}{\lambda_1}|u_i|^2 \geqslant (\lambda^i-\lambda_1)\,|u_i|_1^2 = (Tv_i, Kv_i)-\lambda_1(Sv_i, Kv_i) \geqslant \frac{\lambda_2-\lambda_1}{\lambda_2}|v_i|^2. \tag{4·51}$$

In virtue of (4·51) and relation (4·42), theorem 4·2 (*a*) shows that

$$\lim_{i\to\infty}\sin^2[u_i, w] = \lim_{i\to\infty}\frac{|v_i|^2}{|u_i|^2} \leqslant \lim_{i\to\infty}\left(\frac{\lambda_2}{\lambda_1}\frac{\lambda^i-\lambda_1}{\lambda_2-\lambda_1}\right) = 0;$$

i.e. $\{u_i\}$ converges in direction to w. The error estimate (4·47) follows also from (4·51). The proof of theorem 4·2 is thus complete.

We see from the proof of our theorem that if the space H_0^1 is not one-dimensional then the limit of the sequence $\{u_i\}$ depends on the choice of u_0.

Let us now consider an important case when $C = S$ with $D_S = D_C = D_T$. This will happen for example when we deal with finite matrices or bounded operators in H. Let us first summarize in Lemma 4·5 below the following properties which are valid in this case.

Lemma 4·5. *If $C = S$ and all conditions satisfied by C are now valid for S, then*

(*a*) $(Sh, Ku) = 0$, *where u and h are determined by* (4·8) *to* (4·10);

(*b*) $a = 1$, $b = 0$, $c = d$, $p(u,t) = \dfrac{2t-st^2}{1+ct^2}$;

(*c*) $\dfrac{dp}{dt} = \dfrac{2(1-st-dt^2)}{(1+dt^2)^2} = 0$ for $\tilde\theta = \dfrac{2}{s-\sqrt{\{s^2+4d\}}} < 0$ *and* $\theta = \dfrac{2}{s+\sqrt{\{s^2+4d\}}} > 0$;

(*d*) $\min_t p = p(u,\tilde\theta) = \tilde\theta < 0$ *and* $\max_t p = p(u,\theta) = \theta > 0$;

(*e*) $\lambda(h)-\lambda(u-\theta h) = \lambda(h)-\lambda(u)+\lambda(u)-\lambda(u-\theta h) = s+d\theta = 1/\theta$;

(*f*) *The graph of p is determined by figures* 1, 2*j*, *and* 3*j*;

(*g*) *Lemma* 4·1 *remains valid for t in the interval*: $0 < t < 2/(M_1-m_1)$.

THEOREM 4·3. *If $C = S$ and all other conditions of theorem 4·2 are satisfied, then*
(*a*) *$\{\lambda^i\}$ converges to $\lambda^* = \lambda_1$;*
(*b*) *$\{u_i\}$ converges in direction to w in the H_1- and H_0-metric and*

$$|u_i - \xi_i w|_1 \leqslant \left(\frac{\lambda^i - \lambda_1}{\lambda_2 - \lambda_1}\right)^{\frac{1}{2}} |u_i|_1. \tag{4·52}$$

Proof. Let us first note that $(Su_0, Kw) = (1/m)(Tu_0, Kw) > 0$. Furthermore, in virtue of lemma 4·4 and relation (4·46), we easily obtain the formula

$$(Su_{i+1}, Kw) = \prod_{j=0}^{i} (1 + \tilde{\tau}_j)(Su_0, Kw), \tag{4·53}$$

where $\tilde{\tau}_j = t_j(\lambda_j - \lambda_1) \geqslant 0$. This implies that $\{(Su_{i+1}, Kw)\}$ forms a monotonically increasing sequence of positive numbers which, in view of lemma 4·3 (*b*), is bounded. Repeating the argument in the proof of theorem 4·2 (*a*) we derive the validity of theorem 4·3 (*a*).

To prove theorem 4·3 (*b*) we use the same argument as in the proof of theorem 4·2 (*b*) to get the inequality

$$\sin^2 [u_i, w]_1 = \frac{|v_i|_1^2}{|u_i|_1^2} \leqslant \frac{\lambda^i - \lambda_1}{\lambda_2 - \lambda_1}$$

from which we conclude that $\{u_i\}$ converges in direction to w in the H_1-metric. Using the property of λ_1 and the inequality (4·3) for $C = S$ we easily show that

$$\sin^2 [u_i, w] = \frac{|v_i|^2}{|u_i|^2} \leqslant \frac{M_1}{\lambda_1} \frac{\lambda^i - \lambda_1}{\lambda_2 - \lambda_1} \to 0 \quad \text{as} \quad i \to \infty;$$

i.e. $\{u_i\}$ converges to w also in direction in the H_0-metric.

4·5. *Construction of sequences $\{t_i\}$ and special cases*

The general method described in the previous sections is not precise until the choice of the sequence $\{t_i\}$ and of the operators K and C has been made. In this section we shall describe a few methods of constructing the sequence $\{t_i\}$ so that the conditions (4·29) and (4·34) are satisfied. Then by specializing the operators K, C, and S we obtain, in particular, the most important iterative methods used in the approximate calculation of eigenvalues and eigenvectors of symmetric matrices and bounded operators as special cases. This will generalize at the same time these particular schemes to the solution of eigenvalue problems with unbounded operators. In what follows we shall always assume that conditions (4·3) and (4·36) are satisfied.

(i) *The method of a constant factor.* If $t_i = t$, t is a constant, then the scheme (4·8) reduces to the method of a constant factor

$$u_{i+1} = u_i - th_i \quad (i = 0, 1, 2, \ldots), \tag{4·8$_1$}$$

where h_i is the solution of (4·10) and u_0 is an initial approximation in D_T. We see from lemma 4·1 that for $\{\lambda^i\}$ to form a decreasing sequence whatever the initial vector u_0 it is necessary to limit t to the interval (4·24) or, if additional condition (4·26) is fulfilled, to the interval (4·27). Let us add that the sequence will not itself, in general, be independent of u_0. To show that such a t satisfies conditions (4·29) and (4·34) let us put t in the form $t = \alpha/M_1$, $0 < \alpha < 2$. Then, by (4·28), $a_i s_i t \leqslant \alpha$ or $2 - a_i s_i t \geqslant 2 - \alpha > 0$. Furthermore,

$$1 - 2b_i t + c_i t^2 = |u_i - th_i|_1^2/|u_i|_1^2 \leqslant \left(1 + t\frac{|h_i|_1}{|u_i|_1}\right)^2 = (1 + tc_i^{\frac{1}{2}})^2$$

and therefore,
$$p(u_i, t) = \frac{t(2 - a_i s_i \alpha/M_1)}{1 - 2b_i t + c_i t^2} \geqslant \frac{t(2-\alpha)}{(1 + tc_i^{\frac{1}{2}})^2}. \tag{4·54}$$

Thus, to verify the inequality (4·29) it is sufficient to estimate the bound of c_i independent of i. Using (4·36) and the fact that for $t = \alpha/M_1$, $0 < \alpha < 2$, $\{\lambda^i\}$ is a monotonically decreasing sequence we obtain

$$c_i = \frac{(Sh_i, Kh_i)}{(Su_i, Ku_i)} = \frac{(Sh_i, Kh_i)}{(Ch_i, Kh_i)}\,\frac{(Ch_i, Kh_i)}{(Tu_i, Ku_i)}\,\lambda^i \leqslant \frac{\lambda^0}{l_2}\,\frac{(Ch_i, Kh_i)}{(Tu_i, Ku_i)}. \tag{4·55}$$

On the other hand, in virtue of (4·3) and (4·36), the operator $T^{-1}S$ defined on D_T is bounded in H_0-norm. Hence, using (4·10) and (4·3), we get

$$|h_i|_2^2 = (Tu_i, Kh_i) - \lambda^i(Su_i, Kh_i) \leqslant (1 + \lambda^0 |T^{-1}S|)\,|u_i|\,|h_i|$$
$$\leqslant M_1^{\frac{1}{2}}(1 + \lambda^0 |T^{-1}S|)\,|u_i|\,|h_i|_2. \tag{4·56}$$

Consequently, by (4·55) and (4·56),

$$(1 + tc_i^{\frac{1}{2}})^2 \leqslant [1 + t(\lambda^0 M_1/l_2)^{\frac{1}{2}}(1 + \lambda^0 |T^{-1}S|)]^2 \equiv [1 + t\tilde{M}]^2 \equiv M, \tag{4·57}$$

where $\tilde{M}$ denotes the upper bound of $c_i^{\frac{1}{2}}$. This and (4·54) imply the validity of (4·29) with $\delta = t(2-\alpha)/M$. The inequality (4·34) is satisfied trivially. Let us add that the inequality (4·40) is, of course, also satisfied.

Thus, for the method of a constant factor t all the results derived in the previous sections remain valid.

Special case. If $K = S = C = I$ and H is a finite dimensional space, then T is a selfadjoint and positive definite matrix. In this case the scheme $(4·8_1)$ reduces to an analogous scheme for such eigenvalue problems [6, 37]. However, even in this simple case our results are more general. If $C = I$ and K is a symmetric and positive definite matrix (or bounded operator in H), then T and S are symmetrizable matrices (or operators) and the scheme $(4·8_1)$ is valid in this case [11]. Also, many other choices of the operators C and K are possible.

(ii) *The method with relative minimal norms.*† If

$$t_i = 1/e_i \equiv (Ch_i, Kh_i)/(Th_i, Kh_i) \equiv \tilde{e}_i,$$

then (4·8) reduces to the procedure

$$u_{i+1} = u_i - \tilde{e}_i h_i \quad (i = 0, 1, 2, \ldots), \tag{4·8_2}$$

with h_i determined by (4·10). This choice of t_i is suggested by the form of the function $p(u, t)$ in (4·15) since one of the basic requirements on t_i is that condition (4·29) be not only valid but also easily verifiable. As will be seen below this is the case when $t_i = \tilde{e}_i$. Let us also observe that when the operator $C \neq T$ and S has also the property that $(Su_i, Kh_i) = 0$ $(i = 0, 1, 2, \ldots)$ then the method $(4·8_2)$ has the following geometrical meaning: The choice

† In case $C = K = S = I$ and T is a bounded, selfadjoint, and positive definite operator in the real space H the method $(4·8_2)$ was suggested by Krasnoselsky [20] and investigated by Pugachev [29, 30] and Bessmertnykh [2]; see also Altman [1].

$t_i = \tilde{e}_i$ implies that the vector $t_i h_i$ is the orthogonal projection in the sense of the H_0-metric of the iterant u_i on h_i. Consequently, u_{i+1} is obtained from u_i by subtracting from it the projection of u_i on h_i so that the norm $|u_{i+1}|^2$ decreases at each step of this process in such a way that its magnitude is a minimum. For that reason we call it *the method with relative minimal norms*. To verify our conditions (4·29) and (4·34) note that by a simple calculation

$$p(u_i, \tilde{e}_i) = \frac{\tilde{e}_i(1+\{\lambda^i/\lambda(h_i)\}}{1-2b_i\tilde{e}_i+c_i\tilde{e}_i^2} \geqslant \frac{\tilde{e}_i}{1-b_i\tilde{e}_i+c_i\tilde{e}_i^2} \geqslant \frac{1}{M_1 M^0} \equiv \delta_2 > 0, \qquad (4{\cdot}29)$$

where we have used the fact that by (4·3)

$$0 < \frac{1}{M_1} \leqslant \tilde{e}_i \leqslant \frac{1}{M_2} \qquad (4{\cdot}34_2)$$

and that, in virtue of the discussion in (i),

$$(1-2b_i\tilde{e}_i+c_i\tilde{e}_i^2) \leqslant (1+\tilde{e}_i c_i^{\frac{1}{2}})^2 \leqslant \left[1+\frac{1}{M_2}.\tilde{M}\right] \equiv M^0.$$

Thus, we see that at the same time we have verified conditions (4·34) and (4·40) so that for this method all the results derived in §§ 4·3 and 4·4 are valid.

Special cases. If we choose $C = T$, then $t_i = \tilde{e}_i = 1$ for each i and the method $(4{\cdot}8_2)$ is essentially equivalent to the method

$$Tu_{i+1} = \lambda^i S u_i, \quad \lambda^i = \frac{(Tu_i, Ku_i)}{(Su_i, Ku_i)} \quad (i = 0, 1, 2, \ldots), \qquad (4{\cdot}8_2')$$

which could be regarded as a generalization of the Birger–Kolomy scheme [22] for the unbounded eigenvalue problem (4·1). In particular, if $K = I$, then T and S are Hermitian positive definite operators and the procedure $(4{\cdot}8_2')$ reduces to the method used in [22]. Obviously, by choosing K, T, and S properly we obtain all the schemes of the form $(4{\cdot}8_2')$ considered in [22]. Hence, if u_0 satisfies (4·39), theorem 4·2 gives the convergence of these procedures. The method $(4{\cdot}8_2')$ is also related to the Schwarz constants method [4]. Let us add that using theorem 2·7 with slightly stronger conditions on S one can derive for the method $(4{\cdot}8_2)$ with $C = T$ an estimate for $\lambda^i - \lambda_1$ which is analogous to the estimate derived by Temple [35] for a differential eigenvalue problem.

Another suitable and important special case of $(4{\cdot}8_2)$ is obtained when we choose $C = S$. In that case $D_S = D_C = D_T$, $(Su_i, Kh_i) = 0$ for each i, and, in view of (4·3) and (4·25), the choice $t_i = (Sh_i, Kh_i)/(Th_i Kh_i) = 1/\lambda(h_i)$ has the property that $0 < 1/M_1 \leqslant t_i \leqslant 1/m_1$. Thus, for this case theorem 4·3 is valid provided u_0 is of the form (4·39). At the end let us remark that the choice of K is still at our disposal.

(iii) *Acceleration of the new method.* We introduce a parameter $\alpha > 0$ in the process $(4{\cdot}8_2)$ so as to get the scheme

$$u_{i+1} = u_i - \alpha\tilde{e}_i h_i \quad (i = 0, 1, 2, \ldots). \qquad (4{\cdot}8_3)$$

Again a simple calculation and the result from (ii) show that for $\delta_2' \equiv [M_1(1+(\alpha/M_2)\,\tilde{M})^2]^{-1}$

$$p(u_i, \alpha\tilde{e}_i) = \frac{2\alpha\tilde{e}_i - \alpha^2\tilde{e}_i + \alpha^2\tilde{e}_i\lambda_i/\lambda(h_i)}{1-2b_i\alpha\tilde{e}_i+c_i\alpha^2\tilde{e}_i^2} > \frac{\tilde{e}_i\alpha(2-\alpha)}{1-2b_i\alpha\tilde{e}_i+c_i(\alpha\tilde{e}_i)^2}$$

$$\geqslant \delta_2'\alpha(2-\alpha). \qquad (4{\cdot}29_3)$$

Hence conditions (4·29) and (4·34) are satisfied for any α in the interval

$$0 < \alpha < 2. \tag{4·58}$$

However, if additionally we know that there exist constants $Q > q > 0$ so that

$$0 < q \leqslant \frac{(Tu, Ku)}{(Su, Ku)} \leqslant Q \quad (u \in D_T), \tag{4·59}$$

then the interval (4·58) can be somewhat extended. Indeed, from the equality ($4·29_3$) and the inequality (4·59) we obtain the inequality

$$p(u_i, \alpha\tilde{e}_i) = \frac{\alpha\tilde{e}_i[2-\alpha(1-\lambda^i/\lambda(h_i))]}{1-2b_i\alpha\tilde{e}_i+c_i\alpha^2\tilde{e}_i^2} \geqslant \frac{\tilde{e}_i\alpha[2-\alpha(1-q/Q)]}{1-2b_i\alpha\tilde{e}_i+c_i\alpha^2\tilde{e}^2}.$$

Combining this with the inequality ($4·29_2$) we derive

$$p(u_i, \alpha\tilde{e}_i) > \delta_2'\alpha[2-\alpha(1-q/Q)] > 0 \tag{$4·29_3^0$}$$

for α lying in the interval
$$0 < \alpha < 2(1-q/Q)^{-1}. \tag{4·60}$$

Special cases. (j) If $K = C = S = I$ and T is a bounded, selfadjoint, and positive definite operator satisfying condition (4·59), then our results yield in this case the corresponding results of Altman [1].

(jj) If in addition to the conditions in (j) we take $\alpha = 2$, which clearly belongs to the interval (4·60), then the procedure ($4·8_3$) reduces to the method of *normal chords* [19].

(iv) *The generalized method of steepest descent.*† If $t_i = \theta_i$, where θ_i is determined by (4·23), then in view of Remark 7 the choice $t_i = \theta_i$ has the property of minimizing the functional $\lambda(u_{i+1})$ so that the scheme (4·8) reduces in this case to the generalized method of steepest descent

$$u_{i+1} = u_i - \theta_i h_i \quad (i = 0, 1, 2, \ldots), \tag{$4·8_4$}$$

which is a well-known procedure in the case of finite symmetric and positive definite matrices [15, 3]. Let us observe that according to the graph of p the roots θ_i can be, in general, both positive and negative. Since $p_i \equiv p(u_i, \theta_i) > 0$ and, by (4·22)

$$1/p_i = 1/\theta_i - b_i \leqslant |1/\theta_i| + |b_i|$$

we see that to verify the validity of (4·29) it is sufficient to show that

$$|1/\theta_i| = \tfrac{1}{2}|a_i s_i + \surd\{(a_i s_i)^2 - 4(a_i b_i s_i - c_i)\}| \quad \text{and} \quad |b_i|$$

possess upper bounds which are independent of i.

To that end note that by definition of b_i and (4·57) we have $|b_i| \leqslant |h_i|_1/|u_i|_1 \leqslant \tilde{M}$ while by (4·28), (4·3), (4·36), and Remark 7 we have $|a_i s_i| \leqslant M_1 + \lambda^0/l_2$. Furthermore, this implies that

$$|a_i b_i s_i - c_i| \leqslant |b_i|\,|a_i s_i| + c_i \leqslant \frac{|h_i|_1}{|u_i|_1}\left(|a_i s_i| + \frac{|h_i|_1}{|u_i|_1}\right) \leqslant \tilde{M}\left(M_1 + \frac{\lambda^0}{l_2} + \tilde{M}\right).$$

Thus, the above inequalities imply that $|1/\theta_i|$ and $|b_i|$ have upper bounds independent of i. Consequently, there exists a constant $\delta_4 > 0$ so that

$$p(u_i, \theta_i) \geqslant \delta_4 > 0 \quad (i = 0, 1, 2, \ldots); \tag{$4·29_4$}$$

i.e. condition (4·29) is satisfied. Hence, lemma 4·2 is valid in this case.

† This method was investigated for the unbounded operators by Samokish [33] for the case when $K = S = I$ and T^{-1} and C^{-1} are compact selfadjoint operators.

REMARK 10. Let us observe that, under the present very general conditions on C and the arbitrariness of u_0, the investigation of the graph of $p(u,t)$ and especially of the Cases $3i$ and $3j$ show not only that θ_i can be positive and negative but also that they can be arbitrarily large since the denominator in the expression γ'' can be arbitrarily small. Consequently, without further conditions on either the choice of C or the initial element u_0 or on both we cannot expect the condition (4·34) to be satisfied. We shall return to this problem at some later time.

(v) *Accelerated method of steepest descent*. If $t_i = \alpha\theta_i$, then we obtain the accelerated method of steepest descent

$$u_{i+1} = u_i - \alpha\theta_i h_i \quad (i = 0, 1, 2, \ldots). \tag{4·8$_5$}$$

Using the fact that $b_i a_i s_i \theta_i^2 = s_i a_i \theta_i = c_i \theta_i - 1$, it is easy to see that

$$p(u_i, \alpha\theta_i) = \frac{\alpha\theta_i}{1-b_i\theta_i}\,\frac{(1-b_i\theta_i)\,(2-a_i s_i \alpha\theta_i)}{(1-2b_i\alpha\theta_i + c_i\alpha^2\theta_i^2)}$$

$$= \frac{\alpha\theta_i}{1-b_i\theta_i}[1+(1-\alpha)\,g(u_i,\theta_i,\alpha)], \tag{4·61}$$

where

$$g(u_i,\theta_i,\alpha) \equiv \frac{1-2b_i\theta_i + \alpha c_i s_i^2}{1-2b_i\alpha\theta_i+\alpha^2 c_i\theta_i^2} \quad (i = 0, 1, 2, \ldots). \tag{4·62}$$

Formulas (4·61) and (4·62) suggest that we take α in the range: $0 < \alpha < 1$. Furthermore, in order for g to be positive, α must also satisfy the condition

$$\frac{2b_i\theta_i - 1}{c_i^2\theta_i^2} < \alpha < 1. \tag{4·63}$$

Since $1 > (2b_i\theta_i - 1)/c_i\theta_i^2$ and $(2b_i\theta_i-1)/c_i\theta_i^2 > 0$ for $\theta_i < 0$ we must conclude from this and (4·63) that we cannot take α so that $0 < \alpha \ll 1$. Thus, we have the justification of the suggestion derived experimentally [37] that we take α so that $0 \ll \alpha < 1$. As a matter of fact, we should take α so that $0 \ll \alpha < 1$ if $1 - 2b_i\theta_i + \alpha c_i \theta_i^2 > 0$ and $\alpha = 1$, otherwise.

If these conditions are satisfied, then from (4·61) and (4·29$_4$) we get

$$p(u_i, \alpha\theta_i) \geqslant \alpha\delta_4 > 0 \quad (i = 0, 1, 2, \ldots).$$

Special case. An important special case is obtained when $C = S$, i.e. when $b = 0$. In this case $p(u_i, \alpha\theta_i) = \alpha\theta_i[1+(1-\alpha)\,(1+\alpha c_i\theta_i^2)/(1+\alpha^2 c_i\theta_i^2)] \geqslant \alpha\theta_i$ for α in $0 < \alpha < 1$. The form of p suggests that in this case we also take α to be near 1.

(vi) *A modified method of steepest descent*. We have seen in (iv) that without further conditions on C and u_0 the choice $t_i = \theta_i$ will not satisfy the inequality (4·34). However, the following modification of procedure (4·8$_4$)†, called here a *modified method of steepest descent*, will be shown to satisfy all our conditions.

Let us determine a sequence $\tilde{t}_i = t(u_i)$ as follows: Choose an arbitrary fixed constant $\delta_6 > 0$ and define t_i by

$$\tilde{t}_i = \begin{cases} \theta_i, & \text{if } 0 < \theta_i \leqslant \delta_6, \\ \delta_6, & \text{if no such root exists.} \end{cases} \tag{4·64}$$

† In case $K = I$ and other operators are finite symmetric and positive matrices the method was proposed by Hestenes & Karush [11].

This sequence is simple to construct since it involves only the solution of (4·19) and a comparison of numbers. It follows from our construction of $\tilde{t}_i$ and the discussion in (iv) that there exists a constant $\delta_5 > 0$ such that

$$\delta_5 \leqslant \tilde{t}_i \leqslant \delta_6 \quad (i = 0, 1, 2, \ldots).$$

Thus, conditions (4·34) and (4·40) are fulfilled. Furthermore, the graphs of $p(u,t)$ and (4·64) show that $p(u,t)$ is an increasing function on the interval $0 \leqslant t \leqslant t(u_i)$. Consequently, we can select a number $\tilde{\delta}_5 > 0$ so that $\tilde{\delta}_5 < \delta_5$ and $\tilde{\delta}_5 \in (0, 2M_1^{-1})$ and for which, in virtue of (4·28) and (4·57), we have

$$p(u_i, \tilde{t}_i) \geqslant p(u_i, \tilde{\delta}_5) = \frac{\tilde{\delta}_5(2 - a_i s_i \tilde{\delta}_5)}{1 - 2b_i \tilde{\delta}_5 + c_i(\tilde{\delta}_5^2)} \geqslant \frac{\tilde{\delta}_5(2 - \tilde{\delta}_5 M_1^{-1})}{(1 + \tilde{\delta}_5 \tilde{M})^2} > 0.$$

Thus, we see that in this case conditions (4·29) and (4·34) and even (4·40) are satisfied.

We shall end this chapter by observing that the form (4·15) of $p(u,t)$ suggests another choice of t_i, namely, $t_i = 1/|a_i s_i|$, so that (4·8) reduces in this case to

$$u_{i+1} = u_i - h_i/|a_i s_i| \quad (i = 0, 1, 2, \ldots). \tag{4·8$_6$}$$

It is easy to see that when $|h_i|_1/|u_i|_1$ is small, as will be the case when u_i is near the solution, then the choice $t_i = 1/|a_i s_i|$ is a good estimate of the optimum value of θ_i. It can be shown that when we choose the operators K and C appropriately and assume that u_0 is so chosen that for some i we have $\lambda^i \leqslant \lambda_2$, then the sequence $\{1/|a_i s_i|\}$ satisfies conditions (4·29), (4·34), and (4·40). The discussion of the last statement as well as the investigation of the closely connected problem indicated in Remark 10 will be taken up at some later time.

4·6. *Applications to ordinary differential eigenvalue problems*

In this section we illustrate the applicability and the numerical effectiveness of the iterative method (4·8) to (4·10) by calculating the smallest eigenvalue λ_1 by means of *the method with relative minimal norms* (4·8$_2$) and its *acceleration* (4·8$_3$) for two ordinary differential eigenvalue problems arising in the problems of elastic stability.

Observe first that, by lemma 4·2 (a) and theorem 4·1, if the initial approximation u_0 in $D(T)$ is so chosen that $\lambda^i \leqslant \lambda_2$ for some i $(i = 0, 1, 2, \ldots)$, then $\lambda^i \rightarrow \lambda_1$ monotonically. Thus, to calculate λ_1 approximately we need to know a lower estimate k_2 for λ_2. The latter will be obtained from the following *comparison theorem* assuming, of course, that the eigenvalues of the comparison problem are easily obtainable.

THEOREM 4·4 (a). *Consider a pair of eigenvalue problems*

$$Tu - \lambda Su = 0, \quad Tu - \lambda^* S^* u = 0, \tag{4·65}$$

where T is K-p.d. and the operators S and S^* satisfy the inequality

$$(S^*u, Ku) \geqslant (Su, Ku) > 0 \quad \text{for all } u \text{ in } D(T). \tag{4·66}$$

Then the eigenvalues λ_i and λ_i^* of the respective problems (4·65) *are such that $\lambda_i^* \leqslant \lambda_i$ for $i = 1, 2, 3, \ldots$.*

(b) *Suppose that T and T^+ are K-p.d. with $D(T) = D(T^+)$,*

$$(T^+u, Ku) \leqslant (Tu, Ku) \quad \text{for all } u \text{ in } D(T), \tag{4·67}$$

and S is K-positive on $D(T)$. Then the eigenvalues λ_i and λ_i^+ of the corresponding problems

$$Tu - \lambda Su = 0, \quad T^+u - \lambda Su = 0 \tag{4·68}$$

are such that $\lambda_i^+ \leqslant \lambda_i$ $(i = 1, 2, 3, \ldots)$.

Proof. The proof of theorem 4·4 follows immediately from Courant's [5] maximum-minimum principle which, in virtue of theorem 2·2 and 2·3, can be proved in exactly the same way as it was done in [4].

(A) *Selfadjoint eigenvalue problem.* As a first example from the theory of elastic stability we consider the question of determining the critical load in the buckling problem for a column, with variable moment of inertia, which is simply supported at the top and clamped at the bottom. After a simple transformation of axis this question (see [42, 43]) leads to the problem of determining the smallest eigenvalue λ_1 of the 4th order ordinary differential eigenvalue problem:

$$(D(x)\,y'')'' + \lambda y'' = 0, \quad y(0) = y(1) = y'(0) = y''(1) = 0, \tag{4·69}$$

where we assume that $D(x) \in C^2(0, 1)$, $0 < m \leqslant D(x) \leqslant M$ on $[0, 1]$ and $m \leqslant 1$.

Let $L_2(0·1)$ be the real Hilbert space of square-summable functions $u(x)$ defined on $[0, 1]$ with the inner product and norm given by

$$(u, v) = \int_0^1 u(x)\, v(x)\, \mathrm{d}x, \quad \|u\|^2 = \int_0^1 u^2\, \mathrm{d}x \quad (u, v \in L_2(0, 1)).$$

To apply our iterative method with relative minimal norms put

$$Ty = (D(x)\,y'')'', \quad Cy = y^{(4)}, \quad Sy = -y'' \quad \text{and} \quad Ky = y \quad \text{for} \quad y \in D(T) \tag{4·70}$$

where $D(T) = D(C)$ is the set of all functions $y(x)$ in $C^4(0, 1)$ satisfying the boundary conditions (4·69). It is easy to verify that T, C, and S thus defined are symmetric. Furthermore, since for each $j = 0, 1, 2, \ldots$

$$\|u^{(j+1)}\|^2 \geqslant 2\|u^{(j)}\|^2 \quad \text{for every } u \text{ in } C^{j+1}(0, 1) \text{ with } u^{(j)}(0) = 0 \tag{4·71}$$

and $m \leqslant D(x) \leqslant M$, it follows that T, C and S are also positive definite and that the inequalities (4·3) and (4·36) hold with $M_2 = m$, $M_1 = M$ and $l_2 = 2$. Thus, theorem 4·1 is applicable to the eigenvalue problem (4·69) with $t_i = \tilde{e}_i$. For the application of the comparison theorem 4·4 consider the eigenvalue problem

$$Cy - \lambda^+ Sy = 0, \quad \text{i.e.} \quad \{y^{(4)} + \lambda^+ y'' = 0,\ y(0) = y(1) = y'(0) = y''(1) = 0\}. \tag{4·72}$$

Since, as is easily seen, for $\lambda^+ \neq 0$ with $\lambda^+ = k^2$ the eigenvalues λ^+ of (4·72) are determined from the equation $\tan k = k$, it follows [4] that $\lambda_1^+ \cong 20·19644$ and $\lambda_2^+ \cong 59·68026$. In virtue of the inequality (4·3) satisfied by T and C, theorem 4·4 (*b*) implies that

$$k_2 = m\lambda_2^+ \cong m59·68026 \leqslant \lambda_2.$$

According to the *iterative method with relative minimal norms* the basic iteration formula for the approximate calculation of the eigenvalues of (4·69) is given by

$$\lambda^i = \frac{(Ty_i, y_i)}{(Sy_i, y_i)} \quad (i = 1, 2, 3, \ldots), \tag{4·73}$$

where y_0 is a given initial approximation in $D(T)$ and

$$y_i = y_{i-1} - t_{i-1} h_{i-1} \quad (i = 1, 2, 3, \ldots), \tag{4·74}$$

where, for each i, h_{i-1} is the solution of the boundary-value problem

$$Ch_{i-1} = Ty_{i-1} - \lambda^{i-1} Sy_{i-1}, \quad \text{i.e.} \begin{cases} h_{i-1}^{(4)} = (D(x)\, y''_{i-1})'' + \lambda^{i-1} y''_{i-1} \\ y_{i-1}(0) = y_{i-1}(1) = y'_{i-1}(0) = y''_{i-1}(1) = 0 \end{cases} \tag{4·75}$$

$$t_{i-1} = \frac{(Ch_{i-1}, h_{i-1})}{(Th_{i-1}, h_{i-1})} \quad (i = 1, 2, 3, \ldots) \tag{4·76}$$

If y_0 is so chosen that $\lambda^0 \leqslant k_2$, then by theorem 4·1 the sequence $\{\lambda^i\}$ constructed by (4·73) to (4·76) converges monotonically to λ_1.

Particular example (i). For the numerical calculations, as a special case of (4·69), consider the eigenvalue problem

$$((1 - 0{\cdot}5x^2)\, y'')'' + \lambda y'' = 0, \quad y(0) = y(1) = y'(0) = y''(1) = 0. \tag{4·77}$$

In this case $D(x) = 1 - 0{\cdot}5x^2$, $Ty = ((1 - 0{\cdot}5x^2)\, y'')''$, $m = 0{\cdot}5$, $M = 1$ and $k_2 = 29{\cdot}84012 \leqslant \lambda_2$. If $y_0(x) = 3x^2 - 5x^3 + 2x^4$, then $y_0 \in D(T)$ and $\lambda^0 = 18{\cdot}25 < k_2$. Hence the sequence $\{\lambda^i\}$ constructed for (4·77) by (4·73) to (4·76) converges monotonically to λ_1. The first six approximations to λ_1 were computed on Maniac III and, when rounded to six decimal places, are given by

$$\lambda^1 = 16{\cdot}727846, \quad \lambda^2 = 16{\cdot}558719, \quad \lambda^3 = 16{\cdot}535116,$$

$$\lambda^4 = 16{\cdot}532073, \quad \lambda^5 = 16{\cdot}531726 \quad \text{and} \quad \lambda^6 = 16{\cdot}531691.$$

To apply the accelerated method with relative minimal norms $(4{\cdot}8_3)$ to (4·77) we first note that (4·29) and (4·34) are satisfied for any parameter α in

$$0 < \alpha < 2. \tag{4·78}$$

The approximations λ^i to λ_1 are also computed by the method (4·73) to (4·76) except that, in accordance with $(4{\cdot}8_3)$, y_i in (4·74) are given by

$$y_i = y_{i-1} - \alpha t_{i-1} h_{i-1} \quad (i = 1, 2, 3, \ldots). \tag{4·79}$$

The calculations were carried out for $\alpha = 0{\cdot}50$, $\alpha = 0{\cdot}75$, $\alpha = 1{\cdot}00$, $\alpha = 1{\cdot}25$, $\alpha = 1{\cdot}50$ and $\alpha = 1{\cdot}75$ with $\alpha = 1{\cdot}00$ giving, of course, the original method (4·74). The obtained results are shown in table 1.

Table 1

α	λ^1	λ^2	λ^3	λ^4	λ^6	λ^7
0·50	17·234877	16·838594	16·670429	16·595594	16·561380	16·545512
0·75	16·921443	16·635284	16·560261	16·539510	16·533778	16·532210
1·00	16·727846	16·558719	16·535116	16·532073	16·531726	16·531691
1·25	16·645627	16·540947	16·532651	16·531793	16·531699	16·531688
1·50	16·665706	16·553612	16535253	16·532335	16·531810	16·531712
1·75	16·778675	16·609669	16·560058	16·543194	16·536655	16·533924

The above table shows that for the problem (4·77) the best approximations λ^i to λ_1 are obtained for the accelerated parameter $\alpha = 1{\cdot}25$.

(B) *Non-selfadjoint eigenvalue problem.* As another application we consider the question of determining the critical load in the stability problem of a compressed bridge belt and in the problem of buckling of a bar under distributed axial load. As is known [43] in both of the above problems, our question leads to the problem of determining the smallest eigenvalue of the problem of the form

$$-y'''-\lambda p(x)\,y'=0, \quad y(0)=y'(0)=y''(1)=0, \tag{4·80}$$

where we assume that $p(x)$ is a polynomial in x such that $0\leqslant p(x)\leqslant 1$ on $[0,1]$. In this case we define the operators T, S, K and C by

$$Ty=Cy=-y''', \quad Sy=p(x)\,y' \quad \text{and} \quad Ky=y' \tag{4·81}$$

with T and C having the domain $D(T)=D(C)$ comprising the set of all y in $C^3(0,1)$ satisfying the boundary conditions (4·80) while the operators S and K are defined, for example, for all y in $C^1(0,1)$ satisfying the condition $y(0)=0$.

Using the same arguments as in [40] it can be shown that T (and C) is K-symmetric and K-p.d., i.e. K is closeable, $KD(T)$ is dense in $L_2(0,1)$, and

$$(Tu,Kv)=-\int_0^1 u'''v'\,\mathrm{d}x=-\int_0^1 u'v'''\,\mathrm{d}x=(Ku,Tv) \quad (u,v\in D(T)), \tag{4·82}$$

$$(Tu,Ku)=\int_0^1 (u'')^2\,\mathrm{d}x\geqslant 2\,\|u'\|^2=2\|Ku\|^2, \quad (Tu,Ku)\geqslant 4\|u\|^2 \quad (u\in D(T)). \tag{4·83}$$

Furthermore, S is clearly K-symmetric and $(Sy,Ky)>0$ for $y\neq 0$ in $D(S)$ for, if

$$(Sy,Ky)=\int_0^1 p(x)\,(y')^2\,\mathrm{d}x=0,$$

then $y'(x)\equiv 0$ on $[0,1]$, $y(x)=\text{const.}$ and $y(x)\equiv 0$ since $y(0)=0$. Note that (4·3) holds with $M_1=M_2=1$ while, in virtue of (4·71), (4·83) and the equality $p_M=\max\limits_{0\leqslant x\leqslant 1} p(x)$, (4·36) holds with $l_2=2/p_M$. Thus, theorem 4·1 or rather (since $T=C$) theorem 4·2 is applicable to the eigenvalue problem (4·80) with $t_i=\tilde{e}_i=1$ for each i.

To get the lower estimate k_2 for λ_2 of (4·80) we use the comparison problem

$$-y'''-\lambda^*y'=0, \quad y(0)=y'(0)=y''(1)=0, \tag{4·84}$$

whose eigenvalues are known to be $\lambda_i^*=(2i-1)^2\,4^{-1}\pi^2$ for $i=1,2,3,\ldots$. Because $S^*y=y'$ and $0\leqslant p(x)\leqslant 1$, $(Sy,Ky)\leqslant(S^*y,Ky)$ for y in $D(T)$. Hence, theorem 4·4 (*a*) implies that $\lambda_i^*\leqslant\lambda_i$, where λ_i are the eigenvalues of (4·80). It follows therefore that

$$k_2=22{\cdot}20661\simeq 9{\cdot}4^{-1}\pi^2=\lambda_2^*\leqslant\lambda_2.$$

The method with relative minimal norms (4·8$_2$) when applied to (4·80) with $C=T$ reduces to the method

$$\lambda^i=\frac{(Ty_i,Ky_i)}{(Sy_i,Ky_i)} \quad (i=1,2,\ldots), \tag{4·85}$$

where y_0 is a given function in $D(T)$ and, since $t_i=1$ for each i,

$$y_i=y_{i-1}-h_{i-1} \quad (i=1,2,3,\ldots), \tag{4·86}$$

where, for each i, h_{i-1} is the solution of the boundary-value problem

$$Th_{i-1}=Ty_{i-1}-\lambda^{i-1}Sy_{i-1}, \quad \text{i.e.} \left.\begin{cases} h'''_{i-1}=y'''_{i-1}+\lambda^{i-1}p(x)\,y'_{i-1} \\ h_{i-1}(0)=h'_{i-1}(0)=h''_{i-1}(1)=0 \end{cases}\right\} \quad (i=1,\ldots). \tag{4·87}$$

If y_0 is so chosen that $\lambda^0 \leqslant k_2$, then by theorem 4·2 the sequences $\{\lambda^i\}$ and $\{y_i\}$ constructed by (4·85) to (4·87) are such that $\{\lambda^i\}$ converges monotonically to λ_1 and $\{y_i\}$ converges in the H_0-metric to an eigenvector of (4·80) belonging to λ_1. Here H_0 is the completion of $D(T)$ in the metric

$$[u, v] = (Tu, Kv) = \int_0^1 u''v''\,dx, \quad |u|^2 = \int_0^1 (u'')^2\,dx. \tag{4·88}$$

Particular example (ii). For the numerical calculation, as an example of (4·80), consider the problem of determining approximately the critical load in the stability problem of a compressed bridge belt. This reduces to finding the smallest eigenvalue λ_1 of

$$-y''' - \lambda(1-x^2)\,y' = 0 \quad y(0) = y'(0) = y''(1) = 0. \tag{4·89}$$

In this case $p(x) = 1-x^2$ and $Sy = (1-x^2)\,y'$ with $p(x)$ and S clearly satisfying our condition on $[0, 1]$. If $y_0 = -6x^2+2x^3$, then $y_0 \in D(T)$ and $\lambda^0 = 5{\cdot}185185 < k_2$. Hence, our assertions of the preceding paragraph apply. The first six approximations to λ_1, when rounded to six decimal places, are given by

$$\lambda^1 = 5{\cdot}121985, \quad \lambda^2 = 5{\cdot}121673, \quad \lambda^3 = 5{\cdot}121669,$$

$$\lambda^4 = 5{\cdot}121669, \quad \lambda^5 = 121669 \quad \text{and} \quad \lambda^6 = 5{\cdot}121669.$$

To apply the accelerated method $(4{\cdot}8_3)$ to (4·89) note that by (4·58) the parameter α must lie in the open interval $(0, 2)$. As before we compute λ^i by (4·85) to (4·87) except that y_i in (4·86) are given by $y_i = y_{i-1} - \alpha h_{i-1}$. The obtained results for various values of α are shown in table 2.

TABLE 2

α	λ^1	λ^2	λ^3	λ^4	λ^5	λ^6
0·50	5·139284	5·126621	5·123073	5·122070	5·121784	5·121703
0·75	5·127121	5·122167	5·121717	5·121674	5·121670	5·121669
1·00	5·121985	5·121673	5·121669	5·121669	5·121669	5·121669
1·25	5·123752	5·121756	5·121673	5·121670	5·121669	5·121669
1·50	5·132295	5·123643	5·122055	5·121747	5·121685	5·121673
1·75	5·147486	5·132911	5·126771	5·124049	5·122802	5·122216

The above numerical results show that with $C = T$ the original method $(4{\cdot}8_2)$ seems to be best.

Finally let us observe that the numerical results obtained for (4·77) and (4·80) by the method with relative minimal norms and its acceleration compare very favourably with similar numerical results obtained by the method of Schwarz constants, variational method of Ritz, Galerkin method, finite difference method, and others (see, for example, [43]). The additional advantage of our method is that it converges monotonically and that at each step of the process we are solving a boundary-value problem of the form $Cy = f$, where f is given and C is a simple operator, say, with constant coefficients in case of differential eigenvalue problems. Furthermore, in many cases the programming of the problem on an electronic computer is quite simple.

The author is greatly indebted to E. Isaacson and P. D. Lax for their interest and suggestions offered in the course of preparation of this paper. Thanks are also due to C. De Prima for numerous stimulating discussions concerning some topics of this paper.

The author expresses his thanks to Mr W. W. Mixon from the Institute for Computer Research at the University of Chicago for his help in the derivation of the numerical results presented in the final section.

The author is grateful to the referee for his suggestion of a number of improvements and for his help in eliminating errata.

References

[1] Altman, M. 1961 An iterative method for the eigenproblem of linear bounded operators and matrices. *Bull. Polska Akad. Nauk.* **9**, 775–778.

[2] Bessmertnykh, G. A. 1961 On the behaviour of corrections in certain methods for the approximate calculation of the bounds of spectrum of a selfadjoint operator. *Dokl. Akad. Nauk. SSSR* **136**, 1265–1268.

[3] Birman, M. Sh. 1952 On calculation of eigenvalues by the method of steepest descent. *Zapisky, Leningrad Univ.* **27**, 209–216.

[4] Collatz, L. 1949 *Eigenwertaufgaben mit technischen Anwendungen.* Leipzig: Akademische Verlagsgesellschaft.

[5] Courant, R. & Hilbert, D. 1938 *Methoden der Mathematischen Physik.* band 2. Berlin: Springer-Verlag.

[6] Faddeev, D. K. & Faddeeva, V. M. 1960 Computational methods of linear algebra. *Gos. Izd. Fiz.-Mat. Lit.* Moskva.

[7] Friedrichs, K. O. 1934 Spektraltheorie hallbeschränkter Operatoren. *Math. Annln.* **109**, 465–487, 685–713.

[8] Friedrichs, K. O. 1947 On the boundary-value problem of the theory of elasticity and Korn's inequality. *Ann. Math.* **48**, 441–471.

[9] Gokhberg, I. C. & Krein, M. G. 1957 Fundamental aspects of defect numbers, root numbers and indices of linear operators. *Usp. Mat. Nauk.* **12**, no. 2, 43–118.

[10] Gould, S. H. 1957 *Variational methods for eigenvalue problems.* University of Toronto Press.

[11] Hestenes, M. R. & Karush, W. 1951 Solutions of $Ax = \lambda Bx$. *J. Res. Nat. Bur. Stand.* **47**, 471–478.

[12] Heuser, H. 1960 Zur Eigenwerttheorie einer Klasse symmetrischer Operatoren. *Math. Z.* **74**, 167–185.

[13] Huzino, S. 1958 On calculating eigenvalues by the gradient method. *Mem. Fac. Sci., Kyusyn Univ.* ser. A **12**, 30–39.

[14] Kamke, E. 1939, 1940 Über die definiten selbstadjugierten Eigenwertaufgaben bei gewähnlichen linearen Differentialgleichungen. I: *Math. Z.* **45** (1939), 759–787; II. *Math. Z.* **46** (1940), 231–250; III: *Math. Z.* **46** (1940), 251–286.

[15] Kantorovich, L. V. 1948 Functional analysis and applied mathematics. *Usp. Mat. Nauk.* **3**, no. 6, 89–185.

[16] Kantorovich, L. V. 1956 Approximate solution of functional equations. *Usp. Mat. Nauk.* **11**, no. 6, 99–116.

[17] Karush, W. 1951 Determination of the extreme values of the spectrum of a bounded self-adjoint operator. *Proc. Am. Math. Soc.* **2**, 980–989.

[18] Kharazov, D. F. 1957 On certain properties of linear operators insuring the validity of Hilbert–Schmidt theory. *Usp. Mat. Nauk.* **12**, no. 4, 201–207.

[19] Kostarchuk, V. N. 1957 The method of normal chords for the solution of operator equations. *Trudy, Sem. on Functional Anal., Voronezh.* no. 3, 54–76 (see also *Dokl. Akad. Nauk. SSSR* **98**, 531–534.)

[20] Krasnoselsky, M. A. 1956 On some procedures of approximate calculation of eigenvalues and eigenvectors of a positive definite matrix. *Usp. Mat. Nauk.* **11**, no. 3, 151–158.

[21] Lax, P. D. & Milgram, A. N. 1954 Parabolic equations. *Ann. Math. Stud.* **33**, 167–190.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

[22] Marek, I. 1960 On iterations of bounded linear operators and Kellogg's iterations in not self-adjoint eigenvalue problems. *Commentationes Math. Univ. Carolinae* **1**, 12–21.
[23] Martyniuk, A. E. 1959 Certain new applications of the methods of Galerkin type. *Mat. Sb.* **49**, 85–108.
[24] Martyniuk, A. E. 1959 Variational methods in boundary-value problems for weakly elliptic equations. *Dokl. Akad. Nauk. SSSR* **129**, 1222–1225.
[25] Mikhlin, S. G. 1952 *Minimum problem for quadratic functionals.* Moscow-Leningrad: Gostekhizdat.
[26] Petryshyn, W. V. 1962 Direct and iterative methods for the solution of linear operator equations in Hilbert space. *Trans. Am. Math. Soc.* **105**, 136–175.
[27] Polsky, N. I. 1955 On the convergence of certain approximate methods of analysis. *Ukrain. Mat. Zh.* **7**, 56–70.
[28] Polsky, N. I. 1962 Projection methods in applied mathematics. *Dokl. Akad. Nauk. SSSR* **143**, 787–790.
[29] Pugachev, B. P. 1956 On two procedures of approximate calculation of eigenvalues and eigenvectors. *Dokl. Akad. Nauk. SSSR*, **110**, 334–337.
[30] Pugachev, B. P. 1956 Investigation of a method of approximate calculation of eigenvalues and eigenvectors. *Trudy, Sem. Functional Anal. Voronezh*, no. 4.
[31] Rellich, F. 1953 Perturbation theory of eigenvalue problems. *Lecture Notes.* New York University.
[32] Riesz, F. & Sz.-Nagy, B. 1955 *Functional analysis.* New York: F. Ungar Publ. Co.
[33] Samokish, B. A. 1958 The method of steepest descent in the problem of eigenelements of semi-bounded operators. *Izv. Vysshikh Ushebnykh Zavedenii: Matematika*, no. 5(6), 105–114.
[34] Sobolevsky, P. E. 1957 On equations with operators forming an acute angle. *Dokl. Akad. Nauk. SSSR* **116**, 754–757.
[35] Temple, G. 1929 The computation of characteristic numbers and characteristic functions. *Proc. Lond. Math. Soc.* (2), **29**, 257–280.
[36] Wielandt, H. 1951 Über die Eigenwertaufgaben mit reellen diskreten Eigenwerten. *Math. Nachr.* **4**, 308–314.
[37] White, P. A. 1958 The computation of eigenvalues and eigenvectors of a matrix. *J. Soc. Industr. Appl. Math.* **6**, 393–437.
[38] Zaanen, C. 1953 *Linear analysis.* Amsterdam: North Holland.

Additional references

[39] Petryshyn, W. V. 1965 On a class of K-p.d. and non-K-p.d. operators and operator equations. *J. Math. Anal. Appl.* **10**, 1–24.
[40] Petryshyn, W. V. 1965 On two variants of a method for the solution of linear equations with unbounded operators and their applications. *J. Math. Phys.* **44**, 297–312.
[41] Shalov, V. M. 1965 The minimum principle for quadratic functional for hyperbolic equations. *Differential Equations*, **1**, 1338–1365.
[42] Timoshenko, S. P. & Gere, J. M. 1961 *Theory of elastic stability.* New York: McGraw-Hill.
[43] Mikeladze, Sh. Ye. 1951 *New Methods of integration of differential equations and their application to problems in the theory of elasticity.* Moscow-Leningrad: Gostekhizdat.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES