Chapter Two

The Rayleigh–Ritz Method

1. Poincaré's Inequalities: The Theoretical Foundation of the Rayleigh–Ritz Method

We now develop the inequalities of Poincaré, which in our terminology relate the eigenvalues of an operator of class $\mathscr{S}$ to the eigenvalues of its projection on a finite-dimensional subspace (see Definition A.19). Incidentally, this presentation will provide a simplification of the older proofs.

Let $\mathfrak{V}_n$ (or, for short, $\mathfrak{V}$) be any n-dimensional supspace of $\mathfrak{D}$ and let V_n (or V) be the orthogonal projection operator onto $\mathfrak{V}$. Consider the equation

$$VAVu = \Lambda u. \tag{1}$$

By construction, VAV is hermitian (or symmetric) and has all of $\mathfrak{H}$ for its domain. Therefore, by Definitions A.10 and A.11, VAV is selfadjoint. Moreover, VAV is also a transformation from the finite-dimensional space $\mathfrak{V}$ into itself, so that the spectrum of VAV, considered as an operator on $\mathfrak{V}$ consists of n real eigenvalues, $\Lambda_1 \leq \Lambda_2 \leq \cdots \leq \Lambda_n$, some of which may be zero. Denote by $w_1, w_2, \ldots, w_n$ the corresponding orthonormal set of eigenvectors in $\mathfrak{V}$. The eigenvalues $\{\Lambda_i\}$ and eigenvectors $\{w_i\}$ will be called *nontrivial*. If, however, we consider VAV as an operator on $\mathfrak{H}$, its spectrum consists of the eigenvalues $\Lambda_1, \Lambda_2, \ldots, \Lambda_n$ as well as the eigenvalue $\lambda = 0$. If $\mathfrak{H}$ is infinite-dimensional, zero is an eigenvalue of infinite multiplicity.

Theorem 1. *(Poincaré) For any n-dimensional space $\mathfrak{V}_n$, the eigenvalues $\{\Lambda_i\}$ of VAV satisfy the inequalities*

$$\lambda_1 \leq \Lambda_1, \quad \lambda_2 \leq \Lambda_2, \ldots, \quad \lambda_n \leq \Lambda_n. \tag{2}$$

PROOF. For a given index j ($1 \leq j \leq n$), we consider a *test function*

$$w_0 = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_j w_j \neq 0,$$

where $\alpha_1, \alpha_2, \ldots, \alpha_j$ are adjusted so that the $j - 1$ orthogonality conditions $(w_0, u_i) = 0$ $(i = 1, 2, \ldots, j - 1)$ are satisfied. Without loss of generality, we can assume that $\|w_0\| = 1$.

It follows from Eq. (1.3.2)† that

$$\lambda_j = \min_{\substack{u \in \mathfrak{D} \\ (u, u_i) = 0 \\ (i = 1, 2, \ldots, j-1)}} R(u) \leq R(w_0).$$

On the other hand, since $Vw_0 = w_0$ and $(w_0, w_0) = 1$, we have

$$R(w_0) = \frac{(Aw_0, w_0)}{(w_0, w_0)} = (VAVw_0, w_0) = \sum_{i=1}^{j} |\alpha_i|^2 \Lambda_i \leq \Lambda_j.$$

The original proof of Theorem 1 by Poincaré [P4], repeated by Pólya and Schiffer [PS1], is somewhat more complicated than ours due to the fact that Poincaré began with an arbitrary basis $v_1, v_2, \ldots, v_n$ for $\mathfrak{V}_n$.

The inequalities of Poincaré are of theoretical importance in that they lead to an independent characterization of eigenvalues, *the minimum–maximum principle*.

These inequalities (2) also have an important practical application, usually called the *Rayleigh–Ritz method*, which we shall discuss in detail in Section 3.

2. The Minimum–Maximum Principle

While the classical characterization is very important as a theoretical tool, it has the disadvantage that it cannot be used to determine higher eigenvalues without using explicitly all preceding eigenvectors. The following approach overcomes this difficulty.

Let us focus our attention on the inequality

$$\lambda_n \leq \Lambda_n \tag{1}$$

from the set of inequalities (1.2). Since each $v \in \mathfrak{V}_n$ can be written as a linear combination

$$v = \beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_n w_n$$

† That is, Eq. (2) of Chapter 1, Section 3. This form of notation is used throughout the book.

it follows that

$$(Av, v) = \sum_{i=1}^{n} |\beta_i|^2 \Lambda_i \leq \Lambda_n (v, v) \tag{2}$$

for any choice of v. Moreover, the particular choice $v = w_n$ yields the equality $R(v) = \Lambda_n$. Combining this equality with inequality (1) we see that

$$\lambda_n \leq \Lambda_n = \max_{u \in \mathfrak{V}_n} R(u).$$

We can now give a characterization of λ_n which is independent in the sense that it does not explicitly use the eigenvectors to characterize the eigenvalues.

Theorem 1. *(Fischer–Pólya) The minimum–maximum principle. Let $\mathfrak{V}_n$ denote any n-dimensional subspace of $\mathfrak{D}$ $(n = 1, 2, \ldots)$. Then, the eigenvalues of A are characterized by the equation*

$$\lambda_n = \min_{\mathfrak{V}_n} \max_{u \in \mathfrak{V}_n} R(u). \tag{3}$$

PROOF. From the above discussion, we have

$$\lambda_n \leq \max_{u \in \mathfrak{V}_n} R(u) \tag{4}$$

for any $\mathfrak{V}_n \subset \mathfrak{D}$, $\dim \mathfrak{V}_n = n$. Let us now consider the subspace $\mathfrak{U}_n = \operatorname{sp}\{u_1, u_2, \ldots, u_n\}$, where sp { } denotes the subspace consisting of all linear combinations of the vectors enclosed. In this case, we have $R(v) \leq \lambda_n$, for any $v \in \mathfrak{U}_n$, and $R(u_n) = \lambda_n$. Therefore,

$$\lambda_n = \max_{u \in \mathfrak{U}_n} R(u). \tag{5}$$

Using inequality (4) and equality (5), we obtain the characterization (3).

We can take into account the possibility that the multiplicity of a given λ_n is greater than 1 in the following more explicit way.

Theorem 2. *For a fixed n, let $m(n)$ (for short, m) be defined by*

$$m = \min\{j \mid \lambda_j = \lambda_n\} \tag{6}$$

and let $\mathscr{G}_n$ be the set of all subspaces $\mathfrak{V} \subset \mathfrak{D}$ such that $m \leq \dim \mathfrak{V} < \infty$. Then, the eigenvalues of A are given by the equation

$$\lambda_n = \min_{\mathfrak{V} \in \mathscr{G}_n} \max_{u \in \mathfrak{V}} R(u). \tag{7}$$

PROOF. For any $\mathfrak{V} \in \mathscr{G}_n$, let $r = \dim \mathfrak{V}$. Since $r \geq m$, we have the inequality

$$\lambda_n = \lambda_m \leq \Lambda_m \leq \Lambda_r \leq \max_{u \in \mathfrak{V}} R(u).$$

It is clear that $\mathfrak{U}_n \in \mathscr{G}_n$, and thus the equality

$$\lambda_n = \max_{u \in \mathfrak{U}_n} R(u)$$

is obtained, which proves (7).

3. Upper Bounds for Eigenvalues

Based upon the theoretical investigation just discussed in Sections 1 and 2, a famous method for obtaining upper bonds for eigenvalues has been developed, which is known today as the *Rayleigh–Ritz method.* This method provides a straightforward and efficient means of computing non-increasing upper bounds for an arbitrary but finite number of eigenvalues of any operator in class $\mathscr{S}$.

The main idea of the method is to restrict a given operator to a finite-dimensional subspace of its domain, yielding a matrix problem for which the eigenvalues are numerically computable. It then follows from Poincaré's result (1.2) that the computed eigenvalues are upper bounds for the initial eigenvalues of the given operator.

Let us now describe the method in detail. Consider an operator $A \in \mathscr{S}$ for which the eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots$ are unknown. For an arbitrary integer n, choose any n linearly independent vectors $v_1, v_2, \ldots, v_n \in \mathfrak{D}$. In applications, these vectors are usually called *test functions* and satisfy the boundary conditions of the original problem. Such vectors generate a subspace $\mathfrak{V}_n$. If we now restrict the operator A to an operator from $\mathfrak{V}_n$ to $\mathfrak{V}_n$, we have the eigenvalue problem

$$V_n A V_n u = \Lambda u, \tag{1}$$

where V_n is the orthogonal projection onto $\mathfrak{V}_n$. From (1.2), the eigenvalues $\Lambda_1 \leq \Lambda_2 \leq \cdots \leq \Lambda_n$ satisfy the inequalities

$$\lambda_1 \leq \Lambda_1, \quad \lambda_2 \leq \Lambda_2, \ldots, \quad \lambda_n \leq \Lambda_n.$$

It remains to solve explicitly for $\Lambda_1, \Lambda_2, \ldots, \Lambda_n$. For the moment, let us introduce an orthonormal basis $v_1' v_2', \ldots, v_n'$ for $\mathfrak{V}_n$. For $u \in \mathfrak{V}_n$, we

have $u = V_n u$, so that we can write

$$u = V_n u = \sum_{i=1}^{n} (u, v_i')v_i', \qquad AV_n u = \sum_{i=1}^{n} (u, v_i')Av_i', \tag{2}$$

and

$$V_n A V_n u = \sum_{j=1}^{n} \sum_{i=1}^{n} (u, v_i')(Av_i', v_j')v_j'. \tag{3}$$

Substituting (2) and (3) in (1), we obtain the equation

$$\sum_{j=1}^{n} \sum_{i=1}^{n} (u, v_i')(Av_i', v_j')v_j' - \Lambda \sum_{i=1}^{n} (u, v_i')v_i' = 0. \tag{4}$$

Taking inner products of both sides of (4) with v_k', we have the set of equations

$$\sum_{i=1}^{n} (u, v_i')(Av_i', v_k') - \Lambda(u, v_k') = 0 \qquad (k = 1, 2, \ldots, n). \tag{5}$$

The system (5) is a matrix eigenvalue problem which has nontrivial solutions if and only if

$$\det\{(Av_i', v_k') - \Lambda\, \delta_{ik}\} = 0 \qquad (i, k = 1, 2, \ldots, n). \tag{6}$$

We now transform Eq. (6) into an equation that involves only the arbitrary basis $v_1, v_2, \ldots, v_n$. In fact, we have

$$v_i' = \sum_{j=1}^{n} \beta_{ij} v_j \qquad (i = 1, 2, \ldots, n), \tag{7}$$

where $\{\beta_{ij}\}$ is a nonsingular matrix. Therefore, we can write

$$(Av_i', v_k') = \sum_{j=1}^{n} \beta_{ij} \sum_{h=1}^{n} \beta_{kh}(Av_j, v_h) \tag{8}$$

and

$$\delta_{ik} = (v_i', v_k') = \sum_{j=1}^{n} \beta_{ij} \sum_{h=1}^{n} \beta_{kh}(v_j, v_h). \tag{9}$$

Using Eqs. (8) and (9) in (6), we see that

$$\det\{(Av_i', v_k') - \Lambda\, \delta_{ik}\} = \det[\{\beta_{ij}\}\{(Av_j, v_h) - \Lambda(v_j, v_h)\}\{\beta_{hk}^T\}].$$

Since $\{\beta_{ij}\}$ is nonsingular, the system (5) has a nontrivial solution if and only if

$$\det\{(Av_j, v_h) - \Lambda(v_j, v_h)\} = 0 \qquad (j, h = 1, 2, \ldots, n). \tag{10}$$

Equation (10) in principle gives an explicit determination of the eigenvalues $\Lambda_1, \Lambda_2, \ldots, \Lambda_n$.

A natural question which arises in the consideration of this method is whether or not the upper bounds $\Lambda_1, \Lambda_2, \ldots, \Lambda_n$ are improved (that is, decreased) by taking more vectors, say $v_{n+1}, v_{n+2}, \ldots, v_{n+r}$. The answer is that the new eigenvalues are not "worse" upper bounds than the old eigenvalues. To be more precise, let $\Lambda_1 \leq \Lambda_2 \leq \cdots \leq \Lambda_n$ and $w_1, w_2, \ldots, w_n$ denote the eigenvalues and eigenvectors of (1) and let $\Lambda_1' \leq \Lambda_2' \leq \cdots \leq \Lambda_n'$ denote the first n eigenvalues of $V_{n+r} A V_{n+r} u = \Lambda u$. Then, from the minimum–maximum principle, the jth eigenvalue ($j = 1, 2, \ldots, n$) of $V_{n+r} A V_{n+r}$ is given by

$$\Lambda_j' = \min_{\mathfrak{V}_j \subset \mathfrak{V}_{n+r}} \max_{u \in \mathfrak{V}_j} R(u).$$

For the particular subspace $\mathfrak{W}_j = sp\{w_1, w_2, \ldots, w_j\}$, we have

$$\Lambda_j' \leq \max_{u \in \mathfrak{W}_j} R(u) = \Lambda_j,$$

which means that the upper bounds are not increased by increasing the dimension of the finite-dimensional subspace. However, there is no guarantee that the bounds are better. There is no improvement, for instance, if we take eigenfunctions as test functions.

REMARK. There is no doubt that the inequality

$$\lambda_1 \leq R(u) \qquad (u \in \mathfrak{D})$$

was given by Rayleigh. Although the inequalities for higher eigenvalues are usually attributed to Ritz, it has been noted that seemingly both Rayleigh and Ritz were familiar only with the inequality $\lambda_1 \leq \Lambda_1$ [W10, P5]. In fact, in the paper of Ritz [R2] where he is supposed to have obtained upper bounds for 17 eigenvalues for a free plate, he actually has only found 17 upper bounds for λ_1. In order to have upper bounds for 17 eigenvalues, Ritz would have been required to consider the determinant of a 17×17 matrix, but we were unable to find in Ritz's papers such a determinant, which could not be easily overlooked. However, since Ritz emphasized the numerical applications, although without sufficient theoretical foundation, there is a historical justification for calling this result the *Rayleigh–Ritz method.*

Let us mention that the equations of the Rayleigh–Ritz method can also be obtained in the following way, which is familiar in applications. Consider the minimum of (Av, v) for all $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ for fixed test functions $v_1, v_2, \ldots, v_n \in \mathfrak{D}$ under the side condition $(v, v) = 1$.

(The choice of the test functions is left to the discretion of the computer.) Introduce a Lagrange multiplier Λ and write the equations

$$(\partial/\partial\alpha_k)[(Av, v) - \Lambda(v, v)] = 0 \qquad (k = 1, 2, \ldots, n). \tag{11}$$

Then, the determinant of (11) is exactly (10). Of course, this procedure would not of itself yield the results given in this section for higher eigenvalues.

4. A Necessary and Sufficient Criterion in the Minimum–Maximum Theory

As in the case of the new maximum–minimum theory [W18], which was actually developed earlier, the classical choice of the subspace, by which we mean here $\mathfrak{V} = \mathfrak{U}_n$, is not the only case for which the minimum is attained in (2.4). We shall now give a set of necessary and sufficient conditions on the subspace $\mathfrak{V}$ for equality (2.5) to hold [S5]. They are much simpler than the corresponding conditions for the maximum–minimum theory (see Section 7.4) which were given earlier.

Theorem 1. *A necessary condition for the equality*

$$\lambda_n = \max_{u \in \mathfrak{V}} R(u) \tag{1}$$

to hold is that the dimension r of $\mathfrak{V}$ *satisfy the inequality*

$$r \leq M(n), \tag{2}$$

where $M(n)$ *(or, for short,* M*) is defined by*

$$M = \max\{j \mid \lambda_j = \lambda_n\}. \tag{3}$$

Let $v_1, v_2, \ldots, v_r$ *denote any basis for* $\mathfrak{V}$*. Assuming that* (2) *holds, we have the following necessary and sufficient condition for equality* (1)*: the quadratic form with the symmetric matrix*

$$\{([A - \lambda_n I]v_i, v_k)\} \qquad (i, k = 1, 2, \ldots, r) \tag{4}$$

is negative-semidefinite.

Proof. In order to establish our necessary condition, suppose that equality (1) holds for some $\mathfrak{V}$ such that

$$M < r = \dim \mathfrak{V}. \tag{5}$$

Applying Theorem 2.1, we have

$$\lambda_r \leq \max_{u \in \mathfrak{V}} R(u) = \lambda_n = \lambda_M, \tag{6}$$

which contradicts inequality (5).† In developing the proof of the necessary and sufficient condition, let us for the moment consider as our basis for $\mathfrak{B}$ the set of eigenvectors $w_1, w_2, \ldots, w_r$ of the operator VAV. In this case, matrix (4) becomes the diagonal matrix

$$\begin{Bmatrix} \Lambda_1 - \lambda_n & 0 & 0 & \cdots & 0 \\ 0 & \Lambda_2 - \lambda_n & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & & & \Lambda_r - \lambda_n \end{Bmatrix}. \tag{7}$$

If equality (1) holds, it follows that

$$\Lambda_i = (Aw_i, w_i) = R(w_i) \leq \lambda_n \qquad (i = 1, 2, \ldots, r). \tag{8}$$

Therefore, the quadratic form defined by matrix (7) is negative-semi-definite. Conversely, if the quadratic form of matrix (7) is negative-semi-definite, then we have (8). Since every $u \in \mathfrak{B}$, $(u \neq 0)$ can be written as

$$u = \sum_{i=1}^{r} \gamma_i w_i,$$

we have

$$R(u) = \left\{ \sum_{i=1}^{r} |\gamma_i|^2 (Aw_i, w_i) \right\} \Big/ \left\{ \sum_{i=1}^{r} |\gamma_i|^2 \right\} \leq \lambda_n. \tag{9}$$

Combining inequalities (2.4) and (9), we obtain

$$\lambda_n \leq \max_{u \in \mathfrak{B}} R(u) \leq \lambda_n,$$

so that equality (1) holds. Let us now obtain the criterion for a general basis $v_1, v_2, \ldots, v_r$. Since $\{v_i\}$ and $\{w_i\}$ are both bases for the space $\mathfrak{B}$, there exists a matrix $\{\alpha_{ij}\}$ such that

$$v_i = \sum_{j=1}^{r} \alpha_{ij} w_j \qquad (i = 1, 2, \ldots, r).$$

Therefore, matrix (4) is equal to

$$\left\{ \left([A - \lambda_n I] \sum_{j=1}^{r} \alpha_{ij} w_j, \sum_{h=1}^{r} \alpha_{kh} w_h \right) \right\}$$
$$= \{\alpha_{ij}\}\{([A - \lambda_n I]w_j, w_h)\}\{\alpha_{hk}^T\} \qquad (i, k = 1, 2, \ldots, r).$$

† It could happen that λ_M is the largest isolated eigenvalue at the beginning of the spectrum, in which case λ_∞ would play the role of λ_r in (6); see Theorem 3.

From this, we see that a change in basis of $\mathfrak{V}$ has the same effect as changing the basis upon which matrix (7) is represented. Therefore, the quadratic form given by matrix (4) will be negative-semidefinite if and only if the quadratic form given by (7) is negative-semidefinite.

REMARK. It should be pointed out that in the above criterion the value λ_n appears explicitly. This is not surprising, since even in the classical theory (1.3.2), which gives only sufficient conditions for equality, the eigenvectors $u_1, u_2, \ldots, u_n$ appear explicitly (see also Chapter 7).

In some of the earlier papers on this subject [S5, S8], the criterion was formulated in the following way.

Theorem 2. *A necessary condition for* (1) *to hold is that the dimension r of* $\mathfrak{V}$ *satisfy* (2). *Assuming that* (2) *holds, a necessary and sufficient condition for* (1) *to hold is that for every* $\varepsilon \geqq 0$ *the quadratic form*

$$\{([A - (\lambda_n + \varepsilon)I]v_i, v_k)\} \qquad (i, k = 1, 2, \ldots, r)$$

be negative-semidefinite.

The proof is virtually identical to the proof given above. The criterion with $\varepsilon \geq 0$ most closely resembles the analogous criterion for the earlier new maximum–minimum theory, as we shall see in Section 7.4.

EXAMPLE 1. Using our criterion, we can now give a nonclassical subspace $\mathfrak{V}$ for which equality (1) holds. Let $\lambda_1 < \lambda_2 < \lambda_3$ and $m = n = M = 2$. Consider the subspace $\mathfrak{V} = sp\{u_2, u_1 + \beta u_3\}$, where

$$0 < |\beta|^2 \leq (\lambda_2 - \lambda_1)/(\lambda_3 - \lambda_2). \tag{10}$$

Every $v \in \mathfrak{V}$ $[(v, v) = 1]$ can be written as

$$v = \xi u_1 + \eta u_2 + \xi\beta u_3,$$

where $|\xi|^2 + |\eta|^2 + |\xi|^2|\beta|^2 = 1$. Therefore, we have, by (10),

$$\begin{aligned} R(v) &= |\xi|^2\lambda_1 + |\eta|^2\lambda_2 + |\xi|^2|\beta|^2\lambda_3 \\ &= |\xi|^2\lambda_1 + (1 - |\xi|^2 - |\xi|^2|\beta|^2)\lambda_2 + |\xi|^2|\beta|^2\lambda_3 \\ &= |\xi|^2(\lambda_1 - \lambda_2) + \lambda_2 + |\xi|^2|\beta|^2(\lambda_3 - \lambda_2) \\ &\leq |\xi|^2(\lambda_1 - \lambda_2) + \lambda_2 + |\xi|^2(\lambda_2 - \lambda_1) = \lambda_2. \end{aligned} \tag{11}$$

This implies

$$\max_{v \in \mathfrak{V}} R(v) \leq \lambda_2.$$

However, from (2.4), we have

$$\lambda_2 \leq \max_{u \in \mathfrak{V}} R(v)$$

and therefore

$$\lambda_2 = \max_{v \in \mathfrak{V}} R(v) \tag{12}$$

holds. Observe that when $\beta = 0$ this example reduces to the classical case $\mathfrak{V} = \mathfrak{U}_2$, in which case (12) is trivially satisfied. When $|\beta|^2 = (\lambda_2 - \lambda_1)/(\lambda_3 - \lambda_2)$, we have equality throughout (11), and therefore

$$\lambda_2 = R(v)$$

for all $v \in \mathfrak{V}$, $v \neq 0$. This is the limiting case, since we have

$$R(u_1 + \beta u_3) > \lambda_2 \qquad \text{if} \qquad |\beta|^2 > (\lambda_2 - \lambda_1)/(\lambda_3 - \lambda_2).$$

For this example, the matrix (7) of Theorem 1 is

$$\begin{pmatrix} 0 & 0 \\ 0 & (\lambda_1 - \lambda_2) + |\beta|^2(\lambda_3 - \lambda_2) \end{pmatrix},$$

which is clearly negative-semidefinite for all β satisfying the inequalities (10).

One might assume that an example such as the above is somewhat exceptional. Surprisingly, it turns out, as in Section 7.6, that there exists a "nonclassical" choice in all cases, with two very special exceptions, which we shall now prove. We begin by giving a more precise definition of a nonclassical choice.

Definition 1. For a given index n ($n = 1, 2, \ldots$), a subspace $\mathfrak{V}_n$ is said to be a nonclassical choice for the minimum of the maximum in (2.3) if $\mathfrak{V}_n \not\subset sp\{u_1, u_2, \ldots, u_{M(n)}\}$.

Then we have the following general result [S13].

Theorem 3. *The only two cases in which there does not exist a nonclassical choice for the minimum of the maximum in* (2.3) *are* (i) *A is an operator on a finite-dimensional space and λ_n is its greatest eigenvalue, and* (ii) *$\lambda_n = \lambda_1$, even if $\mathfrak{H}$ is infinite-dimensional.*

PROOF. If, for a fixed λ_n, there is a point in the spectrum of A, not necessarily an eigenvalue, which is greater than λ_n, then, since λ_n is isolated and the spectrum is closed, we choose the smallest such element, say ξ.

Let $\mathfrak{U} = \text{sp}\{u_1, u_2, \ldots, u_{M(n)}\}$. In view of the existence of ξ, we know that $\mathfrak{U} \neq \mathfrak{D}$. Since $\dim \mathfrak{U} < \infty$ and $\mathfrak{D}$ is dense, it follows from a lemma of Gohberg–Krein (Lemma A.1) that there exists a vector $w \in \mathfrak{U}^{\perp} \cap \mathfrak{D}$ $[(w, w) = 1]$. Letting $\eta = R(w)$, we note that $\lambda_n < \xi \leq \eta$. If $\lambda_1 < \lambda_n$, we choose β so that

$$0 < |\beta|^2 \leq (\lambda_n - \lambda_1)/(\eta - \lambda_n)$$

and set

$$v_1 = (1 + |\beta|^2)^{-1/2}(u_1 + \beta w)$$

and

$$v_2 = u_2, \quad v_3 = u_3, \ldots, \quad v_n = u_n.$$

Then, we have

$$([A - \lambda_n I]v_1, v_1) = (1 + |\beta|^2)^{-1}[(\lambda_1 - \lambda_n) + |\beta|^2(\eta - \lambda_n)] \leq 0$$

and

$$([A - \lambda_n I]v_j, v_j) = \lambda_j - \lambda_n \leq 0 \qquad (j = 2, 3, \ldots, n).$$

The diagonal matrix $\{([A - \lambda_n I]v_i, v_j)\}$ is negative-semidefinite and, therefore, it follows from Theorem 2 that the minimum in (2.3) is attained for the nonclassical choice $\mathfrak{V}_n = \text{sp}\{v_1, v_2, \ldots, v_n\}$. If λ_n is the greatest eigenvalue of an operator on a finite-dimensional space $\mathfrak{H}_N$, then $\mathfrak{H}_N = \text{sp}\{u_1, u_2, \ldots, u_{M(n)}\}$, so that nonclassical choices do not exist. On the other hand, if $A \in \mathscr{S}$ and $\lambda_1 = \lambda_n$, we assume that $\mathfrak{V}_n$ is a nonclassical choice for (2.3). Then, we have

$$\lambda_1 = \min_{u \in \mathfrak{D}} R(u) \leq \min_{u \in \mathfrak{V}_n} R(u) \leq \max_{u \in \mathfrak{V}_n} R(u) = \lambda_n$$

so that $\mathfrak{V}_n \subset \mathfrak{E}_n$ (where $\mathfrak{E}_n$ denotes the eigenspace of λ_n), which contradicts our assumption that $\mathfrak{V}_n$ is nonclassical and completes the proof.

The conclusion in the case $\lambda_1 = \lambda_n$ above is actually a special case of the following theorem, which will be used in Section 8.6.

Theorem 4. *A necessary and sufficient condition on the space $\mathfrak{V}_r$ of Theorem* 2.1 *that the simultaneous equalities*

$$\lambda_1 = \Lambda_1, \quad \lambda_2 = \Lambda_2, \ldots, \quad \lambda_r = \Lambda_r \tag{13}$$

hold is that $\mathfrak{V}_r$ be generated by eigenvectors corresponding to $\lambda_1, \lambda_2, \ldots, \lambda_r$.

PROOF. The sufficiency of the condition is obvious. To prove necessity, assume that the equalities (13) hold. Then, we have

$$VAVw_i = \lambda_i w_i \qquad (i = 1, 2, \ldots, r).$$

Since $VAVw_1 = \lambda_1 w_1$, it follows that

$$\lambda_1 = \min_{u \in \mathfrak{D}} R(u) = R(w_1).$$

From the variational theory (see Section 1.3), this means that w_1 is an eigenvector corresponding to λ_1. Now, suppose that $w_1, w_2, \ldots, w_k$ are eigenvectors corresponding to $\lambda_1, \lambda_2, \ldots, \lambda_k$ $(1 \leq k < r)$. Then, we have

$$\lambda_{k+1} = \min_{\substack{u \in \mathfrak{D} \\ u \perp w_1, w_2, \ldots, w_k}} R(u) = R(w_{k+1})$$

and therefore by the same reasoning as in the case of w_1, w_{k+1} is an eigenvector corresponding to λ_{k+1}, which completes the proof.

5. The Principle of Monotonicity

We shall now derive an important application of the *minimum-maximum principle*, namely to carry over to all λ_n statements that are evident for the first eigenvalue λ_1. Such an application was first conceived by Weyl [W31], who proved the same results as given below, not from Poincaré's inequalities, but from an important inequality which he introduced for this purpose. Weyl's inequality, which will be discussed in Chapter 3, became the foundation of the classical maximum-minimum theory. His approach has been used nearly everywhere in texts since 1910. The presentation given here seems to be of more recent times and has appeared in [W5, FR1, S11].

Let A' be an operator of class $\mathscr{S}$ such that

$$\mathfrak{D}(A') \subset \mathfrak{D}(A) \tag{1}$$

and

$$(Au, u) \leq (A'u, u) \qquad \text{for all} \quad u \in \mathfrak{D}(A'). \tag{2}$$

Then, we say that A is *dominated by* A' and write $A \leq A'$. It is obvious from Rayleigh's principle (1.3.1) that $\lambda_1 \leq \lambda_1'$. The following theorem shows that every eigenvalue of A is not greater than the corresponding eigenvalue of A'.

Theorem 1. *(Monotonicity principle) If A' and A are operators of class $\mathscr{S}$ satisfying conditions* (1) *and* (2), *then the eigenvalues λ_i' and λ_i of A' and A, respectively, satisfy the inequalities*

$$\lambda_i \leq \lambda_i' \qquad (i = 1, 2, \ldots). \tag{3}$$

PROOF. For any index i, consider $\mathfrak{U}_i' = \mathrm{sp}\{u_1', u_2', \ldots, u_i'\}$. Then, by the minimum–maximum principle (2.3), we have

$$\lambda_i \leq \max_{u \in \mathfrak{U}_i'}[(Au, u)/(u, u)] \leq \max_{u \in \mathfrak{U}_i'}(A'u, u)/(u, u)] = \lambda_i', \tag{4}$$

which yields (3).

In the case $(Au, u) < (A'u, u)$ for all $u \in \mathfrak{D}(A')$, it is clear that in (4) we have a strict inequality, and therefore

$$\lambda_i < \lambda_i' \qquad (i = 1, 2, \ldots). \tag{5}$$