*A Minimax Theory for Overdamped Networks**

R. J. DUFFIN

Carnegie Institute of Technology, Pittsburgh, Pennsylvania

Communicated by C. LOEWNER

1. Introduction. Dynamical systems with friction are much more difficult to treat than conservative systems. It is the aim of this paper to show that a degree of simplicity is restored by considering the extreme case of large amounts of friction.

The theory of small vibrations of conservative systems is essentially the theory of a pair of definite quadratic forms: the kinetic-energy quadratic form a and the potential-energy quadratic form c . Two quadratic forms may be brought to diagonal forms simultaneously. This diagonalization defines the normal modes of vibration and their characteristic frequencies. It also reveals the orthogonality property of the normal modes.

The theory of small motions of systems with friction is the theory of the triplet of definite quadratic forms a , b , and c . Here b is the Rayleigh dissipation function. As RAYLEIGH [7] observed, the theory is complicated by the fact that it is usually impossible to diagonalize three quadratic forms simultaneously. It results that the normal modes of motion are not orthogonal.

In this paper dynamical systems are considered which are heavily damped. Such systems may be defined precisely by the inequality $b^2 - 4ac > 0$, termed the *overdamping* condition. This inequality is assumed to hold for any virtual mode of motion of the system. A normal mode for a conservative system is of the form qe^{ikt} where q is independent of the time t . A normal mode for an overdamped system is of the form qe^{kt} . In both cases k is real and may be termed the eigenvalue. (For intermediate amounts of friction, damped oscillations result. This would lead to complex eigenvalues. The overdamping condition rules out this type of behavior.)

* The work on this paper was sponsored by the Office of Ordnance Research, U. S. Army, Contract DA-36-061-ORD-378.

Probably the most elegant treatment of conservative systems is furnished by the variational method. The starting point is the Rayleigh quotient $(c/a)^{\frac{1}{2}}$. This functional evaluated for a normal mode gives the corresponding eigenvalue, that is, $k = (c/a)^{\frac{1}{2}}$. The smallest eigenvalue may be defined as the minimum value of $(c/a)^{\frac{1}{2}}$ for all virtual modes.

For overdamped systems a similar procedure proves possible. The Rayleigh quotient is replaced by the functional $[-b \pm (b^2 - 4ac)^{\frac{1}{2}}]/2a$. The eigenvalues fall into two distinct classes, termed primary and secondary. The minimum of the functional, taking the plus sign, gives the smallest primary eigenvalue. The minimum taking the minus sign gives the smallest secondary eigenvalue. (If a is singular, the last statement must be modified.)

More generally it proves possible to define the ordered sequence of eigenvalues by the well-known minimax principle. For conservative systems it is also possible to define this sequence by a step-by-step procedure. However, the latter method is based on the orthogonality of the various normal modes.

The minimax principle not only defines the eigenvalues but also makes evident several important properties of the system. A consequence for conservative systems is that a constraint tends to increase the frequency of each normal mode. Correspondingly, a constraint on an overdamped system tends to increase k for each normal mode. This notion of constraint justifies the Rayleigh-Ritz method for the approximate determination of the eigenvalues.

The minimax method was developed by FISCHER [3], WEYL [9], and COURANT [1]. The Rayleigh-Ritz method was developed by RAYLEIGH, POINCARÉ [5], RITZ, and GALERKIN. Further ramification of these ideas will be found in references [8], [2], and [6].

Only systems with a finite number of degrees of freedom are considered in this paper.

2. Properties of overdamped dynamical systems. Let Q denote an n -dimensional vector space. If u and v are vectors of Q , then (u, v) denotes their scalar product. The norm of v is $\|v\| = (v, v)^{\frac{1}{2}}$. All quantities introduced will be real valued.

Let A , B , and C be three n by n matrices. Quadratic forms a , b , and c are defined by

$$(1) \quad a(v) = (Av, v), \quad b(v) = (Bv, v), \quad \text{and } c(v) = (Cv, v).$$

The fundamental hypotheses of this paper are:

- I. A , B , and C are symmetric.
- II. $a(v) \geq 0$, $b(v) \geq 0$, and $c(v) \geq 0$.
- III. $b^2(v) - 4a(v)c(v) > 0$.

Here v is an arbitrary non-zero vector.

Consider the vector differential equation

$$(2) \quad A\ddot{q} + B\dot{q} + Cq = f.$$

This may be regarded as the Lagrangian equation of a linear dynamical system (or of a linear electrical network). Here the vector q gives the configuration of the system at time t and $\dot{q} = dq/dt$. The masses, damping, and springs of the system are determined by the matrices A , B , and C respectively. The kinetic energy is $a(\dot{q})/2$, and the potential energy is $c(q)/2$. The vector f gives the forces applied to the system.

From the physical point of view *I* and *II* are not hypotheses, as they are demanded by energy considerations. On the other hand, *III* is an arbitrarily imposed hypothesis. It is termed the *overdamping condition*. The extreme opposite of *III* would be $b = 0$, which implies $B = 0$. This corresponds to an undamped system.

If no forces are applied to the system, then

$$(3) \quad A\ddot{q} + B\dot{q} + Cq = 0.$$

As is well known, the solutions of (2) are determined essentially by the solution of the homogeneous equation (3). The first step in the solution of (3) is the determination of the roots of the *characteristic equation*

$$(4) \quad \nabla(x) \equiv \det(x^2 A + xB + C) = 0.$$

Let r denote the rank of the matrix A . Then:

(a) Counting multiplicity, there are precisely $n + r$ roots of equation (4).

(b) These roots all lie on the negative real axis, that is, $-\infty < x \leq 0$.

(Statements preceded by letters are theorems which will be proved in Section 7 to follow.)

Let $h_1 \leq h_2 \leq \dots \leq h_r$ denote the r smallest roots; these are termed *secondary eigenvalues*. Let $k_1 \leq k_2 \leq \dots \leq k_n$ denote the n largest roots; these are termed *primary eigenvalues*.

$$(c) \quad h_r < k_1.$$

The general solution of (3) may be written in the form

$$(d) \quad q = \sum_1^n \epsilon_i u_i e^{k_i t} + \sum_1^r \delta_i w_i e^{h_i t}.$$

Here ϵ_i and δ_i denote arbitrary constant scalars. The u_i are constant vectors termed *primary eigenvectors*. The w_i are constant vectors termed *secondary eigenvectors*.

(e) $u_1, u_2, \dots, u_n$ are linearly independent.

(f) $w_1, w_2, \dots, w_r$ are linearly independent.

Let $v_1, v_2, \dots, v_m$ be an arbitrary set of m independent vectors. These generate an m -dimensional subspace V of Q . Let

$$(5) \quad A_{ij}^* = (Av_i, v_j)$$

define the matrix elements of an m by m matrix A^* . Clearly A^* is symmetric. Let $\rho_1, \rho_2, \dots, \rho_m$ be a set of scalars. These may be thought of as the rectangular coordinates of a vector q^* in an m -dimensional vector space Q^* . Let $v = \sum_1^m \rho_i v_i$. Then

$$(A^*q^*, q^*) = \sum_1^m \sum_1^m A_{ij}^* \rho_i \rho_j = \sum_1^m \sum_1^m (Av_i, v_j) \rho_i \rho_j = (Av, v) \geq 0.$$

Thus A^* satisfies *II*. Define B^* and C^* likewise. Then it is obvious that the inequality *III* is satisfied by the quadratic forms a^* , b^* , and c^* defined by the matrices A^* , B^* , and C^* .

Term the original dynamical system the unconstrained system S . Suppose that $n - m$ linear homogeneous constraints are imposed on S so that the vector q is confined to the m -dimensional subspace V spanned by the vectors $v_1, v_2, \dots, v_m$. This gives rise to a constrained system S^* with m degrees of freedom. If $\rho_1, \rho_2, \dots, \rho_m$ are taken as the Lagrangian coordinates of S^* , then the Lagrangian equation is of the form $A^*\ddot{q}^* + B^*\dot{q}^* + C^*q^* + f^* = 0$. But A^* , B^* , and C^* satisfy *I*, *II*, and *III*, so the imposition of homogeneous linear constraints on an overdamped dynamical system S results in an overdamped system S^* .

The minimax principle gives the following relationships between the eigenvalues of the unconstrained system S and the constrained system S^* :

- (g) The largest primary eigenvalue of $S^* \leq$ the largest primary eigenvalue of S .
The second largest primary eigenvalue of $S^* \leq$ the second largest primary eigenvalue of S , etc.
- (h) The above statement also applies to the secondary eigenvalues.
- (i) The smallest primary eigenvalue of $S^* \geq$ the smallest primary eigenvalue of S . The second smallest primary eigenvalue of $S^* \geq$ the second smallest primary eigenvalue of S , etc.
- (j) If A is non-singular, then the above statement also applies to the secondary eigenvalues.

If n is "large," it is difficult to solve equation (4) for the eigenvalues. The following analog of the Rayleigh-Ritz procedure could be employed to estimate the eigenvalues. Let m be a "small" number and choose an arbitrary set of m independent vectors $v_1, v_2, \dots, v_m$. Then compute the matrices A^* , B^* , and C^* by use of relation (5). Solve the equation $\det(x^2 A^* + x B^* + C^*) = 0$ in order to determine the eigenvalues of the "constrained system." Then the inequalities (g), (h), (i), and (j) give forbidden intervals for the true eigenvalues. Repeating this procedure for a new set of m vectors would supply more information.

The minimax principle also relates the eigenvalues of two systems having the same number of degrees of freedom. For example, let the *varied system* S'' be the same as S except that $(B''v, v) \geq (Bv, v)$ for any vector v . The varied system S'' may be described as having more damping than S . Then

(k) The primary eigenvalues of $S'' \geq$ the primary eigenvalues of S .

(l) The secondary eigenvalues of $S'' \leq$ the secondary eigenvalues of S .

Similar theorems can be derived for systems in which A , B , and C are all varied.

3. The primary functional. In developing the theory it is somewhat more convenient to replace hypothesis *II* by the weaker hypothesis

$$II' \quad b(v) \geq 0.$$

For each non-zero vector v the *primary functional* is defined as

$$(6) \quad p(v) = -2c(v)/(b(v) + d(v))$$

where

$$(7) \quad d(v) = (b^2(v) - 4a(v)c(v))^{\frac{1}{2}} > 0.$$

If ρ is a non-zero scalar, then obviously

$$(8) \quad p(\rho v) = p(v).$$

It follows that it suffices to consider v to be on the surface of the unit sphere, that is, $\|v\| = 1$.

On the surface of the unit sphere a , b , and c are continuous and uniformly bounded. By condition *III* it follows that d is continuous and has positive upper and lower bounds. By condition *II'* it follows that the same is true for $b + d$. Thus $p(v)$ is a continuous and uniformly bounded functional of the space Q . It is understood that p is not defined for the zero vector.

Lemma 1. $p(v) = x$ if and only if x satisfies

$$(9) \quad x^2 a(v) + xb(v) + c(v) = 0$$

and

$$(10) \quad 2xa(b) + b(v) > 0.$$

Proof. Consider $(b + d)^2(p^2 a + pb + c) = 4c^2 a - 2(b + d)cb + b^2 c + 2bdc + d^2 c = (4ac - b^2 + d^2)c = 0$. Thus $x = p(v)$ satisfies (9). Also $(b + d)(2pa + b) = -4ac + b^2 + bd = (b + d)d$. Thus $2pa + b = d > 0$, and (10) is satisfied.

If x satisfies (9), then

$$(11) \quad (2xa + b)^2 = d^2.$$

Making use of (10) gives $(2xa + b) = d$. If $a \neq 0$, then $x = (d - b)/2a = -2c/(d + b) = p$. If $a = 0$, $x = -c/b = p$. This completes the proof.

For each scalar x let a symmetric matrix T_x be defined as

$$(12) \quad T_x = x^2 A + xB + C$$

so

$$(13) \quad (T_x v, v) = x^2 a(v) + xb(v) + c(v).$$

If $T_x v = 0$, $v \neq 0$, then v is said to be an *eigenvector* and x is said to be an *eigenvalue*. Clearly (9) is satisfied. If $x = p(v)$, then v is said to be a *primary eigenvector* and x is said to be a *primary eigenvalue*.

Lemma 2. *Let u be a primary eigenvector and let z be any vector such that $z \neq 0$ and $z + u \neq 0$.*

$$(14) \text{ If } p(z) = p(u), \text{ then } p(u + z) = p(u).$$

$$(15) \text{ If } p(z) > p(u), \text{ then } p(u + z) > p(u).$$

$$(16) \text{ If } p(z) < p(u), \text{ then } p(u + z) < p(u).$$

Proof. By (8) it is permissible to suppose that u and z are independent. Let $k = p(u)$ and $v = (1 - \rho)u + \rho z$ where ρ is a parameter. By independence v does not vanish for any value of ρ . By definition $T_k u = 0$. Since T_k is symmetric, $(T_k z, u) = (T_k u, z) = 0$. Thus

$$(17) \quad k^2 a + kb + c = (T_k v, v) = \rho^2 (T_k z, z).$$

First consider the case (14). Then $k = p(z)$, and it follows that the right side of (17) vanishes. Thus relation (11) with $x = k$ gives

$$2ka + b = \pm d.$$

This holds for all values of ρ and the left side is a continuous function of ρ . The right side never vanishes, so the left side is always of one sign. For $\rho = 0$ then $v = u$, and by Lemma 1, $2ka + b > 0$. Thus $2ka + b > 0$ for all values of ρ , and by Lemma 1, $p(v) = k$. For $\rho = \frac{1}{2}$ then $v = (u + z)/2$, and it follows that $p(u + z) = p(v) = k$.

Now consider the case (15) and suppose on the contrary that $p(u + z) \leq k$. Since $p(v)$ is a continuous function of ρ , it follows that $p(v) = k$ for some value of ρ in the range $\frac{1}{2} \leq \rho \leq 1$. Thus the left side of (17) vanishes for this value of ρ . Hence $(T_k z, z) = 0$ and the right side of (17) vanishes for all ρ . By the same argument as employed above it follows that $p(v) = k$ for all ρ . For $\rho = 1$ this gives a contradiction and (15) is proved. The proof of (16) is analogous.

Lemma 3. *The sum of independent primary eigenvectors having the eigenvalue k is a primary eigenvector with the eigenvalue k .*

Proof. Since T_k is a linear operation, the sum of the eigenvectors is an eigenvector. It follows from (14) that the sum is a primary eigenvector.

The *multiplicity* of an eigenvalue k is defined to be the maximum number of independent primary eigenvectors with eigenvalue k .

Lemma 4. *Let $u_1, u_2, \dots, u_m$ be primary eigenvectors with corresponding eigenvalues $k_1 < k_2 < \dots < k_m$, ($m > 1$). Then*

$$k_1 < p(u_1 + u_2 + \dots + u_m) < k_m$$

and the vectors are independent.

Proof. Since $p(u_1) \neq p(u_2)$, it follows from (8) that u_1 and u_2 are independent. By (15) and (16) then $k_1 < p(u_1 + u_2) < k_2$. Proceeding by induction, suppose the lemma is true for the case $m - 1$. Let $z = u_1 + \dots + u_{m-1}$. Then $p(z) < k_{m-1}$ and by (16) $p(z + u_m) < k_m$. Let $z_1 = u_2 + \dots + u_m$. Then $p(z_1) > k_2$ and by (15), $p(u_1 + z_1) > k_1$. This completes the induction.

4. The minimax principle. A number $P(Y)$ is associated with each subspace Y of dimension one or greater by the relation

$$(18) \quad P(Y) = \text{l.u.b. } p(y) \text{ for } y \text{ in } Y.$$

The i^{th} primary minimax value k_i is then defined as

$$(19) \quad k_i = \text{g.l.b. } P(Y) \text{ for } Y \text{ of dimension } i.$$

These relations lead to finite numbers because p is bounded.

Lemma 5. *For some vector y in Y*

$$(20) \quad P(Y) = p(y).$$

Proof. If (20) is not true, there is a sequence of vectors $y_1, y_2, \dots$ all contained in Y such that $p(y_m) \rightarrow P(Y)$ as $m \rightarrow \infty$. It may be supposed that $\|y_m\| = 1$. The points y_m have at least one limit point, say y and $\|y\| = 1$. Thus for a suitable subsequence, $y_m \rightarrow y$. By the continuity of p then, $p(y_m) \rightarrow p(y)$. But y is in Y , so (20) is true.

Lemma 6. *For some i -dimensional subspace U*

$$(21) \quad k_i = P(U).$$

Proof. If (21) is not true, there is a sequence of i -dimensional subspaces $Y_1, Y_2, \dots$ such that $P(Y_m) \rightarrow k_i$ as $m \rightarrow \infty$. Let $e_1^m, e_2^m, \dots, e_i^m$ denote an orthonormal basis for Y_m . As $m \rightarrow \infty$ through a suitable subsequence the basis will converge; in other words, $e_1^m \rightarrow e_1, e_2^m \rightarrow e_2, \dots, e_i^m \rightarrow e_i$. The vectors $e_1, e_2, \dots, e_i$ are an orthonormal set and serve as a basis for a space U of dimension i . For some u in U , then, $p(u) = P(U)$ by Lemma 5. If $u = d_1 e_1 + d_2 e_2 + \dots + d_i e_i$, let $y_m = d_1 e_1^m + d_2 e_2^m + \dots + d_i e_i^m$. Then for the subsequence $y_m \rightarrow u$. Hence

$$P(Y_m) \geq p(y_m) \rightarrow p(u) = P(U).$$

Thus $k_i \geq P(U)$. But by (19), $k_i \leq P(U)$, so (21) is true.

By virtue of Lemmas 5 and 6 equivalent forms of (18) and (19) are

$$(18a) \quad P(Y) = \max p(y) \text{ for } y \text{ in } Y$$

$$(19a) \quad k_i = \min P(Y) \text{ for } Y \text{ of dimension } i.$$

In particular

$$(22) \quad k_1 = \min p(y) \text{ for } y \text{ in } Q,$$

$$(23) \quad k_n = \max p(y) \text{ for } y \text{ in } Q.$$

Suppose that $k_i = P(U)$ for U of dimension i . If $i \geq 2$, let U' be an $i - 1$ dimensional subspace of U . Then by the definitions $k_{i-1} \leq P(U') \leq P(U) = k_i$, so

$$(24) \quad k_1 \leq k_2 \leq \cdots \leq k_n.$$

Lemma 7. *A primary minimax value is a primary eigenvalue.*

Proof. Now let k denote the i^{th} minimax value, so $k = P(U)$ where U is i -dimensional. The transformation $J_m = I + m^{-1} T_k$ is nonsingular if m is a sufficiently large positive integer. Here I is the identity transformation. Thus $Y_m = J_m U$ is an i -dimensional subspace. Let $P(Y_m) = p(y_m)$ where y_m is in Y_m and $\|y_m\| = 1$. For some u_m in U then, $y_m = u_m + m^{-1} T_k u_m$. It is clear from this relation that $\|u_m\|$ is bounded as $m \rightarrow \infty$. Thus there is a vector u in U such that $u_m \rightarrow u$ as $m \rightarrow \infty$ through a suitable sequence of integers. Consequently $y_m \rightarrow u$, and so $\|u\| = 1$.

Let v be a given vector. Then if $p = p(v)$, $p^2 a + pb + c = 0$, so if x is an arbitrary number, $-(p - x)(pa + xa + b) = x^2 a + xb + c = (T_x v, v)$. Let $v = w + z$ and let $x = p(w)$. Then

$$(25) \quad -(p - x)(pa + xa + b) = 2(T_x w, z) + (T_x z, z).$$

In (25) take $v = y_m$, $w = u_m$, and $z = m^{-1} T_k u_m$. Then $p(y_m) = P(Y_m) \geq P(U) \geq p(u_m)$, so $p - x = p(y_m) - p(u_m) \geq 0$. Also $pa + xa + b \rightarrow 2p(u)a(u) + b(u) > 0$. Thus for m sufficiently large the left side of (25) can not be positive.

Multiply (25) by m and substitute on the right, so

$$(26) \quad 0 \geq (T_x u_m, T_k u_m) + m^{-1}(T_x T_k u_m, T_k u_m).$$

It was seen above that $p(y_m) \geq k \geq p(u_m)$. Since $y_m \rightarrow u$ and $u_m \rightarrow u$ it follows that $p(u) \geq k \geq p(u)$, so $k = p(u)$. Hence $x = p(u_m) \rightarrow k$. Thus in the limit (26) becomes $0 \geq (T_k u, T_k u)$. Therefore $T_k u = 0$ and u is a primary eigenvector and k is a primary eigenvalue.

Theorem 1. *There is an independent set of n primary eigenvectors $u_1, u_2, \dots, u_n$. The corresponding eigenvalues are the minimax values $k_1, k_2, \dots, k_n$. Any other primary eigenvector u is a linear combination of vectors of the set having the same eigenvalue as u .*

Proof. According to Lemma 7 there is at least one eigenvector associated with each minimax value. If k is a multiple minimax value (j -fold), the eigenvectors are selected as follows. Suppose $k = k_i = k_{i-1} = \cdots = k_{i-j+1}$. Then $k_i = P(U)$ where U is i -dimensional. According to the proof of Lemma 7 there is a primary eigenvector u_i contained in U and $p(u_i) = k_i$. Let U_1 be the subspace of U consisting of vectors orthogonal to u_i . Thus $P(U_1) \leq k_i$. But U_1 has dimension $i - 1$, so $P(U_1) \geq k_{i-1}$. Hence $P(U_1) = k$. Then by the proof of Lemma 7 there is a primary eigenvector u_{i-1} contained in U_1 and $p(u_{i-1}) = k_{i-1}$. This process is repeated $j - 1$ times. The resulting set of j eigenvectors is independent because it is orthogonal.

Thus a primary eigenvector u_i has been associated with each minimax value k_i for $i = 1, 2, \cdots, n$. Consider a linear combination

$$(27) \quad z = \sum_1^n \rho_i u_i.$$

The terms of the series corresponding to a multiple minimax value k are grouped. If one of the coefficients ρ_i does not vanish in a group, then this group has a resultant $v \neq 0$. This follows from the orthogonality condition just derived. By Lemma 3 it follows that v is a primary eigenvector with eigenvalue k . Hence

$$(28) \quad z = \sum v_i$$

where the v_i are eigenvectors with distinct eigenvalues. By Lemma 4 it follows that $z \neq 0$. Thus the n vectors u_i are independent.

An independent set of n vectors is a basis for Q , so any vector may be represented in the form (27). In particular $z = u$, a primary eigenvector, may be represented. It follows that the summation in (28) contains only one term. Otherwise a relation between eigenvectors with distinct eigenvalues would result, in contradiction to Lemma 4.

Corollary. *The ordered set of n primary minimax values and the ordered set of n primary eigenvalues are identical.*

Proof. The first part of Theorem 1 shows that there are n eigenvalues, counting multiplicity. The last part shows that there are no other eigenvalues.

Let the *reflected problem* be defined by replacing A by $-A$ and C by $-C$. Clearly the hypotheses I, II', and III are still satisfied. The primary functional merely has its sign reversed. The primary eigenvectors for the reflected problem are identical with those of the original problem; the corresponding eigenvalues are equal in magnitude but opposite in sign.

Theorem 2. *If Y is a subspace of dimension ≥ 1 , let*

$$(29) \quad \bar{P}(Y) = \min p(y) \text{ for } y \text{ in } Y.$$

Then for $i = 1, 2, \cdots, n$ the primary minimax value k_i is given by

$$(30) \quad k_i = \max \bar{P}(Y)$$

for all subspaces of dimension $n + 1 - i$.

Proof. Let $k'_1, k'_2, \dots, k'_n$ denote the ordered set of primary eigenvalues for the reflected problem. Then $k'_1 = -k_n, k'_2 = -k_{n-1}$, etc. By virtue of the Corollary, these same relations hold for the minimax values. Let $p'(y)$ denote the primary functional for the reflected problem. By relation (18a) then $P'(Y) = \max p'(y) = \max -p(y) = -\min p(y)$. Thus $P'(Y) = -\bar{P}(Y)$. By relation (19a), $k'_{n+1-i} = \min P'(Y) = -\max \bar{P}(Y)$, and the proof is complete.

5. The secondary functional. The *secondary functional* $s(v)$ is defined for a vector v if and only if $a(v) \neq 0$. The definition is

$$(31) \quad 2sa + b = -d.$$

Thus s satisfies $s^2 a + sb + c = 0$. (Note that the primary functional always satisfies $2pa + b = d$.) A vector w is said to be a *secondary eigenvector* if $T_h w = 0$ and $h = s(w)$. Then h is said to be a *secondary eigenvalue*. It is possible for a vector to be both a primary and a secondary eigenvector. Conceivably a secondary eigenvalue could also be a primary eigenvalue, but Theorem 4 to follow will show that this is impossible.

Theorem 3. *Let r be the rank of A . Then there is an independent set of r vectors $w_1, w_2, \dots, w_r$. Each vector of the set is a secondary eigenvector. Any other secondary eigenvector is a linear combination of vectors of the set with the same eigenvalue.*

Proof. Let the *reciprocal problem* be defined by interchanging A and C . It is obvious that *I*, *II'*, and *III* are still satisfied. The primary functional for the reciprocal problem is

$$(32) \quad p^0 = -2a/(b + d).$$

Thus if $a \neq 0$, then $s = 1/p^0$.

Let w be a primary eigenvector of the reciprocal problem with eigenvalue k^0 . If $k^0 \neq 0$, then $h = 1/k^0$ is a secondary eigenvalue of the original problem. The converse statement also holds. If $k^0 = 0$, then w may be termed an improper eigenvector of the original problem. An improper eigenvector satisfies

$$(33) \quad Aw = 0.$$

The vectors which satisfy (33) constitute a subspace. It is known that this subspace has dimension $n - r$. Any non-zero vector of this subspace is a primary eigenvector of the reciprocal problem with eigenvalue zero. Then Theorem 3 follows from Theorem 1.

Theorem 4. *The range of the primary functional and the range of the secondary functional have no common value.*

Proof. The range of the primary functional is the closed interval (k_1, k_n) . The range of the primary functional for the reciprocal problem is the closed interval (k_1^0, k_n^0) . The inversion of the interval (k_1^0, k_n^0) gives the range of the secondary functional. This gives either one or two intervals. These intervals may extend to infinity. Thus if there is a value in common, then either (i), the range of the

secondary functional is contained in (k_1, k_n) , or (ii), the secondary functional takes on the value k_1 or the value k_n . It is no loss of generality to assume that (ii) holds because the situation would be reversed for the reciprocal problem.

Thus suppose there are vectors u and z such that u is a primary eigenvector with eigenvalue k and $s(z) = k = p(u)$. Let $v = \rho z + (1 - \rho)u$, then $k^2 a(v) + kb(v) + c(v) = (T_k v, v) = \rho^2 (T_k z, z)$. For $\rho = 1$ the left side vanishes, hence $(T_k z, z) = 0$ and the left side vanishes for all ρ . For $\rho = 1$, $2ka + b < 0$ by (31). For $\rho = 0$, $2ka + b > 0$. But $2ka + b$ is a continuous function of ρ , so it would vanish for some intermediate value of ρ . For this value of ρ then, $0 = (2ka + b)^2 = d^2$. But $d = 0$ implies $v = 0$. This is a contradiction, as the relation $s(z) = p(u)$ shows that z and u are independent and v can not vanish for any value of ρ .

6. The characteristic equation. Let H be a non-singular n by n matrix, and let H_t be its transpose. The "transformed" matrices $H_t A H$, $H_t B H$, and $H_t C H$ will also satisfy *I*, *II'*, and *III*. Moreover, this transformation does not change the eigenvalues, the characteristic equation, or rank.

Lemma 8. *Let r be the rank of A . Then $\nabla(x) = \det(x^2 A + xB + C)$ is a polynomial of degree $n + r$ in x .*

Proof. Let A be transformed to diagonal form. Then exactly $n - r$ diagonal elements vanish. It may be supposed that these are the first $n - r$ places. Clearly a further transformation may be made which keeps A diagonal and at the same time diagonalizes B in the first $n - r$ places. Thus if B_{ij} denotes a matrix element of B , then $B_{ij} = 0$ for $i \neq j$ in the range $i, j = 1, 2, \dots, n - r$. In this range any diagonal term $B_{ii} \neq 0$; otherwise condition *III* would be violated for a suitable choice of the vector v .

After the above transformation, the first $n - r$ diagonal elements of the matrix $x^2 A + xB + C$ are of the first degree in x . The remaining r diagonal elements are of the second degree. Thus the diagonal contributes a term of degree $(n - r) + 2r = n + r$ to the determinant ∇ . Terms which involve at least one off diagonal element are seen to be of lower degree.

Lemma 9. *For a given constant k , let R be the rank of the matrix $k^2 A + kB + C$. Then $\nabla(x)$ has a zero of multiplicity $n - R$ at $x = k$.*

Proof. Let $x = k + t$, then

$$(34) \quad x^2 A + xB + C = t^2 A + tB_k + C_k$$

and

$$(35) \quad \nabla(x) = \det(t^2 A + tB_k + C_k)$$

where

$$(36) \quad B_k = 2kA + B \text{ and } C_k = k^2 A + kB + C.$$

If $b_k = (B_k v, v)$ and $c_k = (C_k v, v)$, then

$$(37) \quad b_k^2 - 4ac_k = b^2 - 4ac.$$

Hence the matrices A , B_k , and C_k satisfy conditions *I* and *III*. Transform C_k to diagonal form and B_k partly to diagonal form in the same way that A and B were transformed in the proof of Lemma 8. Then the first $n - R$ diagonal elements of the matrix $t^2 A + tB_k + C_k$ have zeros of multiplicity one at $x = k$. The remaining R elements do not have zeros. Thus the diagonal contributes a term having a zero of multiplicity $n - R$ to the determinant. Terms which involve at least one off diagonal element are seen to be of higher multiplicity.

Theorem 5. *A necessary and sufficient condition that k be a zero of $\nabla(x)$ of multiplicity j is that k be an eigenvalue of multiplicity j . This accounts for $n + r$ real zeros of $\nabla(x)$; there are no other zeros.*

Proof. Note that $C_k \equiv T_k$. Let k be a given constant, then the eigenvalue multiplicity j is the maximum number of independent vectors for which $C_k v = 0$. Hence $j = n - R$ where R is the rank of C_k . Thus the first part of Theorem 5 follows from Lemma 9.

According to Theorem 1 there are n primary eigenvalues (counting multiplicity). According to Theorem 3 there are r secondary eigenvalues (counting multiplicity). According to Theorem 4 the primary and secondary eigenvalues are distinct. These observations show that $\nabla(x)$ has $n + r$ real zeros. But $\nabla(x)$ has precisely $n + r$ zeros by Lemma 8, so the zeros of $\nabla(x)$ are all accounted for.

7. Conclusion. In order to prove the statements of Section 2, it is now assumed that hypothesis *II* holds.

Lemma 10. *The eigenvalues are negative or zero valued. Every primary eigenvalue exceeds every secondary eigenvalue.*

Proof. Since $a \geq 0$, $b \geq 0$, and $c \geq 0$, it follows that the ranges of the primary and secondary functionals must be on the negative real axis. This proves the first statement.

If v is a vector such that $a(v) \neq 0$, then $ap(v) - as(v) = d$. Thus $p(v) > s(v)$. It follows from Theorem 4 that if z is any other vector then $p(z) > s(v)$. In particular this proves the second statement.

Thus Theorem 5 and Lemma 10 prove (a), (b), and (c). Theorems 1 and 3 prove (e) and (f). The number of arbitrary constants in (d) is $n + r$, which is the same as the degree of $\nabla(x)$. Thus by a theorem of CHRYSTAL [4] the relation (d) is the general solution of the differential equation.

Consider now the constrained system with matrices A^* , B^* , and C^* . It was seen that there is a one-to-one linear correspondence between the m -dimensional space Q^* and the m -dimensional subspace V of Q . Moreover, if $q^* \sim v$, then $(A^* q^*, q^*) = (Av, v)$, $(B^* q^*, q^*) = (Bv, v)$, and $(C^* q^*, q^*) = (Cv, v)$. Consequently, $p^*(q^*) = p(v)$, and $s^*(q^*) = s(v)$. The minimax values of the constrained system are, therefore, just as well determined by considering the subspace V . So if $1 \leq i \leq m$, then

$$(38) \quad k_i^* = \min P(Y)$$

for all i -dimensional subspaces Y contained in V . So

$$(39) \quad k_i^* \geq k_i$$

and this proves statement (i).

By Theorem 2

$$(40) \quad k_i^* = \max \bar{P}(Y)$$

for all $(m - i + 1)$ -dimensional subspaces Y contained in V . But $k_{i+n-m} = \max \bar{P}(Y)$ for all $(m - i + 1)$ -dimensional subspaces Y contained in Q . Thus

$$(41) \quad k_i^* \leq k_{i+n-m}$$

and this proves statement (g).

The proofs of (h) and (j) are obtained by considering the reciprocal problem.

Now consider the varied system S'' which satisfies $(B''v, v) \geq (Bv, v)$ for any vector v . Then $b'' \geq b$ and $d'' \geq d$; hence $p''(v) \geq p(v)$. Then $k_i'' = P''(Z)$ where Z is i -dimensional. There is a vector z contained in Z such that $P(Z) = p(z)$. This gives the inequalities $k_i'' = P''(Z) \geq p''(z) \geq p(z) = P(Z) \geq k_i$. This proves statement (k). Clearly the same relationship holds for the reciprocal problem. Taking reciprocals changes the direction of the inequality, and this proves statement (l).

The above proof depends, essentially, only on the condition that $p'' \geq p$. It is clear, therefore, how to treat more general variations.

REFERENCES

- [1] R. COURANT & D. HILBERT, *Methoden der mathematischen Physik*, Berlin (1931).
- [2] R. J. DUFFIN & A. SCHILD, *The effect of small constraints on natural vibrations*, Proc. Fifth Symposium in Applied Mathematics, American Mathematical Society (1954).
- [3] E. FISCHER, *Über quadratische Formen mit reellen Koeffizienten*, Monatshefte für Math. und Physik, **16**, pp. 234-249 (1905).
- [4] E. L. INCE, *Ordinary Differential Equations*, Chapter 6, p. 145, New York (1944).
- [5] H. POINCARÉ, *Sur les équations aux dérivées partielles de la physique mathématique*, Amer. Jour. of Math., **12**, pp. 211-294 (1890). (See pp. 259-261.)
- [6] G. PÓLYA & M. SCHIFFER, *Convexity of functionals by transplantation*, O.N.R. Technical Report No. 14, Oct. 19, 1953, Stanford University (1953).
- [7] RAYLEIGH, *The Theory of Sound*, New York (1945). (See pp. 136-138.)
- [8] A. WEINSTEIN, *Les vibrations et le calcul des variations*, Portugaliae Mathematica, **2**, pp. 36-55 (1941).
- [9] H. WEYL, *Ramifications, old and new, of the eigenvalue problem*, Bull. Amer. Math. Soc., **56**, pp. 115-139 (1950).