AN EXTREMUM PROPERTY OF SUMS OF EIGENVALUES[1]

HELMUT WIELANDT

We present in this note a maximum-minimum characterization of sums like $\alpha_3+\alpha_7+\alpha_8$ where $\alpha_1 \geqq \cdots \geqq \alpha_n$ are the eigenvalues of a hermitian $n \times n$ matrix. The result contains the classic characterization of α_m as well as the maximum property of $\alpha_1+\alpha_2+\cdots+\alpha_m$ given recently by Fan [4]. Though the result is valid also for a completely continuous hermitian operator in Hilbert space, we shall for the sake of simplicity assume the dimension to be finite. As an application we obtain linear inequalities relating the eigenvalues of the sum of two hermitian matrices to the eigenvalues of the summands.

THEOREM 1. *Let R_n be a unitary n-space with inner product (x, y). Let A be a hermitian operator on R_n with eigenvalues $\alpha_1 \geqq \alpha_2 \geqq \cdots \geqq \alpha_n$. Let S be a set of distinct natural numbers $\leqq n$ and $i<j<\cdots<l<m$ its elements. Then*

$$\begin{aligned}(1)\qquad & \alpha_i + \alpha_j + \cdots + \alpha_l + \alpha_m \\ & = \max_{R_i \subset \cdots \subset R_m;\ \dim R_\sigma = \sigma}\ \min_{x_\sigma \in R_\sigma;\ (x_\alpha, x_\beta) = \delta_{\alpha\beta}}\ \sum_{\sigma \in S} (Ax_\sigma, x_\sigma).\end{aligned}$$

More explicitly, formula (1) is equivalent to the following statements I, II.

I. Let $R_i \subset R_j \subset \cdots \subset R_l \subset R_m$ be given subspaces of R_n such that the dimensionality of R_σ is σ, for each $\sigma \in S$. Then there are vectors $x_i, x_j, \cdots, x_m$, such that

$$(2)\qquad x_\sigma \in R_\sigma \ (\sigma \in S), \qquad (x_\alpha, x_\beta) = \begin{cases} 1 & (\alpha = \beta), \\ 0 & (\alpha \neq \beta), \end{cases}$$

$$(3)\qquad \sum_{\sigma \in S} (Ax_\sigma, x_\sigma) \leqq \sum_{\sigma \in S} \alpha_\sigma.$$

II. There is a special sequence $E_i \subset E_j \subset \cdots \subset E_l \subset E_m$ of subspaces[2] of R_n such that for every orthonormal set of vectors $x_i, x_j, \cdots, x_m$, with $x_\sigma \in E_\sigma$ $(\sigma \in S)$, we have

$$\sum_{\sigma \in S} (Ax_\sigma, x_\sigma) \geqq \sum_{\sigma \in S} \alpha_\sigma.$$

Received by the editors January 25, 1954.

[1] This paper was prepared under a National Bureau of Standards contract with The American University, Washington, D. C.

[2] Subscripts of symbols for spaces denote the dimensionality throughout this paper.

The proof of II is easy. Let $e_1, \cdots, e_n$ be an orthonormal set of eigenvectors of A corresponding to $\alpha_1, \cdots, \alpha_n$. Define E_σ to be the subspace spanned by $e_1, \cdots, e_\sigma$. Then $(Ax_\sigma, x_\sigma) \geqq \alpha_\sigma$ $(x_\sigma \in E_\sigma, (x_\sigma, x_\sigma) = 1)$, hence

$$\sum_{\sigma \in S} (Ax_\sigma, x_\sigma) \geqq \sum_{\sigma \in S} \alpha_\sigma.$$

We are going to prove I by induction with respect to n.

(a) Let S contain all natural numbers $\sigma \leqq n$. Then we choose $x_1, \cdots, x_n$ satisfying (2) but otherwise arbitrary. The left-hand side of (3) is the trace of the matrix representing A with respect to the basis x_σ. Hence (3) holds. This especially applies to $n=1$.

(b) In what follows we may assume that there is a natural number $\leqq n$ which is not in S. The largest of these "gap numbers" will be denoted by g. We define f to be the largest number in S which is $<g$ if there exists such a number; if not, we define $f=0$. In either case we have $0 \leqq f < g \leqq n$. If $f>0$ then R_f is defined by hypothesis; if $f=0$ we define $R_f=0$, in accordance with our subscript convention.

We begin with the simplest case. Let $n \notin S$, that is $g=n$. Then we choose any subspace $\tilde{R}_{n-1}$ containing R_f. We define $\tilde{A}$ to be the unique hermitian operator on $\tilde{R}_{n-1}$ such that

$$(\tilde{A}x, x) = (Ax, x) \qquad (x \in \tilde{R}_{n-1}). \tag{4}$$

The eigenvalues $\tilde{\alpha}_1 \geqq \cdots \geqq \tilde{\alpha}_{n-1}$ of $\tilde{A}$ are known [3] to satisfy

$$\tilde{\alpha}_\nu \leqq \alpha_\nu \qquad (\nu = 1, \cdots, n-1). \tag{5}$$

By the induction hypothesis, there are orthonormal vectors $x_\sigma \in \tilde{R}_{n-1}$ such that $x_\sigma \in R_\sigma$ $(\sigma \in S)$ and

$$\sum_{\sigma \in S} (\tilde{A}x_\sigma, x_\sigma) \leqq \sum_{\sigma \in S} \tilde{\alpha}_\sigma.$$

This inequality, by virtue of (4) and (5), implies (3). This finishes the case $g=n$.

Now let $n \in S$, that is, $g<n$. We choose orthonormal eigenvectors $e_{g+1}, \cdots, e_n$ of A corresponding to $\alpha_{g+1}, \cdots, \alpha_n$. Together with R_f they span a space of dimension $\leqq (n-g)+f<n$. Hence we can choose some subspace $\tilde{R}_{n-1}$ such that

$$R_f \subseteq \tilde{R}_{n-1}, \ e_\nu \in \tilde{R}_{n-1} \qquad (\nu = g+1, \cdots, n). \tag{6}$$

Since $g+1, \cdots, n \in S$, we have

$$R_f \subseteq R_{g+1} \cap \tilde{R}_{n-1} \subseteq \cdots \subseteq R_{n-1} \cap \tilde{R}_{n-1} \subseteq \tilde{R}_{n-1}.$$

Since the dimension of $R_\nu \cap \tilde{R}_{n-1}$ is at least $\nu - 1$ we can choose subspaces $\tilde{R}_g, \cdots, \tilde{R}_{n-2}$ such that

$$\tilde{R}_g \subset R_{g+1}, \cdots, \tilde{R}_{n-2} \subset R_{n-1}, \tag{7}$$

$$R_i \subset \cdots \subset R_f \subset \tilde{R}_g \subset \cdots \subset \tilde{R}_{n-2} \subset \tilde{R}_{n-1}. \tag{8}$$

We define as before the operator $\tilde{A}$. Then (4) and (5) hold. By the induction hypothesis applied to $\tilde{A}$ and the subspaces (8), there are orthonormal vectors x_σ ($\sigma \in S$) such that

$$x_\sigma \in R_\sigma \ (\sigma < g), \qquad x_\sigma \in \tilde{R}_{\sigma-1} \ (\sigma > g), \tag{9}$$

$$\sum_{\sigma \in S} (\tilde{A} x_\sigma, x_\sigma) \leqq \sum_{\sigma \in S; \sigma < g} \tilde{\alpha}_\sigma + \sum_{\sigma = g}^{n-1} \tilde{\alpha}_\sigma. \tag{10}$$

Using (4), (5), (7), we find that

$$x_\sigma \in R_\sigma \qquad (\sigma \in S), \tag{9'}$$

$$\sum_{\sigma \in S} (A x_\sigma, x_\sigma) \leqq \sum_{\sigma \in S; \sigma < g} \alpha_\sigma + \sum_{\sigma = g}^{n-1} \tilde{\alpha}_\sigma. \tag{10'}$$

To complete the proof of Theorem 1 it is sufficient to show

$$\sum_{\sigma = g}^{n-1} \tilde{\alpha}_\sigma \leqq \sum_{\sigma = g+1}^{n} \alpha_\sigma. \tag{11}$$

From (5) we get only $\sum_{\sigma=g}^{n-1} \tilde{\alpha}_\sigma \leqq \sum_{\sigma=g}^{n-1} \alpha_\sigma$.

However, we know from (6) that $e_{g+1}, \cdots, e_n$ are eigenvectors of $\tilde{A}$, hence $\alpha_{g+1}, \cdots, \alpha_n$ are eigenvalues of $\tilde{A}$. Thus, we find for the smallest eigenvalues of $\tilde{A}$ the inequalities

$$\tilde{\alpha}_{n-1} \leqq \alpha_n, \ \tilde{\alpha}_{n-2} \leqq \alpha_{n-1}, \cdots, \tilde{\alpha}_g \leqq \alpha_{g+1}$$

which prove (11) and Theorem 1.

Applying Theorem 1 to $-A$ instead of A and denoting the subspace orthogonal to R_ν by $T_{n-\nu}$, we find the equivalent

THEOREM 1′. *Under the assumptions of Theorem* 1 *we have*

$$\begin{aligned} &\alpha_i + \alpha_j + \cdots + \alpha_l + \alpha_m \\ &= \min_{T_{i-1} \subset \cdots \subset T_{m-1};\ \dim T_{\sigma-1} = \sigma - 1} \ \max_{x_\sigma \perp T_{\sigma-1};\ (x_\alpha, x_\beta) = \delta_{\alpha\beta}} \sum_{\sigma \in S} (A x_\sigma, x_\sigma). \end{aligned} \tag{12}$$

REMARKS. (a) Theorem 1′ for $i=m$ is a classical result (Weyl [9], Courant [2]), but Theorem 1 for $i=m$ has also been explicitly mentioned by several authors. Theorem 1 for $i=1, j=2, \cdots, l=m-1$ has been published by Fan [4] in a simpler form. These references also apply to Theorem 2.

(b) It seems worth searching for an extremum characterization of $\sum c_\nu \alpha_\nu$ where $c_1, \cdots, c_n$ are arbitrarily given real numbers.

THEOREM 2. *Let A, B, C be hermitian operators on R_n such that $A+B=C$. Let the eigenvalues of A, B, C be α_ν, β_ν, γ_ν, respectively (numbered in decreasing order). Let S be a set of k natural numbers $i<j<\cdots<l<m$ such that $m \leqq n$. Then*

$$\gamma_i + \gamma_j + \cdots + \gamma_l + \gamma_m \leqq \alpha_i + \alpha_j + \cdots + \alpha_l + \alpha_m + \beta_1 + \beta_2 + \cdots + \beta_{k-1} + \beta_k. \tag{13}$$

PROOF. By II there are subspaces $R_i \subset R_j \subset \cdots \subset R_m$ such that

$$\gamma_i + \gamma_j + \cdots + \gamma_l + \gamma_m = \min_{x_\sigma \in R_\sigma;\ (x_\alpha, x_\beta) = \delta_{\alpha\beta}} \sum_{\sigma \in S} (Cx_\sigma, x_\sigma). \tag{14}$$

Keeping the subspaces R_σ fixed we choose vectors y_σ such that

$$y_\sigma \in R_\sigma, \qquad (y_\alpha, y_\beta) = \delta_{\alpha\beta}, \tag{15}$$

$$\sum_{\sigma \in S} (Ay_\sigma, y_\sigma) = \min_{x_\sigma \in R_\sigma;\ (x_\alpha, x_\beta) = \delta_{\alpha\beta}} \sum_{\sigma \in S} (Ax_\sigma, x_\sigma). \tag{16}$$

To prove Theorem 2 we estimate

$$\sum_{\sigma \in S} (Cy_\sigma, y_\sigma) = \sum_{\sigma \in S} (Ay_\sigma, y_\sigma) + \sum_{\sigma \in S} (By_\sigma, y_\sigma)$$

in two different ways. The left-hand side is $\geqq \gamma_i + \gamma_j + \cdots + \gamma_l + \gamma_m$ by (14) and (15). On the right-hand side we have by Theorem 1

$$\sum_{\sigma \in S} (Ay_\sigma, y_\sigma) \leqq \alpha_i + \alpha_j + \cdots + \alpha_l + \alpha_m,$$

and by Fan's theorem (i.e., Theorem 1 in case $i=1, j=2, \cdots$)

$$\sum_{\sigma \in S} (By_\sigma, y_\sigma) \leqq \beta_1 + \beta_2 + \cdots + \beta_{k-1} + \beta_k.$$

This completes the proof.

Theorem 2 can be shown to be equivalent to the following statement due to Lidskiĭ [7].[3]

Let α, β, γ be the points with coordinates α_ν, β_ν, γ_ν $(\nu = 1, \cdots, n)$ respectively. Define Γ to be the convex closure of the $n!$ points $\alpha + P\beta$ where P runs over all $n!$ permutation matrices. Then under the assumptions of Theorem 2 we have

$$\gamma \in \Gamma. \tag{17}$$

[3] It ought to be mentioned that the author did not succeed in completing the interesting sketch of the proof given by Lidskiĭ. This failure gave rise to the present investigation.

We prove that (13) and (17) imply each other. By a theorem of Hardy, Littlewood, and Pólya [5], the validity of the inequalities (13) (for every choice of $i, j, \cdots, l, m$) is a necessary and sufficient condition in order that there exist an $n\times n$ matrix $M=(m_{\mu\nu})$ such that

$$m_{\mu\nu} \geqq 0, \qquad \sum_{\mu} m_{\mu\nu} = \sum_{\nu} m_{\mu\nu} = 1, \tag{18}$$

$$\gamma - \alpha = M\beta. \tag{19}$$

On the other hand, the set of all matrices M satisfying (18) is known to be the convex closure of the $n!$ permutation matrices P (Birkhoff [1]; see also [6]). Hence the range of the points $M\beta$, where β is fixed and M runs through all matrices satisfying (18), is the convex closure of the $n!$ points $P\beta$. So (19) is true if, and only if, $\gamma-\alpha$ is in the convex closure of the points $P\beta$, that is, if and only if $\gamma\in\Gamma$.

REMARK. From (19) we see that *under the assumptions of Theorem* 2

$$F(\gamma_1 - \alpha_1, \cdots, \gamma_n - \alpha_n) \leqq F(\beta_1, \cdots, \beta_n) \tag{20}$$

holds for every S-convex function F of n variables (for definition and criteria of S-convexity, see Ostrowski [8]). *For instance, if $f(x)$ is a convex function then*

$$\sum_{\nu=1}^{n} f(\gamma_\nu - \alpha_\nu) \leqq \sum_{\nu=1}^{n} f(\beta_\nu). \tag{21}$$

BIBLIOGRAPHY

1. G. Birkhoff, *Three observations on linear algebra*, Universidad Nacional de Tucumán, Revista. Serie A. Matemáticas y Fisica Teórica vol. 5 (1946) pp. 147–151.
2. R. Courant, *Über die Eigenwerte bei den Differentialgleichungen der mathematischen Physik*, Math. Zeit. vol. 7 (1920) pp. 1–57, Theorem 3a.
3. R. Courant and D. Hilbert, *Methoden der mathematischen Physik*, Berlin, 1931, p. 28.
4. Ky Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations*. I, Proc. Nat. Acad. Sci. U.S.A. vol. 35 (1949) pp. 652–655, Theorem 1.
5. G. Hardy, J. Littlewood, and G. Pólya, *Inequalities*, Cambridge, 1934, p. 49.
6. A. Hoffman and H. Wielandt, *The variation of the spectrum of a normal matrix*, Duke Math. J. vol. 20 (1953) pp. 37–40, Lemma.
7. V. Lidskiĭ, *The proper values of the sum and product of symmetric matrices*, Doklady Akad. Nauk SSSR. vol. 75 (1950) pp. 769–772 (in Russian), Theorem 1.
8. A. Ostrowski, *Sur quelques applications des fonctions convexes et concaves au sens de I. Schur*, J. Math. Pures Appl. vol. 31 (1952) pp. 253–292, Paragraph 12.
9. H. Weyl, *Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen*, Math. Ann. vol. 71 (1912) pp. 441–479, Theorem 1.

AMERICAN UNIVERSITY