AN ABSTRACT OF THE THESIS OF

CHARLES GORDON LINDBERG for the Master of Science
(Name) (Degree)

in Mathematics presented on August 31, 1967
(Major) (Date)

Title A CONE ASSOCIATED WITH THE LYAPUNOV MAPPING

Abstract approved: Redacted for Privacy
(C. S. Ballantine)

In this paper we investigate the Lyapunov mapping

$$P \to AP + PA^*,$$

where A is a positive stable matrix and P is a hermitian matrix. In particular, for special positive stable A we characterize the image of the cone of positive definite matrices under this mapping. In Section V we give five conditions on an $n \times n$ hermitian matrix H and a general $n \times n$ positive stable matrix A that are equivalent to the condition: $H = AP + PA^*$ for some positive definite (hermitian) P.

A Cone Associated with the Lyapunov Mapping

by

Charles Gordon Lindberg

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

June 1968

APPROVED:

Redacted for Privacy

Associate Professor of Mathematics

In Charge of Major

Redacted for Privacy

Chairman of Department of Mathematics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented August 31, 1967

Typed by Carol Baker for Charles Gordon Lindberg

ACKNOWLEDGMENT

This author is indeed indebted to Dr. C.S. Ballantine for his guidance and encouragement during the work on this thesis.

TABLE OF CONTENTS

A CONE ASSOCIATED WITH THE LYAPUNOV MAPPING

I. INTRODUCTION

A large amount of research in matrix theory in the past few years has been concerned with the matrix equation

$$Y = AX + XA^* . \tag{1.1}$$

Lyapunov in his 1892 monograph The General Problem of Stability of Motion[1/] [8] established that the equation (1.1) could be used to obtain information pertaining to the location of the characteristic values of A. More recently, interest in equation (1.1) has taken a variety of directions. A survey of results linked with this matrix equation can be found in Givens' report Elementary Divisors and Some Properties of the Lyapunov Mapping $X \to AX + XA^*$ [5, p. 62-70].

One direction of study was suggested by Olga Taussky in the Bulletin of the American Mathematical Society [15, p. 711] when she queried:

> Let A be an $n \times n$ matrix with complex elements and characteristic roots with negative real parts. It is known (theorem of Lyapunov) that such matrices are

---

1/ The translation of the title of Lyapunov's monograph used here is taken from Givens [5, p. 3].

> characterized by the fact that a positive definite G exists with $AG + GA^*$ negative definite. What is the range of $AG + GA^*$ if G runs through all positive definite $n \times n$ matrices? A^* is the complex conjugate and transposed matrix.

This thesis addresses itself to this question. After considering some of the properties of the matrix equation (1.1) and the related mapping, we characterize the set $AP + PA^*$ for some special A, as P ranges over the set of positive definite matrices. In Section V (Theorem 6) we give five conditions on an $n \times n$ hermitian matrix H and a positive stable A that are equivalent to the condition: $H = AP + PA^*$ for some positive definite P.

A brief background of equation (1.1) is given now for the sake of completeness. As noted, Lyapunov is credited with the use of the equation (1.1) for the location of characteristic values. What he did was to establish a connection between the location of the characteristic values of a general $n \times n$ matrix and the signature of a quadratic form. The equation (1.1) results when one considers the matrices associated with the quadratic forms $H(x, x)$ and $P(x, x)$ which are related in the following way:

$$\frac{d}{dt} P(x, x) = H(x, x),$$

where

$$\frac{dx}{dt} = xA.$$

A is a real $n \times n$ matrix, and x is a real $1 \times n$ row vector. If we differentiate and make the appropriate substitutions, the following equation results:

$$P(xA, x) + P(x, xA) = H(x, x). \tag{1.2}$$

In matrix form, equation (1.2) can be expressed as

$$(xA)Px^T + xP(xA)^T = xHx^T,$$

or as

$$AP + PA^T = H. \tag{1.3}$$

$P = [p_{ij}]$ $(i, j = 1, 2, \cdots, n)$ and $H = [h_{ij}]$ $(i, j = 1, 2, \cdots, n)$ in (1.3) are the symmetric matrices formed respectively from the coefficients of the forms $H(x, x)$ and $P(x, x)$, and A^T is the transpose of the matrix A. The central result of Lyapunov is given in Theorem 1.

Theorem 1. The $n \times n$ real matrix A is stable (all of the characteristic values of A have negative real part) if and only if a positive definite $n \times n$ matrix $P = P^T$ exists such that

$$AP + PA^T = -I .$$

Proofs for Theorem 1 can be found in [10, 16].

We can immediately generalize Lyapunov's work to the case of an arbitrary complex matrix $A = [a_{ij}]$ $(i, j = 1, 2, \cdots, n)$. The quadratic forms $H(x,x)$ and $P(x,x)$ are replaced by the hermitian forms

$$H(x,x) = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij} x_i \overline{x}_j$$

and

$$P(x,x) = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} x_i \overline{x}_j \quad .$$

Correspondingly, the matrix equation (1.3) is replaced by the matrix equation

$$H = AP + PA^{*} ,$$

where $H = [h_{ij}]$ $(i, j = 1, 2, \cdots, n)$ and $P = [P_{ij}]$ $(i, j = 1, 2, \cdots, n)$ are the hermitian matrices formed respectively from the coefficients of the hermitian forms $H(x,x)$ and $P(x,x)$. A^{*} denotes the conjugate transpose of A. We follow Ballantine [1, p. 2] in that we will consider positive stable A rather than stable A. An $n \times n$ matrix is called <u>positive stable</u> provided each of its characteristic values has positive real part. We note that if A is positive stable, $-A$ is stable, and $-H$ is negative definite for positive definite H.

If we consider positive stable A rather than stable A, Theorem 1 can be generalized to complex matrices A in the following way:

Theorem 1*. The $n \times n$ complex matrix A is positive stable if and only if a positive definite matrix $P = P^*$ exists such that

$$AP + PA^* = I .$$

II. SOME PROPERTIES OF THE MAPPING $P \to AP + PA^*$

Since we wish to characterize the set $AP + PA^*$ as P ranges over the positive definite hermitian matrices, it will at times be more descriptive to utilize the mapping

$$[A]: P \to AP + PA^*,$$

rather than the equation

$$H = AP + PA^* .$$

The fact that the mapping $P \to AP + PA^*$ is completely determined by specification of $A = [a_{ij}]$ $(i, j = 1, 2, \cdots, n)$ justifies the use of the notation $[A]$ for the mapping. (Since brackets are also used in this thesis to enclose the components of a matrix and for grouping, every effort will be made to make it clear from context which meaning is intended.) We now list a number of properties of the mapping $[A]$, which we will call upon later in the thesis.

Lemma 1. The mapping $[A]$ in the space χ of complex matrices of order n is a linear transformation.

Proof: The mappings $P \to AP$ and $P \to PA^*$ are clearly linear. Thus the mapping $[A]: P \to AP + PA^*$ is linear.

It follows from Lemma 1 that, with respect to each ordered basis of χ, $[A]$ determines an $n^2 \times n^2$ matrix (namely, the matrix of $[A]$ relative to this basis). We will now show that if the ordered basis of χ is chosen properly, the matrix determined by $[A]$ is $A \otimes I + I \otimes \bar{A}$, where $\otimes$ denotes the Kronecker product. For a definition and some of the properties of the Kronecker product (which is also known as the direct or tensor product) see [9, p. 81-88] or [2, p. 227-229]. In each of the cited references the authors found no need to define the Kronecker product for nonsquare matrices. We will now define the Kronecker product of a row vector by a column vector, since this concept will be used to exhibit the particular basis of which we have spoken. Let z be an $n \times 1$ column vector and y a $1 \times n$ row vector. Then the Kronecker product $z \otimes y$ is defined to be the ordinary matrix product zy. We shall now describe the above mentioned basis of χ. Let e_i be an $n \times 1$ column vector with 1 in the i^{th} position and zeros elsewhere. Similarly, let f_j be a $1 \times n$ row vector with 1 in the j^{th} position and zeros elsewhere. Then

$$e_i \otimes f_j = \begin{bmatrix} 0 & \cdots & 0 \\ \cdot & & \cdot \\ \cdot & 1_{ij} & \cdot \\ \cdot & & \cdot \\ 0 & \cdots & 0 \end{bmatrix}, \qquad (2.1)$$

where the matrix on the right side of (2.1) designates the $n \times n$

matrix with 1 in the (i, j) position and zeros elsewhere (this matrix is sometimes denoted E_{ij}). If we require that $e_i \otimes f_j$ precede $e_i' \otimes f_j'$ for $i < i'$ or for $i = i'$ and $j < j'$, then $\{e_i \otimes f_j : i, j = 1, 2, \cdots, n\}$ forms an ordered basis of χ. The ordering described above is called <u>standard lexicographic ordering</u>. To show that the matrix of $[A]$ with respect to this basis is $A \otimes I + I \otimes \bar{A}$, we will consider the mapping $P \to AP$ and $P \to PA^*$ individually. The mapping $P \to AP$ is linear and its matrix with respect to the basis $\{e_i \otimes f_j : i, j = 1, 2, \cdots, n\}$ is determined by its action on the basis elements $e_i \otimes f_j$ $(i, j = 1, 2, \cdots, n)$. We have

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 \\ \cdot & & \cdot \\ \cdot & 1_{ij} & \cdot \\ \cdot & & \cdot \\ 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} 0 & \cdots & a_{1i} & \cdots & 0 \\ \cdot & & a_{2i} & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & & a_{ni} & \cdots & 0 \end{bmatrix} \quad (\text{the } a_{ki} \text{ in the } j^{th} \text{ column}),$$

so that

$$e_i \otimes f_j \to \sum_{k=1}^{n} a_{ki} (e_k \otimes f_j) \quad (i, j = 1, 2, \cdots, n) .$$

Therefore the matrix of $P \to AP$, ordering the $\{e_i \otimes f_j : i, j = 1, 2, \cdots, n\}$

standard lexicographically, is $A \otimes I$. Proceeding in a similar fashion with the transformation $P \to PA^*$, we find that a typical basis element maps as follows:

$$e_i \otimes f_j \to \sum_{k=1}^{n} \bar{a}_{kj} (e_i \otimes f_k) \quad (i, j = 1, 2, \cdots, n) .$$

Thus the matrix of this mapping with respect to the indicated basis is $I \otimes \bar{A}$. Hence the matrix of $[A]$ with respect to the $\{e_i \otimes f_j : i, j = 1, 2, \cdots, n\}$, ordered standard lexicographically, is

$$A \otimes I + I \otimes \bar{A} .$$

Lemma 2. If A has characteristic values λ_i $(i = 1, 2, \cdots, n)$, then $\lambda_i + \bar{\lambda}_j$ $(i, j = 1, 2, \cdots, n)$ are the characteristic values of $A \otimes I + I \otimes \bar{A}$ and hence are the characteristic values of the mapping $[A]$.

Proof: It is well known that we can choose P nonsingular such that $A_1 = PAP^{-1}$ is triangular with λ_i $(i = 1, 2, \cdots, n)$ on the diagonal of A_1, provided the λ_i $(i = 1, 2, \cdots, n)$ are the characteristic values of A [9, Theorem 41.3, p. 75]. $A_1 \otimes I + I \otimes \bar{A}_1$ is clearly triangular with $\lambda_i + \bar{\lambda}_j$ on the diagonal and thus has characteristic values $\lambda_i + \bar{\lambda}_j$ $(i, j = 1, 2, \cdots, n)$. The following calculations verify that $A_1 \otimes I + I \otimes \bar{A}_1$ is similar to $A \otimes I + I \otimes \bar{A}$:

$$A_1 \otimes I + I \otimes \overline{A}_1 = PAP^{-1} \otimes I + I \otimes \overline{PAP}^{-1}$$

$$= (P \otimes \overline{P})(A \otimes I)(P^{-1} \otimes \overline{P}^{-1}) + (P \otimes \overline{P})(I \otimes \overline{A})(P^{-1} \otimes \overline{P}^{-1})$$

$$= (P \otimes \overline{P})(A \otimes I + I \otimes \overline{A})(P \otimes \overline{P})^{-1},$$

since $(P \otimes \overline{P})$ is nonsingular. (The properties of the Kronecker product used above are proved or references to their proofs are given in [9, p. 82-83].) It follows that the characteristic values of $A \otimes I + I \otimes \overline{A}$ are $\lambda_i + \overline{\lambda}_j$ $(i, j = 1, 2, \cdots, n)$.

Lemma 3. If A is an $n \times n$ positive stable matrix, then the matrix $A \otimes I + I \otimes \overline{A}$ and thus the mapping $[A]$ are nonsingular.

Proof: Let A be positive stable with characteristic values λ_i $(i = 1, 2, \cdots, n)$. Then the characteristic values of $I \otimes A + I \otimes \overline{A}$ are $\lambda_i + \overline{\lambda}_j$ for $i, j = 1, 2, \cdots, n$ (Lemma 2). We have that $Re(\lambda_i + \overline{\lambda}_j) > 0$ $(i, j = 1, 2, \cdots, n)$ and therefore $A \otimes I + I \otimes \overline{A}$ is nonsingular. Here $Re(\lambda_i + \overline{\lambda}_j)$ means the real part of the complex number $\lambda_i + \overline{\lambda}_j$.

Lemma 4. Let A be any matrix of χ and let $\mathcal{H}$ be the set of all $n \times n$ hermitian matrices (thus $\mathcal{H}$ is a subset of χ). Then $[A]$ is a (real) linear mapping of $\mathcal{H}$ into $\mathcal{H}$ and at least one (real) matrix representing this real linear mapping is similar (over the complex field) to $A \otimes I + I \otimes \overline{A}$.

Proof: First, for P_1 and P_2 in $\mathcal{H}$ $r_1P_1 + r_2P_2$ is in $\mathcal{H}$ for all real r_1 and r_2. Thus $\mathcal{H}$ is a real linear space. Further, $P = P^*$ implies $(AP + PA^*)^* = AP + PA^*$, so that $[A]$ maps $\mathcal{H}$ into $\mathcal{H}$. The fact that the restriction of $[A]$ to $\mathcal{H}$ is a linear mapping follows from Lemma 1, $\mathcal{H}$ being a subset of χ and the real field being a subfield of the complex field. One can routinely verify that the matrices E_{kk}, $E_{jk}+E_{kj}$, $i(E_{jk}-E_{kj})$ $(j, k = 1, 2, \cdots, n,\ j < k)$ form a basis of $\mathcal{H}$ and also a basis of χ (over the complex field). Thus the (real) matrix $M[A]$ associated with this basis and the matrix $A \otimes I + I \otimes \bar{A}$ are matrices of $[A]: \chi \rightarrow \chi$ with respect to different bases and hence are similar.

Before proceeding we make the following definition: Let $\mathcal{R}$ be a subset of a real linear space. Then $\mathcal{R}$ is a <u>cone</u> provided $r_1R_1 + r_2R_2$ is in $\mathcal{R}$ for all R_1 and R_2 in $\mathcal{R}$ and all positive (real) r_1 and r_2. We note that the set $\mathcal{P}$ of positive definite matrices forms a cone. Further, if A is positive stable (or any complex $n \times n$ matrix), $[A]\mathcal{P}$ is a cone. To show this let H_1 and H_2 be in $[A]\mathcal{P}$, and let r_1 and r_2 be positive. Then $H_1 = [A]P_1$ and $H_2 = [A]P_2$ for some P_1 and P_2 in $\mathcal{P}$, so we have

$$r_1H_1 + r_2H_2 = r_1[A]P_1 + r_2[A]P_2 = [A](r_1P_1 + r_2P_2),$$

which is in $[A]\mathcal{P}$, since $\mathcal{P}$ is a cone.

Since we are primarily interested in how $[A]$ maps the cone of positive definite matrices $\mathcal{P}$, we will call the mapping $[A]: P \to AP + PA^*$, subject to the restriction that P be positive definite and A be positive stable, the Lyapunov mapping. The Lyapunov mapping is thus $[A]$ (for some positive stable A) restricted to $\mathcal{P}$, and its range is $[A]\mathcal{P}$. $[A]\mathcal{P}$ is the cone we are interested in characterizing. We note that both $[A]\mathcal{P}$ and the Lyapunov mapping depend on the positive stable A under consideration at the time.

We now state three theorems which give information concerning the cone $[A]\mathcal{P}$.

Theorem 2. If A is positive stable and H is positive definite, then

$$AP + PA^* = H \tag{2.2}$$

can be solved with P positive definite.

Proof: We follow Taussky [13]. Since H is positive definite, $H = RR^*$ for some nonsingular R. Then we have (to solve)

$$AP + PA^* = RR^*,$$

which is equivalent to

$$R^{-1}ARR^{-1}P(R^{-1})^{*} + R^{-1}P(R^{-1})^{*}R^{*}A^{*}(R^{*})^{-1} = I,$$

which can be written

$$A_1P_1 + P_1A_1^{*} = I, \tag{2.3}$$

where

$$A_1 = R^{-1}AR$$

$$P_1 = R^{-1}P(R^{-1})^{*} .$$

Now since A_1 is positive stable (it has the same characteristic values as A), it follows from Theorem 1* that (2. 3) can be solved with P_1 positive definite. Hence it follows that $P = RP_1R^{*}$ is positive definite and satisfies (2. 2).

Theorem 3. $H = AP + PA^{*}$ for some positive stable A and some positive definite P if and only if H has at least one positive characteristic value.

Stein gives a proof of Theorem 3 in his paper On the Ranges of Two Functions of Positive Definite Matrices [11]. From Theorem 2 and Theorem 3 it follows that the range of the Lyapunov mapping contains the cone of positive definite matrices but does not contain any nonpositive definite matrices. In Lemma 3, when we established that $[A]$ was nonsingular for positive stable A, we assured the existence of $[A]^{-1}$. This existence of $[A]^{-1}$ allows us to characterize the range $[A]\mathcal{P}$ of the Lyapunov mapping in the following way.

<u>Theorem 4.</u> Let A be positive stable and H be hermitian. Then H is in $[A]\mathcal{P}$ if and only if $[A]^{-1}H$ is positive definite.

In the next section (and in later sections) we use Theorem 4 to characterize the range of the Lyapunov mapping for some special positive stable matrices.

III. $[A]\mathcal{P}$ FOR DIAGONAL A

In this section we consider only positive stable matrices A which are diagonal. These matrices display their characteristic values, and since $[A]P$ and $[A]^{-1}H$ are easily calculated, we can use Theorem 4 to characterize the image cone $[A]\mathcal{P}$. The simplest kind of diagonal matrix is a scalar matrix, which is just a scalar multiple of the identity. For this reason we first examine the case where A is a positive stable scalar matrix λI. The restriction that $A = \lambda I$ is rather a severe one, and the results are indicative of the severity of the restriction. Let $A = \lambda I$ be positive stable and let H be hermitian. Then $\mathrm{Re}(\lambda) > 0$, and so H is in $[A]\mathcal{P}$ if and only if, for some $P \varepsilon \mathcal{P}$,

$$H = [A]P = (\lambda I)P + P(\lambda I)^{*} = (\lambda + \bar{\lambda})P,$$

i.e., if and only if H is positive definite (since $\lambda + \bar{\lambda} > 0$ and $\mathcal{P}$ is a cone). Thus for scalar A the cone $\mathcal{P}$ is mapped onto itself by $[A]$.

When A is an $n \times n$ nonscalar diagonal matrix, the range of the Lyapunov mapping is not so conveniently described. Before considering the $n \times n$ case, it is interesting to examine the 2×2 case in detail. A 2×2 complex matrix is hermitian if and only if it is of the form

$$\begin{bmatrix} w & x+iy \\ x-iy & z \end{bmatrix} ,$$

where w, x, y, z are real and $i = \sqrt{-1}$. If we fix the basis (for the 2×2 hermitian matrices) to be the one mentioned in Lemma 4, with the ordering E_{11}, $E_{12}+E_{21}$, $i(E_{12}-E_{21})$, E_{22}, then each 2×2 hermitian matrix is associated with a 4-tuple by the following mapping:

$$\begin{bmatrix} w & x+iy \\ x-iy & z \end{bmatrix} \rightarrow (w, x, y, z),$$

which is an isomorphism from the space of 2×2 hermitian matrices to the space of real 4-tuples. Now if we apply the usual distance to the space of 4-tuples, we can describe the cones $\mathcal{P}$ and $[A]\mathcal{P}$ geometrically since they are subsets of the set of 2×2 hermitian matrices. First, a point (w, x, y, z) (this represents a 2×2 hermitian matrix) corresponds to a positive definite (hermitian) matrix if and only if

$$w > 0 \quad \text{and} \quad wz - (x^2+y^2) > 0 \ .$$

(We get from the above conditions that $z > 0$ also.) The set of points satisfying this system of inequalities forms the interior of (one nappe of) a four-dimensional cone. The traces of the boundary of this cone in each of the three-dimensional coordinate hyperplanes (a coordinate hyperplane is the set obtained by setting one of the coordinates equal to zero and is labeled by the remaining coordinates) are given by the following:

(1) If $x = 0$, we have the w, y, z hyperplane. The conditions which describe the boundary of the cone in this hyperplane are

$$w > 0 \quad \text{and} \quad wz - y^2 = 0.$$

This is an elliptic cone of one nappe with axis $y = w - z = 0$ and which is tangent to the wy plane and the yz plane respectively along the positive w and z axes.

(2) If $y = 0$, we get the w, x, z hyperplane. The conditions that define the boundary in this hyperplane are

$$w > 0 \quad \text{and} \quad wz - x^2 = 0,$$

and the locus is similar to the one described in (1).

(3) If $w = 0$ or $z = 0$, we get the hyperplanes x, y, z or w, x, y. The trace in either of these hyperplanes is just a half-line: the positive half of the z axis in the former and the positive half of the w axis in the latter.

Now if we apply the mapping $[A]$ to the cone $\mathcal{P}$ of positive definite matrices, Lemma 4 assures that the matrices in the image $[A]\mathcal{P}$ will be hermitian. Thus the image can be described geometrically in the same manner as $\mathcal{P}$.

A 2×2 diagonal matrix is positive stable and nonscalar if and only if it is of the form $\begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$ with $\lambda \neq \mu$, $\mathrm{Re}(\lambda) > 0$ and $\mathrm{Re}(\mu) > 0$. Thus let A be of this form and let

$$P = \begin{bmatrix} w & x+iy \\ x-iy & z \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} a & b+ic \\ b-ic & d \end{bmatrix}$$

be hermitian. Then H is in the range of the Lyapunov mapping if and only if

$$P = [A]^{-1}H$$

is positive definite. The existence of $[A]^{-1}$ is guaranteed by Lemma 3. With A, P and H as given we have

$$H = \begin{bmatrix} a & b+ic \\ b-ic & d \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix} \begin{bmatrix} w & x+iy \\ x-iy & z \end{bmatrix} + \begin{bmatrix} w & x+iy \\ x-iy & z \end{bmatrix} \begin{bmatrix} \bar{\lambda} & 0 \\ 0 & \bar{\mu} \end{bmatrix}$$

$$= \begin{bmatrix} (\lambda+\bar{\lambda})w & (\lambda+\bar{\mu})(x+iy) \\ (\mu+\bar{\lambda})(x-iy) & (\mu+\bar{\mu})z \end{bmatrix} .$$

From this we get

$$P = \begin{bmatrix} w & x+iy \\ x-iy & z \end{bmatrix} = \begin{bmatrix} \frac{a}{\lambda+\bar{\lambda}} & \frac{b+ic}{\lambda+\bar{\mu}} \\ \frac{b-ic}{\mu+\bar{\lambda}} & \frac{d}{\mu+\bar{\mu}} \end{bmatrix} = [A]^{-1}H,$$

and it follows that H is in the range of the Lyapunov mapping if and only if

$$\frac{a}{\lambda+\bar{\lambda}} > 0 \quad \text{and} \quad \frac{ad}{(\lambda+\bar{\lambda})(\mu+\bar{\mu})} - \frac{b^2+c^2}{(\lambda+\bar{\mu})(\mu+\bar{\lambda})} > 0 .$$

Thus the point with coordinates (a, b, c, d) corresponds to a 2×2 hermitian matrix in $[A]\mathscr{P}$ if and only if

$$a > 0 \quad \text{and} \quad Kad-(b^2+c^2) > 0 ,$$

where

$$K = \frac{(\lambda+\bar{\mu})(\bar{\lambda}+\mu)}{(\lambda+\bar{\lambda})(\mu+\bar{\mu})} .$$

The following calculations verify that $K > 1$ and conversely that every $K > 1$ is of this form (with $Re(\lambda) > 0$, $Re(\mu) > 0$ and $\lambda \neq \mu$). Namely, we have that the following inequalities are pairwise equivalent:

$$|\lambda - \mu|^2 > 0$$

$$\lambda\bar{\lambda} + \mu\bar{\mu} > \lambda\bar{\mu} + \mu\bar{\lambda}$$

$$\lambda\bar{\lambda} + \mu\bar{\mu} + \bar{\lambda}\bar{\mu} + \lambda\mu > \lambda\bar{\mu} + \mu\bar{\lambda} + \bar{\lambda}\bar{\mu} + \lambda\mu$$

$$(\lambda + \bar{\mu})(\bar{\lambda} + \mu) > (\lambda + \bar{\lambda})(\mu + \bar{\mu})$$

$$\frac{(\lambda + \bar{\mu})(\mu + \bar{\lambda})}{(\lambda + \bar{\lambda})(\mu + \bar{\mu})} > 1 \quad .$$

The conditions which H must satisfy to be in the range of the Lyapunov mapping resemble those conditions a matrix P must satisfy to be in $\mathcal{P}$, the one difference being the coefficient K. Therefore, how this K affects the cone is of prime interest. For K = 1 (i.e., A scalar) the cone $[A]\mathcal{P} = \mathcal{P}$, as we have already seen. To see what happens as K gets large, we shall look at the traces of the boundary locus in the respective coordinate hyperplanes. When a or d equals zero, K does not affect the locus and thus it is the same as that of $\mathcal{P}$ in the corresponding hyperplanes. If we let c = 0 and consider the boundary in the a, b, d hyperplane, the effect of K can be seen. The necessary and sufficient conditions for H to be on the boundary of this trace are the following:

$$a > 0 \quad \text{and} \quad Kad - b^2 = 0.$$

It is clear from the equation $Kad-b^2 = 0$ that H need not be positive definite. As K increases, $b^2/(ad)$ must increase. For any (2×2) hermitian H such that the diagonal entries of H are positive we could find an A (2×2 positive stable and diagonal) such that K was large enough to assure that H be in the range of the corresponding Lyapunov mapping (see Lemma 1: (1) $\leftrightarrow$ (2) in [1, p. 3]). A similar locus arises when b is set equal to zero. Thus the range of the Lyapunov mapping when $n = 2$ is a cone which contains the cone $\mathcal{P}$ for all diagonal A and whose size depends on A, when A is diagonal, in the manner just described.

When $n > 2$ it is no longer necessary that the characteristic values be distinct for the matrix to be nonscalar. In general a diagonal positive stable matrix A can be denoted by $\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, where the $\mathrm{Re}(\lambda_i) > 0$ $(i = 1, 2, \cdots, n)$ and the $\lambda_i (i=1, 2, \cdots, n)$ are not necessarily distinct. Now if $H = [h_{ij}]$ $(i, j = 1, 2, \cdots, n)$ is hermitian and $H = AP + PA^*$ for some positive definite $P = [p_{ij}]$ $(i, j = 1, 2, \cdots, n)$, we have

$$\begin{bmatrix} h_{11} h_{12} \cdots h_{1n} \\ h_{21} h_{22} \cdots h_{2n} \\ \vdots \\ h_{n1} h_{n2} \cdots h_{nn} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \cdots 0 \\ 0 & \lambda_2 \cdots 0 \\ \vdots & \vdots \\ 0 & 0 \cdots \lambda_n \end{bmatrix} \begin{bmatrix} p_{11} p_{12} \cdots p_{1n} \\ p_{21} p_{22} \cdots p_{2n} \\ \vdots \\ p_{n1} p_{n2} \cdots p_{nn} \end{bmatrix}$$

$$+ \begin{bmatrix} p_{11} p_{12} \cdots p_{1n} \\ p_{21} p_{22} \cdots p_{2n} \\ \vdots \\ p_{n1} p_{n2} \cdots p_{nn} \end{bmatrix} \begin{bmatrix} \bar{\lambda}_1 & 0 \cdots 0 \\ 0 & \bar{\lambda}_2 \cdots 0 \\ \vdots & \vdots \\ 0 & 0 \cdots \bar{\lambda}_n \end{bmatrix}$$

$$= \begin{bmatrix} (\lambda_1+\bar{\lambda}_1)p_{11} & (\lambda_1+\bar{\lambda}_2)p_{12} & \cdots & (\lambda_1+\bar{\lambda}_n)p_{1n} \\ (\lambda_2+\bar{\lambda}_1)p_{21} & (\lambda_2+\bar{\lambda}_2)p_{22} & \cdots & (\lambda_2+\bar{\lambda}_n)p_{2n} \\ & & & \\ (\lambda_n+\bar{\lambda}_1)p_{n1} & (\lambda_n+\bar{\lambda}_2)p_{n2} & \cdots & (\lambda_n+\bar{\lambda}_n)p_{nn} \end{bmatrix},$$

i.e.,

$$h_{ij} = (\lambda_i+\bar{\lambda}_j)p_{ij} \qquad (i, j = 1, 2, \cdots, n) .$$

Hence the (i, j) entry of the positive definite matrix P is given by

$$p_{ij} = (\lambda_i + \overline{\lambda}_j)^{-1} h_{ij} \quad (i, j = 1, 2, \cdots, n).$$

One of the well known conditions for positive definiteness applied to P would give a system of requirements that H must satisfy to be in the range of the Lyapunov mapping.

To get more information we now change the notation as follows: let λ_i $(i = 1, 2, \cdots, k)$ be the distinct characteristic values of A (so $k \leq n$). Then A is similar to a block diagonal matrix, each diagonal block of which is a scalar matrix, with distinct blocks corresponding to distinct scalars. Thus we may assume A itself is in this form. Let the i^{th} diagonal block of A have order m_i $(i = 1, 2, \cdots, k,$ where $m_i \geq m_j$ for $i < j)$ and let P and H be partitioned into blocks P_{ij} and H_{ij} $(i, j = 1, 2, \cdots, k)$ respectively, where P_{ij} and H_{ij} are each of dimensions $m_i \times m_j$. We find that H is in the cone $[A]\mathcal{P}$ if and only if

$$H_{ij} = (\lambda_i + \overline{\lambda}_j) P_{ij} \quad (i, j = 1, 2, \cdots, k)$$

for some positive definite P. Note that the $m_1 \times m_1$ leading principal submatrix

$$H_{11} = (\lambda_1 + \overline{\lambda}_1) P_{11}$$

is positive definite, since $\lambda_1 + \overline{\lambda}_1 > 0$ and P_{11} is the $m \times m$ leading principal submatrix of a positive definite matrix. It is easily verified that thus H has at least m_1 positive characteristic values. We recall that m_1 is the multiplicity of λ_1 and that $m_1 \geq m_i$ $(i = 1, 2, \cdots, k)$. This result is a special case of a theorem of P. Stein and A. M Pfeffer.[2/] That theorem is given next without proof in the present author's wording.

Theorem 5. Let A be an $n \times n$ positive stable matrix with m equal to the maximum number of Jordan blocks associated with one characteristic value. If P is positive definite, then

$$H = AP + PA^*$$

has at least m positive roots; further, for every hermitian H with at least m positive characteristic values there exists a B similar to A and a positive definite P such that

$$H = BP + PB^* .$$

2/ Mentioned by O. Taussky in an address at the Argonne Matrix Conference March 30 - 31, 1967; to appear soon in the Bulletin of the International Computation Centre.

IV. $[A]\mathcal{P}$ FOR A GENERAL 2×2 A

In this section we use Theorem 4 to characterize the cone $[A]\mathcal{P}$, when A is a 2×2 general positive stable matrix. In this case we hoped that $[A]^{-1}H$ might be familiar and thus we might be able to generalize our results to larger n $(n \geq 3)$. We were unable to see any such generalization of the results obtained, and thus we feel that Theorem 4 is limited as a practical method of characterizing the cone $[A]\mathcal{P}$ in the general $n \times n$ case if $n \geq 3$.

We now proceed to find the inverse of a 4×4 matrix associated with $[A]$, where A is a general 2×2 positive stable matrix. Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

be a positive stable matrix. The matrix of $[A]$ with respect to the basis $\{e_i \otimes f_j : i, j = 1, 2, \cdots, n\}$ (ordered standard lexicographically) then will be

$$A \otimes I + I \otimes \bar{A} = \begin{bmatrix} a+\bar{a} & \bar{b} & b & 0 \\ \bar{c} & a+\bar{d} & 0 & b \\ c & 0 & d+\bar{a} & \bar{b} \\ 0 & c & \bar{c} & d+\bar{d} \end{bmatrix} .$$

We will use the adjoint method to find the inverse of $A \otimes I + I \otimes \bar{A}$. The adjoint matrix is defined to be the transpose of the cofactor matrix, and thus

$$[A \otimes I + I \otimes \bar{A}]^{-1} = [\det(A \otimes I + I \otimes \bar{A})]^{-1} \operatorname{adj}(A \otimes I + I \otimes \bar{A})$$

The factor $ad-bc = \det A$ occurs in many of the components of the adjoint matrix; therefore, we will use the substitution

$$\Delta = ad-bc.$$

We also find that expressions of the form $\Delta - \bar{\Delta}$ and $\Delta(a+\bar{d}) + \bar{\Delta}(d+\bar{a})$ occur frequently. Thus a substitution of the form $A \to A'$ (not the transpose of A) such that $\Delta' = \det A'$ is real and such that $A' \otimes I + I \otimes \bar{A}' = A \otimes I + I \otimes \bar{A}$, would also simplify the adjoint considerably. We note that subtracting the same imaginary number from a and d does not change $A \otimes I + I \otimes \bar{A}$, so we let $a' = a - \varepsilon$ and $d' = d - \varepsilon$, where ε is pure imaginary. Then set

$$A' = \begin{bmatrix} a' & b \\ c & d' \end{bmatrix} ,$$

so that $\Delta' = a'd' - bc$. Before solving the equation $\Delta' = \bar{\Delta}'$ for ε, we let

$$\tau = a+d + \overline{a+d} = \operatorname{tr} A + \operatorname{tr} \bar{A} = \lambda + \mu + \bar{\lambda} + \bar{\mu} > 0 ,$$

where λ and μ are the characteristic values of A. Now, since ε is pure imaginary we find by routine calculation that $\Delta' = \overline{\Delta}'$ only if

$$\varepsilon = \frac{\Delta - \overline{\Delta}}{\tau}$$

(which is obviously pure imaginary). Conversely, if

$$\varepsilon = \frac{\Delta - \overline{\Delta}}{\tau}$$

then the same calculation shows that $\Delta' = \overline{\Delta}'$. Using this value for ε, we see that indeed

$$A' \otimes I + I \otimes \overline{A}' = A \otimes I + I \otimes \overline{A}$$

and hence that $[A'] = [A]$. We will now determine the sixteen components of the adjoint of $A \otimes I + I \otimes \overline{A}$ $(= A' \otimes I + I \otimes \overline{A}')$ and denote them a_{ij} $(i, j = 1, 2, 3, 4)$. (During these calculations we repeatedly use the fact that $\overline{\Delta}' = \Delta'$.) To simplify the notation during the calculation we will drop primes until otherwise specified.

$$\begin{aligned} a_{11} &= (a+\overline{d})(d+\overline{a})(d+\overline{d}) - \overline{b}\overline{c}(a+\overline{d}) - cb(d+\overline{a}) \\ &= (\overline{a}\overline{d}-\overline{b}\overline{c})(a+\overline{d}) + (ad-cb)(d+\overline{a}) + d\overline{d}(\overline{a}+d + a+\overline{d}) \\ &= \overline{\Delta}(a+\overline{d}) + \Delta(d+\overline{a}) + d\overline{d}(a+d + \overline{a+d}) \\ &= (\Delta + d\overline{d})\tau , \end{aligned}$$

$$\alpha_{12} = -\{\bar{b}[(d+\bar{a})(d+\bar{d}) - \bar{c}\bar{b}] + cb\bar{b}\}$$

$$= -\bar{b}[(d^2 + \bar{a}d + \bar{a}\bar{d} + d\bar{d} - \bar{c}\bar{b} + cb) + ad - ad]$$

$$= -\bar{b}[\bar{\Delta} - \Delta + d(d+\bar{a} + a+\bar{d})]$$

$$= -\bar{b}d\tau ,$$

$$\alpha_{13} = \bar{b}(-\bar{c}b) - (a+\bar{d})\, b(d+\bar{d}) + cb^2$$

$$= b[-c\bar{b} - (ad + a\bar{d} + d\bar{d} + \bar{d}^2 + cb) + \bar{a}\bar{d} - \bar{a}\bar{d}]$$

$$= b[-\bar{d}(\bar{a}+d + a+\bar{d}) - \Delta + \Delta]$$

$$= -b\bar{d}\,\tau ,$$

$$\alpha_{14} = -[-b\bar{b}(d+\bar{a}) - (a+\bar{d})\, b\bar{b}]$$

$$= b\bar{b}\,\tau .$$

The remaining components are found in analogous fashion. They simplify to the following :

$$\alpha_{21} = -\bar{c}d\tau ,$$

$$\alpha_{22} = (\Delta + \bar{a}d)\tau ,$$

$$\alpha_{23} = b\bar{c}\tau ,$$

$$\alpha_{24} = -\bar{a}b\tau ,$$

$\alpha_{31} = -c\bar{d}\tau,$

$\alpha_{32} = c\bar{b}\tau,$

$\alpha_{33} = (a\bar{d}+\Delta)\tau,$

$\alpha_{34} = -a\bar{b}\tau,$

$\alpha_{41} = c\bar{c}\tau,$

$\alpha_{42} = -c\bar{a}\tau,$

$\alpha_{43} = -a\bar{c}\tau,$

$\alpha_{44} = (\Delta + a\bar{a})\tau .$

If we denote the determinant of $A \otimes I + I \otimes \bar{A}$ by D, we have

$$D = (\lambda + \bar{\lambda})(\lambda + \bar{\mu})(\mu + \bar{\lambda})(\mu + \bar{\mu}). \tag{4.1}$$

We can see from (4.1) that D is positive since $\mathrm{Re}(\lambda) > 0$ and $\mathrm{Re}(\mu) > 0$. Thus the inverse of the matrix $A \otimes I + I \otimes \bar{A}$ is

$$\frac{\tau}{D} \begin{bmatrix} \Delta + d\bar{d} & -b\bar{d} & -d\bar{b} & b\bar{b} \\ -c\bar{d} & \Delta + a\bar{d} & b\bar{c} & -a\bar{b} \\ -c\bar{d} & c\bar{b} & \Delta + a\bar{d} & -a\bar{b} \\ c\bar{c} & -c\bar{a} & -a\bar{c} & \Delta + a\bar{a} \end{bmatrix},$$

and we can now find $[A]^{-1}H$. To do this we order H (standard) lexicographically, getting a 4×1 column vector which we denote $\vec{H}$. We can now perform the operation $(A \otimes I + I \otimes \bar{A})^{-1} \vec{H}$ and then rearrange the result in the usual 2×2 matrix form. $[A]^{-1}H$ is given by the following, when $H = \begin{bmatrix} w & u \\ \bar{u} & z \end{bmatrix}$

$$\frac{T}{D} \begin{bmatrix} w(\Delta+d\bar{d}) - u\bar{b}d - \bar{u}\bar{d}b + zb\bar{b} & u(\Delta+a\bar{d}) - w\bar{c}d - za\bar{b} + \bar{u}\bar{c}b \\ \bar{u}(\Delta+\bar{a}d) - wc\bar{d} - z\bar{a}b + uc\bar{b} & z(\Delta+a\bar{a}) - uc\bar{a} - \bar{u}\bar{c}a + wc\bar{c} \end{bmatrix} .$$

It turns out that $[A]^{-1}H$ can be written in the form

$$[A]^{-1}H = \frac{T}{D} (\Delta H + \tilde{A} H \tilde{A}^{*}), \qquad (4.2)$$

where $\tilde{A}$ is the transpose of the cofactor matrix of A. We note that

$$(\tilde{A})^{*} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}^{*} = \begin{bmatrix} \bar{d} & -\bar{c} \\ -\bar{b} & \bar{a} \end{bmatrix} = (\tilde{A^{*}}),$$

and thus there should be no confusion in the expression (4.2). At this point we restore the primes and consider the conditions H must satisfy to be in the cone $[A]\mathcal{P}$. (4.2) becomes

$$[A]^{-1}H = \frac{T}{D} (\Delta' H + \tilde{A}' H \tilde{A}'^{*})$$

$$= \frac{T}{D} \{ \Delta' H + (\tilde{A} - \varepsilon I) H (\tilde{A}^{*} + \varepsilon I) \}$$

$$= \frac{T}{D} \{ (\Delta' - \varepsilon^{2}) H + \tilde{A} H \tilde{A}^{*} + \varepsilon \tilde{A} H - \varepsilon H \tilde{A}^{*} \} . \qquad (4.3)$$

Thus H is in the cone $[A]\mathcal{P}$ if and only if (4.3) is positive definite.

V. CONDITIONS ON H AND A EQUIVALENT TO $H \in [A]\mathcal{P}$

Using the equation $H = AP + PA^*$ or the related mapping is not the only method one might consider to characterize the cone $[A]\mathcal{P}$. In this section (Theorem 6) we give four equivalent methods one might use to characterize the cone. Some motivation for these is as follows:

The integral formula in (2) (in Theorem 6) is essentially given by Bellman [2, Theorem 6, p. 175], where he assumes the existence of the integral, but later fails to discuss that existence as he promises.

The equation in (3) is one used by Stein [11] and by Taussky [13] in a manner very close to that used here.

(4) was motivated by an exercise in Bellman [2, p. 99, Ex. 7]. Our Lemma 5 is a corrected and simplified version of that exercise, and characterizes the cone $\mathcal{P}$ by means of its dual cone ($\mathcal{P}'$), the set of all nonzero nonnegative definite matrices B. The equivalence of (1) and (4) (in Theorem 6) similarly characterizes $[A]\mathcal{P}$ by means of its dual cone ($[A^*]^{-1}\mathcal{P}'$).

In (5) we simply observe that $[A^{-1}]\mathcal{P} = [A]\mathcal{P}$.

In this section we assume A and H are given. We now state and prove Theorem 5.

Theorem 6. Let H be an $n \times n$ hermitian matrix and A be an $n \times n$ positive stable matrix. Then the following five statements are equivalent:

(1) $H = AP + PA^*$ for some positive definite P;

(2) $\int_0^\infty e^{-At} H e^{-A^* t} dt$ is positive definite;

(3) $H = P - CPC^*$ for some positive definite P, where $C = (A+I)^{-1}(A-I)$;

(4) $\operatorname{tr}(KH) > 0$ for every nonzero hermitian K such that $A^*K + KA$ is nonnegative definite;

(5) $H = A^{-1}P + P(A^{-1})^*$ for some positive definite P.

Proof: We will show that (1) is equivalent to each of the other four statements. To show that (1) and (2) are equivalent we will show that the integral in (2) is $[A]^{-1}H$, and thus equivalence follows from Theorem 4. To verify this let

$$P = \int_0^\infty e^{-At} H e^{-A^* t} dt, \qquad (5.1)$$

where the exponential function of a matrix is given by the usual power series, which always converges [2, Theorem 2, p. 166]. Let S be a nonsingular matrix such that $A_1 = S^{-1}AS$ is the direct sum of Jordan blocks of the form

$$\begin{bmatrix} \lambda_i & 0 & 0 & \cdots & 0 \\ 1 & \lambda_i & & & \\ 0 & 1 & \lambda_i & & \\ \cdot & \cdot & \cdot & & \\ \cdot & & \cdot & \cdot & \\ \cdot & & & \cdot & \cdot \\ 0 & & & 1 & \lambda_i \end{bmatrix} = \lambda I + R_i .$$

Then to show that (5. 1) exists it is sufficient to show that the integral

$$P_1 = \int_0^\infty \exp\{-A_1 t\} H_1 \exp\{-A_1^* t\}\, dt$$

exists, where $H_1 = S^{-1}HS^{-1^*}$, since $\exp\{-(SAS^{-1})\} = S(\exp\{-A_1\})S^{-1}$ (c.f. [4, Theorem 10. 2, p. 196]) and hence one easily sees that the given integral equals

$$S\left(\int_0^\infty \exp\{-A_1 t\} H_1 \exp\{-A_1^* t\} dt\right) S^*,$$

i.e., $P = SP_1S^*$, and that $H = SH_1S^*$. Let H_1 be partitioned in such a way as to make the block multiplication $\exp\{-A_1 t\} H_1 \exp\{-A_1^* t\}$ defined. Then the (i, j) block of P_1

will be given by

$$(P_1)_{ij} = \int_0^\infty \exp\{-(\lambda_i I + R_i)t\}(H_1)_{ij}\exp\{-(\bar{\lambda}_j I + R_j^*)t\}\,dt$$

$$= \int_0^\infty \exp\{-(\lambda_i + \bar{\lambda}_j)t\}\exp\{-R_i t\}(H_1)_{ij}\exp\{-R_j^* t\}\,dt \quad .$$

Note that R_i and R_j^* are nilpotent, so the exponentials involving them are polynomials in t. Consequently, the existence of the blocks of P_1 and thus the existence of P follow from the convergence of integrals of the form

$$\int_0^\infty t^k \exp\{-(\lambda_i + \bar{\lambda}_j)t\}\,dt.$$

Churchill in [3, Theorem 1, p. 171] proves that this integral is absolutely convergent for $k \geq 0$ and for $\mathrm{Re}(\lambda_i + \bar{\lambda}_j) > 0$. Since we know that the stability of A guarantees the uniqueness of P (Lemma 3), all that remains is to show that the integral formula does in fact yield the solution.

Calculating, we have

$$AP + PA^* = \int_0^\infty (Ae^{-At}He^{-A^*t} + e^{-At}He^{-A^*t}A^*)dt$$

$$= \int_0^\infty - \frac{d}{dt}(e^{-At}He^{-A^*t})\,dt$$

$$= \lim_{T \to \infty} -[e^{-At}He^{-A^*t}]_0^T$$

$$= H .$$

We used the fact that

$$e^{-AT} \to 0 \quad \text{as} \quad T \to \infty,$$

which is verified in [2, p. 241-242].

We next show (1) is equivalent to (3) using the equivalent transformations

$$C = (A + I)^{-1}(A-I)$$

and

$$A = (I-C)^{-1}(I+C) .$$

To show that (1) implies (3) we need only substitute for A given in terms of C by the above transformation, and we get

$$
\begin{aligned}
H &= (I-C)^{-1}(I+C)P + P(I+C^*)(I-C^*)^{-1} \\
&= (I-C)^{-1}[(I+C)P(I-C^*) + (I-C)P(I+C^*)](I-C^*)^{-1} \\
&= (I-C)^{-1}[2P - 2CPC^*](I-C^*)^{-1} \\
&= 2(I-C)^{-1}P(I-C^*)^{-1} - C\,2(I-C)^{-1}P(I-C^*)^{-1}C^* \\
&= P_1 - CP_1C^*,
\end{aligned}
$$

where $P_1 = 2(I-C)^{-1}P[(I-C)^{-1}]^*$ and thus is positive definite. We used the fact that C commutes with $(I-C)^{-1}$. To show that (3) implies (1) we assume $H = P-CPC^*$ for some positive definite P and make the substitution given above. This yields

$$
\begin{aligned}
H &= P - (A+I)^{-1}(A-I)P(A^*-I)(A^*+I)^{-1} \\
&= (A+I)^{-1}[(A+I)P(A^*+I) - (A-I)P(A^*-I)](A^*+I)^{-1} \\
&= (A+I)^{-1}[2AP + 2PA^*](A^*+I)^{-1} \\
&= AP_1 + P_1A^*,
\end{aligned}
$$

where $P_1 = 2(A+I)^{-1}P[(A+I)^{-1}]^*$ and is thus positive definite.

Before we establish the equivalence of parts (1) and (4), we state and prove an essential lemma.

Lemma 5. Let P be an $n \times n$ hermitian matrix. Then P is positive definite if and only if $\operatorname{tr}(PB) > 0$ for every nonzero nonnegative definite (hermitian) B.

Proof: We first note that for nonsingular T, $tr[(T^*PT)B] = tr[P(TBT^*)]$. Thus in proving the lemma we need only consider diagonal P, since every hermitian matrix is conjunctive with a diagonal matrix, and a matrix is nonzero nonnegative definite if and only if every matrix conjunctive with it is nonzero nonnegative definite. Let $P = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$ be positive definite, and $B = [b_{ij}]$ $(i, j, = 1, 2, \cdots, n)$ be nonzero nonnegative definite. Then $\lambda_i > 0$ $(i = 1, 2, \cdots, n)$, and $b_{ii} \geq 0$ $(i = 1, 2, \cdots, n)$ with $b_{ii} > 0$ for at least one i. Thus

$$tr(PB) = \sum_{i=1}^{n} \lambda_i b_{ii} > 0$$

for any nonzero nonnegative definite B. Now assume that $tr(PB) > 0$ for every nonzero nonnegative definite B. Let $B_k = b_{ij,k}$ $(i, j, k = 1, 2, \cdots, n)$ be the n matrices with elements

$$b_{ij,k} = \begin{cases} 1 & \text{if } i = j = k \\ 0 & \text{otherwise} \end{cases}$$

Then the B_k are nonzero nonnegative definite and

$$tr(PB_k) = \lambda_k \quad (k = 1, 2, \cdots, n) .$$

Therefore $\lambda_k > 0$ for $k = 1, 2, \cdots, n$, and it follows that P is positive definite.

We are now ready to show that (1) implies (4). To do this we must show that $tr(KH) > 0$ for every nonzero hermitian K such that $A^*K + KA$ is nonnegative definite, whenever $H = AP+PA^*$ for some positive definite P. Note that A^*K+KA is $[A^*]K$ and thus is nonzero for all nonzero K. Suppose that $H = AP + PA^*$ for some positive definite P. Then

$$
\begin{aligned}
tr(KH) &= tr\,[K(AP + PA^*)] \\
&= tr\,(KAP + KPA^*) \\
&= tr(KAP) + tr(KPA^*) \\
&= tr\,(KAP) + tr\,(A^*KP) \\
&= tr[(A^*K + KA)P] \\
&> 0 \qquad \text{(Lemma 5)}
\end{aligned}
$$

whenever $A^*K + KA$ is nonzero nonnegative definite. To show that (4) implies (1), let $tr(KH) > 0$ for every nonzero hermitian K such that $A^*K + KA$ is nonnegative definite. Let $P = [A]^{-1}H$. Then $H = [A]P$, and so

$$\begin{aligned} tr(KH) &= tr[(AP + PA^*)K] \\ &= tr(APK) + tr(PA^*K) \\ &= tr(PKA) + tr(PA^*K) \\ &= tr[P(KA + A^*K)] \\ &= tr[P(A^*K + KA)] . \end{aligned}$$

For every nonzero nonnegative definite B, there exists a (nonzero hermitian) K, namely $K = [A^*]^{-1}B$, such that $A^*K + KA = [A^*]K = B$, and hence such that

$$tr(PB) = tr[P(A^*K + KA)] = tr(KH) > 0 .$$

It follows from Lemma 5 that P is positive definite, which completes this part of the proof.

To show (1) implies (5), let (1) be given. Then

$$\begin{aligned} H &= AP + PA^* \\ &= (A^{-1}A)PA^* + AP(A^*A^{*-1}) \\ &= A^{-1}P_1 + P_1(A^{-1})^* , \end{aligned}$$

where $P_1 = APA^*$ and is therefore positive definite. Similarly, if we assume (5) is given, we have

$$H = A^{-1}P + P(A^{-1})^*$$

$$= A^{-1}P[(A^*)^{-1}A^*] + (AA^{-1})P(A^{-1})^*$$

$$= AP_1 + P_1A^* ,$$

where $P_1 = A^{-1}P(A^{-1})^*$ and is thus positive definite. To show that (5) implies (1) we could have applied the part already proven ((1) implies (5)) to the positive stable matrix A^{-1}. This completes the proof of Theorem 5.

BIBLIOGRAPHY

1. Ballantine, C.S. Associate Professor, Department of Mathematics. A note on the matrix equation $H = AP + PA^*$. Corvallis, Oregon State University, 1967. 16 p. (Unpublished typescript)

2. Bellman, Richard. Introduction to matrix analysis. New York, McGraw-Hill, 1960. 328 p.

3. Churchill, Ruel V. Operational mathematics. 2d ed. New York, McGraw-Hill, 1958. 337 p.

4. Finkbeiner, Daniel T. Introduction to matrices and linear transformations. San Francisco, W. H. Freeman, 1960. 246 p.

5. Givens, Wallace. Elementary divisors and some properties of the Lyapunov mapping $X \rightarrow AX + XA^*$. 1961. 80 p. (U.S. Atomic Energy Commission. ANL-6456)

6. Hoffman, Kenneth and Ray Kunze. Linear algebra. 2d ed. Englewood Cliffs, Prentice-Hall, 1964. 322 p.

7. Lang, Serge. Linear algebra. Reading, Addison-Wesley, 1966. 294 p.

8. Lyapunov, A. M. Problème général de la stabilité du mouvement. Princeton, N.J., Princeton University, 1947. 270 p. (Annals of Mathematics Studies, no. 17; Reprint of French translation of Russian original)

9. MacDuffee, C.C. The theory of matrices. Berlin, Springer, 1933. 110 p. (Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 2, no. 5)

10. Ostrowski, Alexander and Hans Schneider. Some theorems on the inertia of general matrices. Journal of Mathematical Analysis and Applications 4: 72-84. 1962.

11. Stein, P. On the ranges of two functions of positive definite matrices. Journal of Algebra 2: 350-353. 1965.

12. Taussky, Olga. A generalization of a theorem of Lyapunov. Journal of the Society of Industrial and Applied Mathematics 9:640-643. 1961.

13. ________________ Matrices C with $C^n \to 0$. Journal of Algebra 1:5-10. 1964.

14. ________________ A remark on a theorem of Lyapunov. Journal of Mathematical Analysis and Applications 2:105-107. 1961.

15. ________________ Research problems: matrix theory. Bulletin of the American Mathematical Society 71:711. 1965.

16. Taussky, Olga and H. Wielandt. On the matrix function $AX+XA'$. Archive for Rational Mechanics and Analysis 9:93-96. 1962.