Czechoslovak Mathematical Journal

Bryan Cain; Daniel Hershkowitz; Hans Schneider

Theorems of the alternative for cones and Lyapunov regularity of matrices

Czechoslovak Mathematical Journal, Vol. 47 (1997), No. 3, 487–499

Persistent URL: http://dml.cz/dmlcz/127372

Czechoslovak Mathematical Journal, 47 (122) (1997), Praha

THEOREMS OF THE ALTERNATIVE FOR CONES AND LYAPUNOV REGULARITY OF MATRICES

BRYAN CAIN, Ames, DANIEL HERSHKOWITZ, Haifa, HANS SCHNEIDER,[1] Madison

(Received January 9, 1995)

Dedicated to Marvin Marcus on the occasion of his retirement

Abstract. Standard facts about separating linear functionals will be used to determine how two cones C and D and their duals C^* and D^* may overlap. When $T\colon V \to W$ is linear and $K \subset V$ and $D \subset W$ are cones, these results will be applied to $C = T(K)$ and D, giving a unified treatment of several theorems of the alternate which explain when C contains an interior point of D. The case when $V = W$ is the space H of $n \times n$ Hermitian matrices, D is the $n \times n$ positive semidefinite matrices, and $T(X) = AX + X^*A$ yields new and known results about the existence of block diagonal X's satisfying the Lyapunov condition: $T(X)$ is an interior point of D. For the same V, W and D, $T(X) = X - B^*XB$ will be studied for certain cones K of entry-wise nonnegative X's.

1. INTRODUCTION

This article has two main objectives. One is to show that a variety of equivalences and theorems stating alternatives may be viewed as corollaries of a general theorem describing how cones and their duals may overlap in abstract, and topological, real vector spaces. The other is to illustrate the scope of this point of view by deriving new results. From a wealth of possibile new results we have selected examples of, or very close to, traditional interests, e.g. characterizing the $n \times n$ complex matrices A (resp. C) such that the Lyapunov (resp. Stein) condition "$L_A(X) = AX + XA^*$

This research was supported by a joint grant No. 90-00434 from the United States—Israel Binational Science Foundation. Jerusalem, Israel.

[1] The research of this author was supported in part by NSF grants DMS-9123318, DMS-8901445, and EMS-878971.

(resp. $S_C(X) = X - C^*XC$) is positive definite" has a Hermitian solution X which lies in a cone of block diagonal matrices or of entry-wise nonnegative matrices or both. Our results include new theorems "of the alternative" and new equivalences. ($P \Leftrightarrow Q$) may always be stated as the alternative: P or not$Q$, but not both.) Here is some background on the results relating to L_A.

A complex square matrix A is said to be (positive) stable if its spectrum lies in the open right half-plane. Lyapunov, while studying the asymptotic stability of solutions of differential systems, proved a theorem in 1892 which, restated for matrices, asserts that A is stable if and only if there exists a positive definite Hermitian matrix G such that the matrix $AG + GA^*$ is positive definite.

Lyapunov's theorem has motivated the study of positive definite Hermitian matrices G such that $AG + GA^*$ is positive definite. Such matrices G are called *stability factors* for A. Stability factors have been studied by Carlson and Schneider [8], by Hershkowitz and Schneider [12], and by others. An interesting special case, which plays an important role in various applications, is the case of matrices A, so called *Lyapunov diagonally stable* matrices, for which there exist diagonal stability factors. Unlike stability, Lyapunov diagonal stability is not merely a spectral property, and in general it is hard to characterize. Recently, Carlson, Hershkowitz and Shasha [7] unified the study of stability and Lyapunov diagonal stability, by characterizing those matrices for which there exist stability factors with given block diagonal structure.

Another related topic is the research on matrices A for which there exists a (not necessarily positive definite) Hermitian matrix G such that the matrix $AG + GA^*$ is positive definite. We call such matrices A *Lyapunov regular*, and we call the corresponding matrix G a *regularity factor* for A. Ostrowski and Schneider showed [16] that a matrix is Lyapunov regular if and only if it has no purely imaginary eigenvalue.

Some studies of matrix stability use theorems of the alternative for cones, e.g. [1] and [7]. Our paper develops and applies theorems of the alternative to obtain characterizations of classes of matrices which have stability factors or regularity factors with given block structure.

Section 2 is devoted to general results, valid for real vector spaces and linear maps or real topological vector spaces and continuous linear maps. It begins discussing how two convex sets C and D and their duals may be situated in space and then specializes to the case where C and D are cones and then further to the case where C is the range or kernel of a linear transformation. Our use of separation theorems connects our results to ones in [17]. We generalize here a result in [2] and a theorem of the alternative in [9].

The rest of the paper (Lemma 5.6 being a noteworthy exception) mainly specializes and applies the results of Section 2 to obtain the results about the L_A and S_C

mentioned above. Some information about the positive semidefinite members of $\mathrm{Ker}(L_A)$ (and of $\mathrm{Ker}(S_C)$) is also obtained, cf. Theorem (4.8) and Lemma (6.2).

2. General results about cones

We shall use the same notations when we consider real vector spaces and linear maps as we do for real topological vector spaces and continuous linear maps. So the meaning of a symbol may depend on the underlying category. Words and remarks in brackets usually pertain to the topological case.

Let V denote a [topological] vector space over $\mathbb{R}$, the real field. Then V' denotes its dual, that is all [continuous] linear maps $f: V \to \mathbb{R}$. If $S \subset V$ then

- S° denotes the radial kernel of S, i.e. $x \in S^\circ$ means: There is a positive function $\delta_x: V \to \mathbb{R}$ such that $x + t(w - x) \in S$ whenever $w \in V$ and $0 \leqslant t < \delta_x(w)$. (cf. page 14 of [13]) [however we define S° to be the topological interior of S, if V has a topology];
- $S^- = V \setminus (V \setminus S)^\circ$;
- $S^* = \{f \in V' \colon f(x) \geqslant 0 \text{ for all } x \text{ in } S\}$;
- $S^\perp = \{f \in V' \colon f(x) = 0 \text{ for all } x \text{ in } S\}$;
- $S^* = S^\perp$, if $S = -S$, e.g. if S is a subspace;
- S is a (convex) cone if and only if $S \neq \emptyset$ and $x+y$ and $tx \in S$ whenever $x, y \in S$ and $t \geqslant 0$;
- $\mathrm{cone}(S)$ is the smallest cone in V, which contains S.

Let W too denote a [topological] vector space over $\mathbb{R}$. If $T: V \to W$ is [continuous and] linear, $T^*: W' \to V'$ is defined by $(T^*g)(x) = g(Tx)$ for every g in W' and x in V. If $S \subset V$ then it is simple to verify that

$$T(S)^* = T^{*-1}(S^*) \quad \text{(i.e. } \{g \in W' : T^*g \in S^*\}).$$

Our first result Theorem 2.1 is pivotal. Our other results are mostly corollaries of it or lemmas aimed at proving or explaining its corollaries. We state it quite generally.

(2.1) Theorem. *Let C and D be convex subsets of V, a real [topological] vector space. If $0 \in C \cap D$ the following are equivalent.*

(i) $C \cap D^\circ \neq \emptyset$.

(ii) $-C^* \cap D^* = \{0\}$ *and* $D^\circ \neq \emptyset$.

P r o o f. (i) $\Rightarrow$ (ii): If $x \in C \cap D^\circ$ and $f \in -C^* \cap D^*$ then $f(x) \leqslant 0 \leqslant f(x)$, so $f(x) = 0$. Were f nonzero at w, the line through w and x would intersect D in a line segment L having x in its interior because $x \in D^\circ$. Then $f(z)$ would change sign as $z \in L$ passed through x, which contradicts $f|_D \geqslant 0$.

(ii) $\Rightarrow$ (i): If not, by 3.8 page 22 [14.2 page 118] of [13] there is a $0 \neq f \in V'$ such that $\sup f(C) \leqslant \inf f(D)$. Since $0 \in C \cap D$, both numbers are zero. So $0 \neq f \in -C^* \cap D^* = \{0\}$. □

Remark. In the case of certain spaces Theorem (2.1) can also be derived from a lemma due to Dubovickii and Miljutin [9], see also page 37 of [11] or page 411 of [20]. In finite dimensional Euclidean space our result is close to Theorem 11.3 on page 97 of [17].

Since S^* is convex when S is any subset of a real vector space, Theorem 1 yields:

(2.2) Corollary. *When C and D are subsets of V, a real [topological] vector space, the following are equivalent.*

(i) $C^* \cap D^{*o} \neq \emptyset$.

(ii) $-C^{**} \cap D^{**} = \{0\}$ *and* $D^{*o} \neq \emptyset$.

The natural imbedding $i\colon V \to V''$ permits us to compare S^{**} with S (or, more precisely, $i(S)$).

(2.3) Lemma. *Let $S \subset V$, a real [locally convex topological] vector space. Then $i^{-1}(S^{**}) = \operatorname{cone}(S)^-$.*

Proof. Since S^{**} is a closed cone containing $i(S)$, $L = i^{-1}(S^{**})$ is a closed cone containing $R = \operatorname{cone}(S)^-$. If x is not in R, by 3.9 page 23 [14.3 page 118] of [13] there is an $f \in V'$ such that

$$\inf[f(\operatorname{cone}(S)^-)] > f(x) = i(x)(f).$$

Since $0 \in \operatorname{cone}(S)$, the infimum is 0. Hence $f \in S^*$ and so $i(x)$ is not in S^{**}, i.e. x is not in L. □

(2.4) Remark. Hence when $V'' = V$ (e.g. when $\dim V < \infty$) it is natural to consider the case where C and D are closed cones and $C = C^{**}$ and $D = D^{**}$. If in addition C is a subspace, $C^* = C^\perp$ and $-C = C$, so Corollary (2.2) becomes

(2.5) Corollary. *Let C and D be subsets of a real [topological] vector space $V = V''$. If $C^{**} = C = -C$ and $D^{**} = D$ (C must be a closed subspace and D a closed cone), then $C^\perp \cap D^{*o} \neq \emptyset$ iff $C \cap D = \{0\}$ and $D^{*o} \neq \emptyset$.*

Proof. This is Theorem (2.1) with $C^* = C^\perp$ in place of C and D^* in place of D. □

Corollary (2.5) is a special case of Corollary (2.2) and it contains Corollary 2.6 of [2].

(2.6) Corollary. *Let V and W be real [topological] vector spaces, and let T: $V \to W$ be [continuous and] linear. Let $C \subset V$ and $D \subset W$. Suppose $0 \in D$, D is convex, $W = W''$ and $T(C)^{**} = T(C)^-$. Then $T^{*-1}(C^*) \cap D^\circ \neq \emptyset$ iff $-T(C)^- \cap D^* = \{0\}$ and $D^\circ \neq \emptyset$.*

Proof. This is Theorem (2.1) with $T(C)^* = T^{*-1}(C^*)$ in place of C. □

(2.7) Theorem. *Let V and W be real [topological] vector spaces, and let T: $V \to W$ be [continuous and] linear. Let $C \subset V$ and $D \subset W$ be convex. If $0 \in C \cap D$ and $D^\circ \neq \emptyset$, then* (i)–(iv) *are equivalent,* (o) $\Rightarrow$ (i), *and if $C^\circ \neq \emptyset$ then* (i) $\Rightarrow$ (o).

(o) $T(C^\circ) \cap D^\circ \neq \emptyset$.
(i) $T(C) \cap D^\circ \neq \emptyset$.
(ii) $T(-C)^* \cap D^* = \{0\}$.
(iii) $-C^* \cap T^*(D^* \setminus \{0\}) = \emptyset$.
(iv) $-C^* \cap T^*(D^*) = \{0\}$ *and* $\mathrm{Ker}(T^*) \cap D^* = \{0\}$.

Proof. (i) $\Leftrightarrow$ (ii): By Theorem 2.1. (ii) $\Leftrightarrow$ (iii): Since $0 \in T(C)^* = T^{*-1}(C^*)$, (ii) is equivalent to $T^{*-1}(-C^*) \cap (D^* \setminus \{0\}) = \emptyset$, which, by routine facts about sets and functions, is equivalent to (iii). (iii) $\Rightarrow$ (iv) is clear, and so is its converse. (o) $\Rightarrow$ (i) is trivial. (i) $\Rightarrow$ (o): Let $w \in C^\circ$ and $x \in C \cap T^{-1}(D^\circ)$. Then for $t > 0$ and small enough $(1-t)x + tw \in C^\circ \cap T^{-1}(D^\circ)$. □

The equivalence of (i) and (iii) maybe stated as a theorem of the alternative "either (i) holds or (iii) fails but not both." In other words:

(2.8) Theorem. *Given the hypotheses of Theorem 2.7 one, but not both, of the following holds*

(a) *There is an $x \in C$ such that $T(x) \in D^\circ$.*
(b) *There is a nonzero $f \in D^*$ such that $-T^*(f) \in C^*$.*

(2.9) Remark. Theorem 2.8 yields many well known theorems of the alternative. For example Theorem 2.10 in [9] is obtained by putting $V = \mathbb{R}^n$, and $W = \mathbb{R}^m$, letting T be an $m \times n$ matrix, and using the nonnegative orthant as a cone.

The well known equation $(\mathrm{Ker}\, T)^\perp = (\mathrm{Range}\, T^*)^-$ which is valid for bounded linear operators on a Hilbert space, takes the (perhaps less well known) form $(\mathrm{Ker}\, T)^\perp = T^*(W')$ (or $(\mathrm{Ker}\, T)^* = (T^*(W'))^-$) when T: $V \to W$ is linear and V and W are vector spaces. Then (2.1) with $C = \mathrm{Ker}(T)$ is:

(2.10) Corollary. *Let $V, W, T\colon V \to W$ come from a category of real vector spaces and linear maps in which* $(\mathrm{Ker}(T))^* = (T^*(W'))^-$. *Let $0 \in D \subset V$ be convex. The following are equivalent.*

(i) $\mathrm{Ker}(T) \cap D^\circ \neq \emptyset$.
(ii) $(T^*(W'))^- \cap D^* = \{0\}$ *and* $D^\circ \neq \emptyset$.

3. Notations, special cones and their duals

We shall need some additional notations:

Let $\alpha = \{\alpha_1, \ldots, \alpha_p\}$ denote a partition of $\langle n \rangle = \{1, \ldots, n\}$ and let $A = (a_{ij})$ be an $n \times n$ matrix. We say that A is α-diagonal if $a_{ij} = 0$ whenever i and j lie in different α_k's. Such an A is permutation similar to a block diagonal matrix. A_{ij} will denote the submatrix of A consisting of the a_{rs} with $r \in \alpha_i$ and $s \in \alpha_j$. Let $H(\alpha)$ denote the α-diagonal members of H, the real vector space of $n \times n$ complex Hermitian matrices endowed with the inner product $\langle X, Y \rangle = \mathrm{trace}(Y^* X)$. Let $E(i,j)$ denote the $n \times n$ matrix with a 1 in the ij-th place and zeroes everywhere else. Set $F(r,s) = E(r,s)$ when $r = s$, $E(r,s) + E(s,r)$ when $r < s$, and $iE(s,r) - iE(r,s)$ when $r > s$. Then the $F(r,s)$ form an orthogonal basis for H.

(3.1) Lemma. *Let $\{F_1, \ldots, F_p\}$ be a partition of $\{F(i,j)\}$. Set $H_i = \mathrm{Span}\, F_i$ (formed with real coefficients), and let $C_i \subset H_i$ be a cone. Then $C = C_1 + \ldots + C_p$ is a cone and $C^* = C_1^* + \ldots + C_p^*$. Both of these sums are orthogonal direct sums.*

Proof. If $f \in H'$, $f = \Sigma f_i$ where $f_i = f|H_i$. Then $f \in C^*$ if and only if each $f_i \in C_i^*$. We omit the rest. □

Let PSD $\subset H$ denote the cone of complex positive semidefinite matrices, NNH $\subset H$, the cone of $n \times n$ Hermitian matrices with nonnegative entries, and PIH $\subset H$ the cone $\mathrm{Span}\{F(r,s)\colon r > s\}$ of $n \times n$ Hermitian matrices with purely imaginary entries. Set $P(\alpha) = \mathrm{PSD}(\alpha) = \mathrm{PSD} \cap H(\alpha)$, $\mathrm{NNH}(\alpha) = \mathrm{NNH} \cap H(\alpha)$, and $\mathrm{PIH}(\alpha) = \mathrm{PIH} \cap H(\alpha)$.

(3.2) Corollary.

$$\begin{aligned}
\mathrm{NNH}^* &= \mathrm{NNH} + \mathrm{PIH}.\\
H(\alpha)^* &= H(\alpha)^\perp = \mathrm{Span}\{F(r,s)\colon r \in \alpha_i \text{ and } s \in \alpha_j \text{ and } i \neq j\}\\
&= \{(G_{ij}) \in H\colon G_{ii} = 0_{k\times k} \text{ where } k \text{ is the number of elements in } \alpha_i\}.\\
P(\alpha)^* &= (H(\alpha) \cap \mathrm{PSD})^* = (H(\alpha)^\perp + \mathrm{PSD})^- = (H(\alpha)^\perp + P(\alpha))^-\\
&= H(\alpha)^\perp + P(\alpha).
\end{aligned}$$

P r o o f. To prove the first and second sentences let each H_i be one of the 1-dimensional subspaces Span$\{F(r,s)\}$. Then C_i is either $\{0\}F(r,s)$, $(0,\infty)F(r,s)$, or $(-\infty,\infty)F(r,s)$. Apply the lemma. The third sentence relies on: $(C \cap D)^* = (C^*+D^*)^-$ whenever C and D are closed cones (cf. page 376 of [3]) and PSD* =PSD. To justify the last equality select anything in the closure and decompose it into a sum of a term M in $H(\alpha)$ and a term in $H(\alpha)^\perp$. Note that then $M \in \mathrm{PSD}^- = \mathrm{PSD}$. □

(3.3) Lemma. $(P(\alpha) \cap \mathrm{NNH})^* = H(\alpha)^\perp + P(\alpha) + \mathrm{NNH}(\alpha) + \mathrm{PIH}(\alpha)$.

P r o o f. Let L [R] denote the left [right] side of the equation we must justify. By [3] and the previous corollary, $L = (P(\alpha)^* + \mathrm{NNH}^*)^- = (H(\alpha)^\perp + P(\alpha) + \mathrm{NNH} + \mathrm{PIH})^- = (H(\alpha)^\perp + P(\alpha) + \mathrm{NNH}(\alpha) + \mathrm{PIH}(\alpha))^- = R^-$. Suppose the sequence $E(k) = A(k) + B(k) + C(k) + D(k) \in R$ converges to E and $A(k) \in H(\alpha)^\perp, B(k) = (b_{ij}(k)) \in P(\alpha), C(k) = (c_{ij}(k)) \in \mathrm{NNH}(\alpha)$, and $D(k) \in \mathrm{PIH}(\alpha)$. Since the sequences $A(k)$ and $F(k) = (f_{ij}(k)) = B(k) + C(k) + D(k)$ lie in orthogonally complementary subspaces they must both converge, say to A and F, respectively. Since $B(k)$ is positive semidefinite and $c_{ii} \geqslant 0$, we have $|b_{rs}(k)|^2 \leqslant [b_{rr}(k)+c_{rr}(k)][b_{ss}(k)+c_{ss}(k)] = f_{rr}(k)f_{ss}(k)$, which is bounded because it converges. Hence $B(k)$ is bounded and so has a convergent subsequence with limit, say B, in $P(\alpha)$. The corresponding subsequence of $F(k)-B(k) = C(k)+D(k)$ must then also converge to $F-B$, and since $C(k)$ and $D(k)$ are in orthogonally complementary spaces, the corresponding subsequences of $C(k)$ and $D(k)$ will both converge, say to $C \in \mathrm{NNH}(\alpha)$ and $D = F-B-C \in \mathrm{PIH}(\alpha)$. Hence $E = A+B+C+D \in R$, so R is a closed cone. □

4. Lyapunov regularity of matrices

(4.1) Definition. Let α be a partition of $\langle n \rangle = \{1,\dots,n\}$ into p nonempty sets and let A be an $n \times n$ complex matrix. Then A is *Lyapunov α-regular* if $L_A(G)$ is positive definite for some $G \in H(\alpha)$. A Lyapunov α-regular A is called *Lyapunov regular* when $p=1$ (then G need not have any zero entries) and *Lyapunov diagonally regular* when $p=n$ (then G must be diagonal).

A characterization of the Lyapunov α-regular matrices A can be obtained from (i) ⇔ (ii) in Theorem 2.7 by setting $V = W = H$, $C = H(\alpha)$, $D = \mathrm{PSD}$, and $T = L_A$. Then C^* is described in Corollary 3.2.

(4.2) Theorem. *Let A be a complex $n \times n$ matrix and α a partition of $\langle n \rangle$ into p nonempty sets. The following are equivalent.*

(i) *A is Lyapunov α-regular.*

(ii) *For every nonzero* $K \in \mathrm{PSD}$ *there exists* $i \in \langle p \rangle$ *such that* $(A^*K + KA)_{ii} \neq 0$.

An equivalent statement of alternative nature is the following.

(4.3) Corollary. *Let* A *be a complex* $n \times n$ *matrix and* α *a partition of* $\langle n \rangle$ *into* p *nonempty sets. Then either* A *is Lyapunov* α*-regular, or there exists a nonzero* $K \in \mathrm{PSD}$ *such that* $(A^*K + KA)_{ii} = 0$ *for all* $i \in \langle p \rangle$, *but not both.*

In the special cases of $\alpha = \{\langle n \rangle\}$ i.e. $p = 1$, we obtain the following characterization for Lyapunov regularity of matrices.

(4.4) Theorem. *Let* A *be a complex* $n \times n$ *matrix. The following are equivalent.*

(i) A *is Lyapunov regular.*

(ii) *For every nonzero* $K \in \mathrm{PSD}$ *we have* $A^*K + KA \neq 0$.

(4.5) Remark. Theorem 4.8, which may be of interest in its own right, will show that (ii) is equivalent to $\mathrm{Spec}(A) \cap i\mathbb{R} = \emptyset$. Thus Ostrowski and Schneider's characterization of Lyapunov regular matrices (stated in the introduction) follows from Theorems 4.4 and 4.8.

(4.6) Remark. A complete description of $\mathrm{Ker}(L_A)$ is obtainable by reducing to the case where A is in Jordan form and solving the resulting block matrix equation $L_A(X) = 0$. Theorem 4.8 can be proven this way, but not as succinctly.

(4.7) Definition. Whenever $Z \subset \mathbb{C}$ and A is a square complex matrix, $g(A; Z)$ will denote the sum of the geometric multiplicities of the eigenvalues of A which lie in Z.

4.8 Theorem. *Let* A *be an* $n \times n$*complex matrix. Set* $g = g(A; i\mathbb{R})$. *Then*

$$\mathrm{rank}(\mathrm{Ker}(L_A) \cap \mathrm{PSD}) = \{0, 1, \ldots, g\}.$$

Proof. If $K \in \mathrm{PSD}$ there is an invertible S such that $L = \mathrm{SKS}^* = \mathrm{Diag}(I_r; 0)$. Set $B = \mathrm{SAS}^{-1} = \begin{pmatrix} C & D \\ E & F \end{pmatrix}$, where C is $r \times r$. If $K \in \mathrm{Ker}(L_A)$ also, then $\begin{pmatrix} 2\,\mathrm{Re}\, C & E^* \\ E & 0 \end{pmatrix} = 2\,\mathrm{Re}(BL) = SL_A(K)S^* = 0$. Thus C is skew Hermitian, so $r = g(C; i\mathbb{R})$, which is at most g because $E = 0$. That is $\mathrm{rank}(K) = r \leqslant g$.

On the other hand, if $0 \leqslant s \leqslant g$ is an integer, there are independent eigenvectors $x_1, \ldots, x_s$ of A having eigenvalues in $i\mathbb{R}$. Then $K = x_1 x_1^* + \ldots + x_s x_s^* \in \mathrm{PSD}$ and $L_A(K) = 0$. Let S be the inverse of a matrix whose first s columns are $x_1, \ldots, x_s$ then $\mathrm{SKS}^* = \mathrm{Diag}(I_s, 0)$. So K has rank s. □

In the special case of $\alpha = \{\{1\}, \ldots, \{n\}\}$ we have the following corollary of Theorem (4.2).

4.9 Theorem. *Let A be a complex $n \times n$ matrix. The following are equivalent.*

(i) *A is Lyapunov diagonally regular.*
(ii) *For every nonzero $K \in$ PSD the matrix $A^*K + KA$ has a nonzero diagonal element.*

Corollary (2.5) yields the following theorem of the alternative.

(4.10) Theorem. *Let A be a complex $n \times n$ matrix and α a partition of $\langle n \rangle$ into p nonempty sets. Then either there exists a positive definite Hermitian $n \times n$ matrix K such that $(A^*K + KA)_{ii} = 0$ for all $i \in \langle p \rangle$, or there exists a $G \in H(\alpha)$ such that $AG + GA^*$ is a nonzero positive semidefinite matrix, but not both.*

Proof. Let $V = H$, $D =$ PSD and $C = L_A(H(\alpha))$. Then the hypotheses of Corollary (2.5) are satisfied, $D^* = D$, and $D^{*o} \neq \emptyset$. Since $C^\perp = {L_{A^*}}^{-1}(H(\alpha)^\perp)$, the first alternative, viz. $H(\alpha)^\perp \cap L_{A^*}(D^\circ) \neq \emptyset$, is equivalent to $C^\perp \cap D^{*o} \neq \emptyset$, and by Corollary (2.5) $C \cap D = \{0\}$. The latter is equivalent to $C \cap (D \setminus \{0\}) = \emptyset$, which is the negation of the second alternative. □

(4.11) Remark. Applying Corollary (2.6) with $V = H$, $D =$ PSD, $C = H(\alpha)$, and $T = L_A^*$ also proves Theorem (4.10).

(4.12) Remark. To see that $A^*K + KA$ cannot be replaced by A^*K or KA in any of the theorems (4.2), (4.3), (4.4), (4.9), and (4.10), set $A = iI$ and note that then $\mathrm{Ker}(L_{A^*}) = H$ and for every nonzero $K \in$ PSD we have $(A^*K)_{ii} \neq 0 \neq (KA)_{ii}$, for some $i \in \langle p \rangle$.

(4.13) Theorem. *Let A be a complex $n \times n$ matrix. The following are equivalent.*

(i) *There exists a positive definite Hermitian K such that $L_A(K) = 0$.*
(ii) *If $G = G^*$ then $L_{A^*}(G)$ is positive semidefinite iff it is 0.*
(iii) *A is similar to a skew Hermitian matrix.*

Proof. (i)$\Leftrightarrow$ (ii) is the special case of Corollary (2.10) with $V = W = H$, $T = L_A$, and $D =$ PSD. (i) $\Leftrightarrow$ (iii) by Theorem (4.8). Note $g(A; i\mathbb{R}) = n$. □

5. Other theorems of the alternative and applications to Lyapunov stability

(5.1) Definition. (see [7]). Let α be a partition of $\langle n \rangle$. A complex $n \times n$ matrix A is said to be *Lyapunov α-stable* if there exists a positive definite Hermitian α-diagonal matrix G such that $L_A(G)$ is positive definite.

The equivalence of (i) and (iii) in the following theorem was proven as Theorem 3.10 in [7].

(5.2) Theorem. *Let A be a complex $n \times n$ matrix and α a partition of $\langle n \rangle$ in p nonempty sets. The following are equivalent.*

(i) *For every nonzero $L \in$ PSD there exists an $i \in \langle p \rangle$, such that $-(A^*L + LA)_{ii}$ is not positive semidefinite.*
(ii) *There exists a $G \in P(\alpha)$ such that $AG + GA^*$ is positive definite.*
(iii) *A is Lyapunov α-stable.*

Proof. Set $V = H(\alpha)$, $W = H$, $C = P(\alpha)$, $D =$ PSD, and $T = L_A$. Then (i), (ii), (iii) are respectively (iii), (i), (o) from Theorem (2.7). □

(5.3) Remark. Actually, if $G \in$ PSD and $L_A(G)$ is positive definite, so is G, because then, for small $t > 0$, $E = G + tI$ and $L_A(E)$ are positive definite. Hence G, A, and E have the same inertia.

The following known characterization of stable matrices is a corollary of Theorem (5.2) for the special case of $\alpha = \{\{1, \ldots, n\}\}$. It already follows from Theorem 3.10 in [7]. It is originally stated as statement (c) of Proposition 5 in [6].

(5.4) Corollary. *Let A be a complex $n \times n$ matrix. Then A is stable if and only if for every nonzero $L \in$ PSD, the matrix $-(A^*L + LA)$ is not positive semidefinite.*

(5.5) Theorem. *Let A be a complex $n \times n$ matrix and α a partition of $\langle n \rangle$ into p nonempty sets. The following are equivalent.*

(i) *The cone $\{L_A(G) \colon G \in P(\alpha)\}^-$ contains no nonzero positive semidefinite matrices.*
(ii) *There exists an $L \in$ PSD such that $-(A^*L + LA)_{ii}$ is positive semidefinite for all $i \in \langle p \rangle$.*

Proof. Set $V = H(\alpha)$, $W = H$, $C = P(\alpha)$, $D =$ PSD, and $T = -L_A$. Apply Corollary (2.6). □

For a related corollary of (2.6), we need the following cone theoretic lemma.

(5.6) Lemma. *Let V and W be real topological vector spaces, $C \subset V$ and $D \subset W$ be cones, and $T: V \to W$ be a continuous linear transformation. Suppose the topology on V is given by a norm $\|\cdot\|$ and $N = \{v \in C : \|v\| \leqslant 1\}$ is compact. Then $T(C)^- \cap D \neq \{0\}$ implies $T(C \setminus \{0\}) \cap D \neq \emptyset$.*

Proof. Let $0 \neq w \in T(C)^- \cap D$. Then there exist $v_n \in C$ such that $\lim T(v_n) = w$ and $\lim \|v_n\| = L \in [0, \infty]$. Since $w \neq 0$, $L \neq 0$. Let v be a limit point of $\{\|v_n\|^{-1} v_n\} \subset N \subset C$. Then $v \neq 0$ and $Tv = 0 \in D$ if $L = \infty$ while $Tv = L^{-1} w \in D$ if $L < \infty$. □

With the additional hypothesis $\operatorname{Ker}(T) \cap C = \{0\}$, the converse of Lemma 5.6 is easy to prove; without it, there are counterexamples such as: C is the first quadrant of $\mathbb{R}^2$, D lies in the third quadrant, and T is the orthogonal projection onto any line contained in quadrants 2 and 4. Or, more simply, let $C \neq \{0\}$, D be any cones in any spaces and let $T|_C = 0$ with $C \neq \{0\}$.

(5.7) Remark. When Corollary 2.6 is combined with Lemma 5.6 we obtain (from the combined hypotheses and $D^\circ \neq \emptyset$): $T^{*-1}(C^*) \cap D^\circ = \emptyset$ implies $-T(C \setminus \{0\}) \cap D^* \neq \emptyset$. Since $C^* \cap T^*(D^\circ) = \emptyset$ implies $T^{*-1}(C^*) \cap D^\circ = \emptyset$, when $V = H(\alpha)$, $W = H$, $C = P(\alpha)$, $D = \mathrm{PSD}$, and $T = -L_A$ we obtain

(5.8) Corollary. *Let A be a complex $n \times n$ matrix and α a partition of $\langle n \rangle$ into p nonempty sets. If there exists no $L \in \mathrm{PSD}^\circ$ such that $-(A^*L + LA)_{ii}$ is positive semidefinite for all $i \in \langle p \rangle$, then there exists a nonzero $H \in P(\alpha)$ such that $AH + HA^*$ is positive semidefinite.*

6. The Stein operator and the cone of nonnegative Hermitian matrices

(6.1) Theorem. *Let $T: H(\alpha) \to H$ be linear. The following are equivalent.*

(o) $T((P(\alpha) \cap \mathrm{NNH})^\circ) \cap \mathrm{PSD}^\circ \neq \emptyset$.

(i) $T(P(\alpha) \cap \mathrm{NNH}) \cap \mathrm{PSD}^\circ \neq \emptyset$.

(ii) $-(H(\alpha)^\perp + P(\alpha) + \mathrm{NNH}(\alpha) + \mathrm{PIH}(\alpha)) \cap T^*(\mathrm{PSD} \setminus \{0\}) = \emptyset$.

(iii) $-(H(\alpha)^\perp + P(\alpha) + \mathrm{NNH}(\alpha) + \mathrm{PIH}(\alpha)) \cap T^*(\mathrm{PSD}) = \{0\}$ *and* $\mathit{Ker}(T^*) \cap \mathrm{PSD} = \{0\}$.

Proof. Let $C = P(\alpha) \cap NNH$ and $D = \mathrm{PSD}$ in Theorem (2.7) and use Lemma (3.3) to find C^*. □

(6.2) Lemma. *Let $S_C(X) = X - C^*XC$ where C is an $n \times n$ complex matrix. Set $g = g(C; \{z \in \mathbb{C}: |z| = 1\})$. Then* $\operatorname{rank}(\operatorname{Ker}(S_C)) \cap \mathrm{PSD}) = \{0, 1, \ldots, g\}$.

Proof. The proof of Theorem (4.8) may be imitated here, or one may use the equivalence described in [19] of features of L_A and S_C. □

(6.3) Corollary. *Let C be an $n \times n$ complex matrix and α be a partition of $\langle n \rangle$. Let $S_C: H(\alpha) \to H$ be defined by $S_C(X) = X - C^*XC$. Then the following are equivalent.*

(o) $S_C((P(\alpha) \cap \mathrm{NNH})^0) \cap \mathrm{PSD}^\circ \neq \emptyset$.
(i) $S_C(P(\alpha) \cap \mathrm{NNH}) \cap \mathrm{PSD}^\circ \neq \emptyset$.
(ii) $-(H(\alpha)^\perp + P(\alpha) + \mathrm{NNH}(\alpha) + \mathrm{PIH}(\alpha)) \cap S_C^*(\mathrm{PSD} \setminus \{0\}) = \emptyset$.
(iii) $-(H(\alpha)^\perp + P(\alpha) + \mathrm{NNH}(\alpha) + \mathrm{PIH}(\alpha)) \cap S_{C^*}(\mathrm{PSD}) = \{0\}$ *and* $|\mu| \neq 1$ *for every* $\mu \in \operatorname{Spec}(C)$.

Proof. Since $(S_C)^* = S_{C^*}$ Theorem (6.1) and the Lemma (6.2) establish the equivalence. □

Remark. In this theorem S_C can be replaced by L_A if the restriction on the spectrum is also changed to: $\operatorname{Re}(\mu) \neq 0$ for every $\mu \in \operatorname{Spec}(A)$.

Acknowledgement. The authors thank Professor Michael Ferris for simplifying part of the proof of Theorem (2.1) and for pointing out the close relation of this theorem to Theorem 11.3 in [17]. They also thank Professor Adi Ben-Israel for drawing their attention to the relation of Theorem (2.1) to the Dubovickii-Miljutin theory, and Dr. Dafna Shasha for providing a reference for Corollary (5.4).

References

[1] *G. P. Barker, A. Berman, and R. J. Plemmons*: Positive diagonal solutions to the Lyapunov equations. Lin. Multilin. Alg. *5* (1978), 249–256.
[2] *G. P. Barker, B. S. Tam, and Norbil Davila*: A geometric Gordan-Stiemke theorem. Lin. Alg. Appl. *61* (1984), 83–89.
[3] *A. Ben-Israel*: Linear equations and inequalities on finite dimensional, real or complex, vector spaces: A unified theory. Math. Anal. Appl. *27* (1969), 367–389.
[4] *A. Berman*: Cones, Matrices and Mathematical Programming. Lecture Notes in Economics and Mathematical Systems, Vol. 79. Springer-Verlag, 1973.
[5] *A. Berman and A. Ben-Israel*: More on linear inequalities with application to matrix theory. J. Math. Anal. Appl. *33* (1971), 482–496.
[6] *A. Berman and R. C. Ward*: ALPS: Classes of stable and semipositive matrices. Lin. Alg. Appl. *21* (1978), 163–174.
[7] *D. H. Carlson, D. Hershkowitz, and D. Shasha*: Block diagonal semistability factors and Lyapunov semistability of block triangular matrices. Lin. Alg. Appl. *172* (1992), 1–25.

[8] *D. H. Carlson and H. Schneider*: Inertia theorems for matrices: the semidefinite case. J. Math. Anal. Appl. *6* (1963), 430–436.
[9] *A. Ja. Dubovickii and A.A. Miljutin*: Extremum problems with certain constraints. Soviet Math. *4* (1963), 759–762; Dokl. Akad. Nauk SSSR *149* (1963), 759–762.
[10] *D. Gale*: The theory of linear economic models. McGraw-Hill, 1960.
[11] *I. V. Girsanov*: Lectures on Mathematical Theory of Extremum Problems [sic]. Lecture Notes in Economics and Mathematical Systems, Vol. 67. Springer-Verlag, 1972.
[12] *D. Hershkowitz and H. Schneider*: Semistability factors and semifactors. Contemp. Math. *47* (1985), 203–216.
[13] *J. L. Kelley, I. Namioka, et al.*: Linear Topological Spaces. van Nostrand, 1963.
[14] *A. N. Lyapunov*: Le problème général de la stabilité du mouvement. Ann. Math. Studies *17* (1949). Princeton University Press.
[15] *H. Nikaido*: Convex Structures and Economic Theory. Mathematics in Science and Engineering, Vol. 51. Academic, 1968.
[16] *A. Ostrowski and H. Schneider*: Some theorems on the inertia of general matrices. J. Math. Analysis and Appl. *4* (1962), 72–84.
[17] *R. T. Rockafellar*: Convex Analysis. Princeton University Press, 1970.
[18] *B. D. Saunders and H. Schneider*: Applications of the Gordan-Stiemke theorem in combinatorial matrix theory. SIAM Rev. *21* (1979), 528–541.
[19] *O. Taussky*: Matrices C with $C^n \to 0$. J. Alg. *1* (1964), 1–10.
[20] *E. Zeidler*: Nonlinear functional analysis and its applications III: Variational methods and optimation. Springer-Verlag, 1985.

Authors' addresses: Bryan Cain, Mathematics Department, Iowa State University, Ames, IA 50011, U.S.A.; Daniel Hershkowitz, Mathematics Department, Technion – Israel Institute of Technology, Haifa 32000, Israel; Hans Schneider, Mathematics Department, University of Wisconsin – Madison, Madison, WI 53706, U.S.A.