PROCEEDINGS OF
SYMPOSIA IN PURE MATHEMATICS
VOLUME VII

# CONVEXITY



AMERICAN MATHEMATICAL SOCIETY
PROVIDENCE, RHODE ISLAND
1963

PROCEEDINGS OF
SYMPOSIA IN PURE MATHEMATICS
VOLUME VII

# CONVEXITY

AMERICAN MATHEMATICAL SOCIETY
PROVIDENCE, RHODE ISLAND
1963

Library of Congress Catalog Number 63-10760

Prepared by the American Mathematical Society under
Contract Number AF 49 (638) – 964 with the United States Air Force

## PREFACE

Of the thirty-two papers in this volume, seventeen were presented at the Symposium on Convexity and the others were submitted later. (Symposium speakers were Besicovitch, Coxeter, Danzer, Davis, Day, Dvoretzky, Fan, Gale, Grünbaum, Hammer, Hoffman, Karlin, Klee, Motzkin, Phelps, Pták, Schaefer, and Valentine.) The thirty-third "paper" included here is a report on unsolved problems, based on the Symposium's session devoted to them, on informal discussions during the Symposium, and on later communications from the participants.

The papers are arranged alphabetically by author, since this seems most convenient for reference purposes. Interrelationships of the various papers, and their relation to the theory as a whole, are discussed in the Introduction. Since some of the individual bibliographies were so long and in such a state of flux, a common list of references did not seem feasible. However, the Author Index (in conjunction with the individual bibliographies) should be a fair substitute for such a list, and also makes it easy to learn which of the thirty-three papers cite the work of a given author. There are also a Subject Index and an Index of Unsolved Problems.

The editor is indebted to Professors Gale and Grünbaum for their assistance in planning the Symposium, to Dr. Pták and Professors Besicovitch, Coxeter, Day, Fan, and Motzkin for presiding at Symposium sessions, and to Dr. Danzer and Professors Besicovitch, Corson, Firey, Grünbaum, McMinn, and Motzkin for refereeing some of the papers. In particular, the advice and assistance of Branko Grünbaum have been invaluable.

The details of publication have been capably handled by Miss Ellen Swanson, Head of the American Mathematical Society's Editorial Department.

Victor Klee

# INTRODUCTION

The systematic study of convex sets was initiated by H. Brunn and H. Minkowski. For most of the important notions in the field, at least a germ can be found in the latter's collected works (1911). Not only does the theory of convexity play a central role in Minkowski's geometry of numbers, but it also shares some of the nontechnical aspects of elementary number theory. Its basic notions are simple, natural, and of strong intuitive appeal. The subject is primarily one of ideas rather than machinery, and does not lend itself readily to unified treatment. It abounds in attractive special problems, and many mathematicians working mainly in other fields have published one or two papers on convexity. These aspects have accounted for the rapid but disorganized growth of the theory.

The 1934 survey by T. Bonnesen and W. Fenchel was an excellent summary of a large body of material, and is still a standard source of information in the field. Though selective in coverage, they cited more than 450 references; a current survey of the same degree of completeness would be a tremendous undertaking, probably not feasible. More than half of their book emphasized various quantitative notions such as diameter, area, volume and mixed volumes. Since 1934 these same notions have continued to play an important role. However, more striking (since less predictable) has been the intensive development of several qualitative aspects of the theory, including the combinatorial geometry associated with intersection and covering properties, the refinement and application (especially in functional analysis and game theory) of such notions as extremal structure and separation properties, the study of convexity in infinite-dimensional spaces, increasing use of convexity as a descriptive tool, and the evolution of various analogues and generalizations of convexity.

Though several quantitative investigations are included here, the Symposium was intended primarily to emphasize the more qualitative aspects of the theory. In particular, the five aspects listed above are all represented in the present volume. Among the unavoidable omissions, two are especially regretted by the editor. There is nothing here about the geometry of convex surfaces and the associated development of metric methods in differential geometry, carried out by A. D. Aleksandrov and his students in the Soviet Union and in this country by H. Busemann. Also omitted are the important results on infinite-dimensional simplexes, boundaries and extremal structure which have been developed in the past few years by G. Choquet and others.

In addition to the wide range of topics treated here, there is much variety of approach. Some of the shorter papers treat a single problem in full detail, while at the other extreme are several long papers which include very few proofs but survey broad areas in the field of convexity.

* * * * * * * * * * * *

Four of the papers are set in the Euclidean plane $E^2$. **BESICOVITCH**'s first paper gives a short proof of the known fact that a set of given constant width has minimum area when it is a Reuleaux triangle. His second paper solves affiirmatively a special case of the following problem: Must a set of constant width $w$ contain a semicircle of diameter $w$? **DANZER** gives a short proof of the known result that if $C$ is a closed convex curve in $E^2$ which does not contain exactly three vertices of any rectangle, then $C$ is a circle. In his first paper, **DAVIS** characterizes rectangles by means of an extremal area property involving inscribed crosses and also discusses a related conjecture of Ungar on extremal perimeters.

**HAMMER**'s first paper is set in an arbitrary Minkowski plane where by the use of outwardly simple line families he is able to give an analytic representation for all convex curves of constant Minkowski width. He also summarizes his earlier work on diametral lines and associated convex bodies.

**BESICOVITCH**'s third paper discusses Coxeter's problem of finding the smallest cage (edges of a convex polyhedron) which will hold a unit-sphere in $E^3$ without permitting it to escape. His other two papers give new proofs of known results concerning smoothness properties of a convex body $K$ in $E^3$ and concerning directions of line segments in the boundary of $K$. In **GALE**'s first paper he uses the Borsuk-Ulam mapping theorem (involving antipodal points) to prove that if a convex body of width $w'$ in $E^n$ is obtained from one of width $w$ by means of a homeomorphism which decreases distances, then $w' \leqq w$.

**COXETER** proposes an exact upper bound for the number of equal non-overlapping spheres in $E^n$ that can touch another of the same size. The difficulty of this problem is indicated by the following quotation: "$\cdots$ Can a rigid material sphere be brought into contact with 13 other such spheres of the same size? Gregory said 'Yes' and Newton said 'No', but 180 years were to elapse before a conclusive answer was given." His historical survey of the problem in $E^n$ extends from a paper by Kepler in 1611 to the latest published works. The problem is treated as the case $\phi = \pi/6$ of the problem of packing $(n-2)$-spheres of angular radius $\phi$ on an $(n-1)$-sphere, and the proposed upper bound is attained when the $(n-2)$-spheres are inscribed in the cells of a regular polytope $\{p, 3, \cdots, 3\}$. Though the bound is not fully established, much supporting evidence is given. Some related material is also discussed, such as the growth of the number of spheres as $n \to \infty$ and the known results for other values of $\phi$.

**PORITSKY** treats a system of linear inequalities of the form $x_1 f_1(\theta) + \cdots + x_n f_n(\theta) \leqq g(\theta)$, where $g$ and the $f_i$'s are real analytic functions of the real variable $\theta$ ranging over a bounded or unbounded interval $I$. He studies the convex region consisting of all points $x = (x_1, \cdots, x_n) \in E^n$ which satisfy the given system of inequalities (for all $\theta \in I$), and is especially concerned with describing the region's boundary in terms of the envelope curve $C$ and its tangent and osculating flats of various dimensions, where $C$ is the set of all points $x$ such that for some $\theta \in I$, $\sum_{i=1}^{n} x_i f_i^{(j)}(\theta) = g^{(j)}(\theta)$ for $0 \leqq j \leqq n-1$ ($^{(j)}$ indicating the $j$th derivative).

**DVORETZKY** reviews his earlier results on near-sphericity in $E^n$, one of which asserts that for each $\varepsilon \in ]0, 1[$ and each positive integer $k$ there exists $N(k, \varepsilon)$ such that every convex body of dimension $\geqq N(k, \varepsilon)$ admits a $k$-dimensional section which is spherical to within $\varepsilon$. He derives new corollaries, including some on orthogonal projections, and discusses some open problems.

Two papers treat the facial structure of convex polyhedra. **GALE**'s second paper is concerned with cyclic polytopes in $R^{2m}$, these being convex polyhedra which are combinatorially equivalent to the convex hull of an $n$-pointed subset of the moment curve $\{(t, t^2, \cdots, t^{2m}): t \in R\}$. They have the remarkable property of being $m$-neighborly in the sense that each $m$ vertices determine a face. He computes the number of $(2m-1)$-dimensional faces of such a polytope and this is conjectured to be the maximum attainable for convex polyhedra in $R^{2m}$ which have $n$ vertices. Certain neighborly polytopes are proved to be cyclic, and regular cyclic polytopes are constructed in $E^{2m}$. **GRÜNBAUM AND MOTZKIN** call an abstract graph $k$-polyhedral provided it is isomorphic with the graph formed by the edges and vertices of a $k$-dimensional convex polyhedron. They prove that each $k$-polyhedral graph contains as subgraph a refinement of $C_{k+1}$, the complete graph with $k+1$ nodes. As Gale's result shows, the graph $C_{k+1}$ is $j$-polyhedral whenever $4 \leqq j \leqq k$; however, this and other sorts of ambiguity are excluded for graphs which are 2-polyhedral or 3-polyhedral.

**VALENTINE** deals mainly with known results on the intersection properties of convex sets. He obtains refinements and new proofs for many of these, his aim being to show what can be accomplished by systematic exploitation of dual cones. His viewpoint is well expressed by the following quotation: "··· since it is a rare coincidence for the proofs of a theorem and its dual to be of equal difficulty, there is a double reason to investigate the dual. One may gain either a simpler proof or a less obvious theorem."

Five of the papers are expository surveys of a sort which should be valuable in any field, and especially in the field of convexity where so many results have been rediscovered so many times and where there are so many elementary unsolved problems. Though including few proofs or none at all, they give rather complete descriptions of known results and existing literature in their respective areas. Some of them include new results as well, and most of them discuss many unsolved problems. Since the papers are themselves summaries, it is hardly feasible to summarize them here, but it may be helpful to list their section headings.

**GRÜNBAUM,** *Borsuk's problem and related questions*— reductions of the problem; partial solutions; universal covers; other results on partitions; coverings by translates; finite sets; related problems.

**GRÜNBAUM,** *Measures of symmetry for convex sets*— distance-functions for spaces of convex sets; invariant points and sets; a property of some measures of symmetry; general methods for geometric definitions of measures of symmetry; known results on special measures of symmetry; some extremal problems which possibly lead to measures of symmetry; an interesting functional; some generalizations.

**DANZER, GRÜNBAUM AND KLEE,** *Helly's theorem and its relatives*— proofs of Helly's theorem; applications of Helly's theorem; the theorems of Carathéodory and Radon; generalizations of Helly's theorems; common transversals; some covering problems; intersection theorems for special families;

other intersection theorems; generalized convexity. (The last section makes little contact with the others. It contains a rather complete survey of existing generalizations of the notion of convex set.)

**KLEE,** *Infinite-dimensional intersection theorems*—intersection theorems for infinite families (also in $R^n$); intersection theorems involving the weak topology; intersection properties of metric cells.

**CUDIA,** *Rotundity*—rotundity and smoothness properties; comparison of properties; product spaces, quotient spaces, and subspaces; duality; geometry and reflexivity.

Like those of Cudia and Klee, the papers by **BISHOP AND PHELPS** and by **PHELPS** are concerned with the geometry of infinite-dimensional convex sets. The principal result of Bishop and Phelps is that if $C$ is a closed convex subset of a Banach space, then the support points of $C$ are dense in the boundary of $C$. They show also that for each bounded closed convex subset $C$ of a Banach space $E$, the members of the conjugate space $E^*$ which attain their maximum on $C$ are dense in $E^*$ (norm topology). Several other interesting results are obtained by the same methods. The paper by Phelps treats some of the more technical points which arise when the space is not normable. In particular, he uses supporting cones to give a new proof of the existence of relative extreme points, where a convex cone $K$ with vertex $x$ is said to support the convex set $C$ provided $C \cap K = \{x\}$.

**CORSON AND KLEE** show that the topological classification problem for closed convex bodies in a normed linear space $E$ can be reduced to that for $E$'s unit cell and its closed linear subspaces of finite deficiency. For all $\aleph_0$-dimensional spaces as well as for a wide variety of infinite-dimensional Banach spaces, the problem is solved by proving that all closed convex bodies in $E$ are homeomorphic with $E$ itself. The main tool is the fact that certain spaces are homeomorphic with their positive cones. Also obtained are some results on uniformly continuous transformations of convex sets.

The remaining papers are not so directly concerned with convex sets as such, though in each case some sort of convexity is essential either in the paper itself or for its motivation. Both Karlin and Davis deal with convex functions. For real intervals $X$ and $Y$, **KARLIN** considers the functional transformation $T$ carrying a real function $f$ on $Y$ into the function $g = Tf$ on $X$ given by the formula $gx = \int_Y K(x, y)f(y)dy$, the kernel $K$ being a bounded measurable function on the rectangle $X \times Y$. He is especcially interested in conditions on $K$ which insure that $g$ is convex whenever $f$ is bounded and convex; a similar problem for monotone functions is also considered. The conditions obtained involve the total positivity or sign-regularity of $K$, where $K$ is said to be sign-regular of order $r$ provided there exists a sequence of numbers $(\varepsilon_m)$, each either $+1$ or $-1$, such that whenever $x_1 < x_2 < \cdots < x_m$, $y_1 < y_2 < \cdots < y_m$, $x_i \in X$, $y_j \in Y$, and $1 \leqq m \leqq r$, then $\varepsilon_m K(x_1, \cdots, x_m; y_1, \cdots, y_m) \geqq 0$, where the expression $K(\cdots;\cdots)$ is the determinant of the matrix which has $K(x_i, y_j)$ in the $i$th row and the $j$th column; $K$ is totally positive of order $r$ provided this condition holds with all the $\varepsilon_m$'s equal to $+1$. Inter-relationships among various classes of kernels are studied, and many examples are given.

In his second paper, **DAVIS** studies various classes of real-valued convex functions (of one or several real variables) where for each class the defining condition involves the class $\boldsymbol{H}_n$ of $n \times n$ (real) symmetric matrices. For example, if $f$ is a function of one real variable and the matrix $A \in \boldsymbol{H}_n$ has spectral representation $A = \sum_1^n \lambda_i P_i$, it is customary to write $f(A) = \sum_1^n f(\lambda_i) P_i$. In this way $f$ can be regarded as a function on $\boldsymbol{H}_n$ to $\boldsymbol{H}_n$. The function $f$ is called matrix-convex provided $f((1-\lambda)A + \lambda B) \leqq (1-\lambda)f(A) + \lambda f(B)$ for all $\lambda \in [0, 1]$ and $A, B \in \boldsymbol{H}_n$, where the ordering is that induced in $\boldsymbol{H}_n$ by agreeing that a member of $\boldsymbol{H}_n$ is non-negative if and only if it is positive semidefinite. The matrix-convex functions form a proper subclass of the ordinary convex functions and are closely related to the matrix-monotone functions of Loewner. The paper is devoted to an exposition of Loewner's theory along with related ideas for several variables due to Korányi, Sherman, and Davis himself.

In addition to the paper of Poritsky mentioned earlier, two other papers are included here because of the close connections between convex sets and linear inequalities. **BELLMAN AND FAN** study systems of linear inequalities in which the variables are Hermitian matrices and the ordering is defined as in the paper of Davis just mentioned. They find consistency conditions for various systems of inequalities, the conditions being quite analogous to those in the classical situation except that in each case the consistency of an auxiliary system must be assumed. Also included are several interesting examples, as well as results on the minimum and maximum of the traces of certain matrices related to the systems in question.

**HOFFMAN** supplies a unified approach to some linear programming problems which are amenable to "obvious" solutions. His guide is the observation by Monge that if unit quantities are to be transported from points $X$ and $Y$ to points $Z$ and $W$ (not necessarily respectively) so as to minimize the total distance traveled, then the two routes cannot intersect. He defines a Monge sequence to be an ordering of the set $\{(i,j): 1 \leqq i \leqq m, 1 \leqq j \leqq n\}$ and introduces the notion of such a sequence being consonant with a given $m \times n$ matrix. An algorithm is given whereby a solution for the transportation problem associated with a given matrix can be derived from a Monge sequence consonant with the matrix. The warehouse problem of Cahn is transformed into one to which this algorithm is applied and many other problems are mentioned to which the same idea is applicable.

The many new notions in **MOTZKIN**'s paper are treated in 79 theorems distributed among 50 sections. The paper is concisely written and can hardly be summarized here, but we shall describe its basic idea. Let $R$ be a (not necessarily commutative) ring with unit 1 and let $V$ be a left module over $R$. Let $\tilde{R}$ be the set of all finite sequences $\lambda = (\lambda_1, \cdots, \lambda_k)$ of members of $R$. When $S \subset V$, the vector $\lambda$ is said to be an endovector of $S$, or $S$ is said to be endo-$\lambda$, provided $S$ includes the point $\sum_{i=1}^k \lambda_i s_i$ for every choice of $s_1, \cdots, s_k \in S$. For $\Lambda \subset R$, $S$ is said to be endo-$\Lambda$ provided $S$ is endo-$\lambda$ for each $\lambda \in \Lambda$. Since the family of all endo-$\Lambda$ sets in $V$ is intersectional, the $\Lambda$-hull of $S$ is defined as the smallest endo-$\Lambda$ set which contains $S$. The set $\Lambda$ is said to be complete provided for some $S$, $\Lambda$ is the set of all endovectors of $S$. These and related notions are studied in some detail, where of course the most important cases are those in which $R$ is the real field and the condition $(\lambda_1, \cdots, \lambda_k) \in \Lambda$ is equivalent to one of the following: (i) $\lambda \in \tilde{R}$; (ii) $\sum_1^k \lambda_i = 1$; (iii) $\lambda_i \geqq 0$; (iv) $\sum_1^k \lambda_i = 1$ and $\lambda_i \geqq 0$. The corresponding endo-$\Lambda$ sets are the linear subspaces ($O$-flats), the affine subspaces (flats), the positive cones (convex cones with vertex $O$), and the convex sets.

**MOTZKIN AND STRAUS** are concerned with representing the points of a set as linear combinations of boundary points. Their principal result asserts that if $\alpha_1 + \cdots + \alpha_n = 1$ and $\sum_{i \neq j} |\alpha_i| \geqq |\alpha_j|$ for $1 \leqq j \leqq n$, then for very general sets $S$ it is true that each point of $S$ can be represented in the form $p = \sum_1^k \alpha_i x_i$ for points $x_i$ of the outer boundary of $S$.

**PTÁK** presents a unified treatment of several important results on weak compactness, all of which are shown to follow from a combinatorial lemma which gives conditions for the existence of certain convex means. For an infinite set $S$, let $M(S)$ denote the set of all functions $\lambda$ on $S$ to $[0, \infty[$ for which the set $N(\lambda)$ is finite and $\sum_{s \in S} \lambda(s) = 1$, where $N(\lambda) = \{s \in S: \lambda(s) > 0\}$. Let $\mathscr{W}$ be a family of subsets of $S$, and for $\varepsilon > 0$ and $H \subset S$ let $M(H, \mathscr{W}, \varepsilon)$ denote the set of all $\lambda \in M(S)$ such that $N(\lambda) \subset H$ and $\sum_{w \in W} \lambda(w) < \varepsilon$ for all $W \in \mathscr{W}$. The lemma asserts the equivalence of the following two conditions: (1) $M(H, \mathscr{W}, \varepsilon) = \varnothing$ for some infinite $H \subset S$ and some $\varepsilon > 0$; (2) there exists a sequence $(s_n)$ of distinct points of $S$ and a sequence $(W_n)$ of members of $\mathscr{W}$ such that $\{s_1, \cdots, s_n\} \subset W_n$ for all $n$. With the aid of this lemma he proves that if $A$ is a subset of a complete convex space $E$ and $A$ satisfies a certain double limit condition, then the closed convex hull of $A$ is weakly compact. This includes the well-known theorems of Krein and Eberlein on weak compactness. The same lemma is employed to yield an extensive series of results on weak convergence and weak compactness in locally convex spaces and especially in spaces of continuous functions.

**SCHAEFER** is concerned with spectral properties in an ordered locally convex algebra $A$, where this is a locally convex algebra (usually over the complex field) with unit $e$ and with an associated positive cone $K \ni e$ such that $K$ is closed, proper, includes the product of any commuting pair of its elements, and is normal in the sense that there is a family of pseudonorms $p$ on $E$ which generate the topology and are such that $p(x + y) \geqq p(x)$ for all $x, y \in K$. The principal motivating example of such an $A$ is the algebra of all continuous endomorphisms of a Hilbert space, where $K$ is the cone of positive Hermitian operators and the topology is that of either bounded or pointwise convergence. (There are other important examples also.) The paper contains much interesting material on such algebras $A$, its principal results showing that the spectral behavior of certain members of $K$ is quite analogous to that in the finite-dimensional case. In particular, the members of $K$ whose spectrum is bounded have spectral behavior like that of positive matrices, while those in the unit interval of $K$ (i.e., those $a \in A$ for which $0 \leqq a \leqq e$—diagonal positive matrices in the classical case) behave spectrally like positive Hermitian operators.

**FAN**'s paper is motivated by the Krein-Milman extreme point theorem. He establishes a general lemma which is purely set-theoretical in character, involving neither topological nor vector space concepts, from which the Krein-Milman theorem follows. (Another lemma, in a sense dual to the first, is shown to imply theorems on filters due to Wallman and Stone.) He then considers a set $\Phi$ of real-valued functions on a set $S$, calling a set $X \subset S$ convex provided $X$ is an intersection of sets of the form $\{x \in S: f(x) \geqq \alpha\}$. Since the family of $\Phi$-convex sets is intersectional, the $\Phi$-hull can be defined in the natural way. The notion of $\Phi$-betweenness is defined for points of $S$ and in terms of this the $\Phi$-extreme points of subsets of $X$ are defined. These notions appear in several theorems which generalize known results on extreme points and are related to the abstract minimum principal of Bauer.

**HAMMER**'s second paper is motivated by his notion of a semispace at a point $p$ in a linear space $L$, this being a maximal convex subset of $L \sim \{p\}$. He reviews some of the known results on semispaces, including their connection with extreme points and the fact that the semispaces form a minimal intersection base for the convex subsets of $L$. He then describes his system of extended topology which arose from an attempt to consider certain processes and concepts associated with convexity (and especially with semispaces) as topological in character. Many new notions are introduced, complications arising mainly from the fact that in place of the usual topological closure operation he considers an arbitrary expansive function $g$—i.e., one associating with each set some superset thereof. After discussing the extended topology, he interprets the various notions in terms of convexity, where $gX$ is the union of $X$ with all the line segments determined by points of $X$. Several unsolved problems are mentioned.

V.K.

# CONTENTS

# ON SYSTEMS OF LINEAR INEQUALITIES IN HERMITIAN MATRIX VARIABLES

BY

RICHARD BELLMAN AND KY FAN[1]

**1. Introduction.** This paper is concerned with systems of linear inequalities, in which the variables are Hermitian matrices. The inequalities between Hermitian matrices are to be understood in the sense of positive semidefiniteness or positive definiteness. More precisely, for two Hermitian matrices $H$ and $K$ of same order, we write $H \geqq K$ to signify that $H - K$ is positive semi-definite. Similarly, the strict inequality $H > K$ means that $H - K$ is positive definite. Consider the system (2) of linear inequalities, where $A_{ij}$ are arbitrary square complex matrices, $B_i$ and $C_j$ are Hermitian matrices (all of same order), and $c$ is a real number. Theorem 1 gives a necessary and sufficient condition for the system (2) to be consistent, i.e., for the existence of Hermitian matrices $X_j$ satisfying (2). Consistency conditions for more special systems (10), (14), (18) and (22) are also explicitly stated. Then we derive two theorems on minimum and maximum (Theorems 2,3). Theorem 2 asserts that the minimum of the trace of $\sum_{j=1}^{n} C_j X_j$, when $\{X_j\}$ varies over all solutions of the system (14), is equal to the maximum of the trace of $\sum_{i=1}^{m} B_i Y_i$, when $\{Y_i\}$ varies over all solutions of the system (18), provided that the systems (14′), (18′) of strict inequalities are consistent.

These results are analogous to the well-known theorems on systems of linear inequalities in real variables (see [2; 3]). However, for the case of Hermitian matrix variables, each of our theorems requires an additional hypothesis which is not needed in the case of real variables. Thus, in Theorem 1 we assume that there exist positive definite Hermitian matrices $Y_i$ satisfying (1); in Theorem 2 we assume that the systems (14′), (18′) of strict inequalities (instead of the systems (14), (18)) are consistent. These hypotheses are indeed essential (see Examples 2, 3).

The systems studied in this paper are quite natural, especially in the case $m = n = 1$. For instance, when $m = n = 1$, system (22) becomes

$$Y \geqq 0\,, \qquad YA + A^*Y = C\,,$$

where $A$ is an arbitrary square complex matrix and $C, Y$ are Hermitian. This differs slightly from the familiar system

$$Y > 0\,, \qquad YA + A^*Y = -I$$

($I$ being the identity matrix), which arises in stability problems of differential equations. It is a classical theorem of Lyapunov (see [1, Chapter 13; 4,

[1] The work of the second author was supported by the U. S. Atomic Energy Commission at Argonne National Laboratory.

Chapter XV, § 5]) that there exists a positive definite Hermitian matrix $Y$ satisfying $YA + A^*Y = -I$ if and only if all eigenvalues of $A$ have negative real parts.

All matrices considered here are square matrices with complex elements. Since the matrices considered in a theorem (except in the proof of Theorem 1) are always of same order, the order will often not be mentioned. Throughout the paper, $A_{ij}$ are arbitrary square complex matrices which are not necessarily Hermitian, $B_i, C_j, X_j$ and $Y_i$ are Hermitian matrices. As usual, the adjoint of a matrix $A$ is denoted by $A^*$, the trace of $A$ is denoted by $\operatorname{tr} A$. We repeat again that, for two Hermitian matrices $H, K$ of same order, $H \geqq K$ means that $H - K$ is positive semi-definite, and $H > K$ means that $H - K$ is positive definite.

**2. Existence theorem.** We begin with a lemma which will be needed in the proof of Theorem 1.

LEMMA. *Let $H_1, H_2, \cdots, H_p$ be $p$ Hermitian matrices of same order. Let $\mathcal{Q}$ denote the convex cone in the Euclidean $p$-space $\mathcal{R}^p$ formed by all points with coordinates $(\operatorname{tr} H_1Z, \operatorname{tr} H_2Z, \cdots, \operatorname{tr} H_pZ)$, when $Z$ varies over all positive semi-definite Hermitian matrices (of the same order as the $H_k$'s). If there exist real numbers $c_1, c_2, \cdots, c_p$ such that $\sum_{k=1}^p c_kH_k$ is positive definite, then $\mathcal{Q}$ is a closed set in $\mathcal{R}^p$.*

PROOF. Let $Z_1, Z_2, \cdots, Z_i, \cdots$ be a sequence of positive semi-definite Hermitian matrices (of the same order as the $H_k$'s) and let $\lim_{i\to\infty} \operatorname{tr} H_kZ_i = t_k$ $(1 \leqq k \leqq p)$. We want to prove that $(t_1, t_2, \cdots, t_p) \in \mathcal{Q}$. Let $H_0 = \sum_{k=1}^p c_kH_k$ be positive definite, and let $a$ be a positive number such that $aI \leqq H_0$, where $I$ denotes the identity matrix. We have $\lim_{i\to\infty} \operatorname{tr} H_0Z_i = \sum_{k=1}^p c_kt_k$, so there exists a positive number $b$ such that $a \cdot \operatorname{tr} Z_i \leqq \operatorname{tr} H_0Z_i \leqq b$ for all $i = 1, 2, 3, \cdots$. Then $0 \leqq \operatorname{tr} Z_i \leqq b/a$ for all $i$. Since $Z_i$ is positive semi-definite, $\operatorname{tr} Z_i^2 \leqq (\operatorname{tr} Z_i)^2 \leqq (b/a)^2$, which implies that the elements of $Z_i$ are bounded. There exists a subsequence $\{Z_{\nu_i}\}$ such that $\lim_{i\to\infty} Z_{\nu_i} = Z$ exists. Then $Z$ is a positive semi-definite Hermitian matrix and $\operatorname{tr} H_kZ = \lim_{i\to\infty} \operatorname{tr} H_kZ_{\nu_i} = t_k$ $(1 \leqq k \leqq p)$. Hence $(t_1, t_2, \cdots, t_p) \in \mathcal{Q}$.

EXAMPLE 1. Let

$$H_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad H_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

The set $\mathcal{Q}$ in $\mathcal{R}^2$ formed by all points with coordinates $(\operatorname{tr} H_1Z, \operatorname{tr} H_2Z)$, when $Z$ varies over all positive semi-definite Hermitian matrices of order 2, is not closed. In fact, $\mathcal{Q}$ consists of the origin $(0, 0)$ and all points $(x, y)$ with $x > 0$ and $-\infty < y < +\infty$.

THEOREM 1. *Let $A_{ij}$ be arbitrary matrices, $B_i$, $C_j$ be Hermitian matrices (all of same order), and let $c$ be a real number. Suppose that there exist positive definite Hermitian matrices $Y_i$ satisfying*

$$(1) \qquad \sum_{i=1}^m (Y_iA_{ij} + A_{ij}^*Y_i) + C_j = 0 \qquad (1 \leqq j \leqq n).$$

*Then the system*

$$\text{(2)}\qquad \sum_{j=1}^{n}(A_{ij}X_j + X_jA_{ij}^*) \geqq B_i \qquad (1 \leqq i \leqq m),$$
$$\operatorname{tr}\sum_{j=1}^{n} C_jX_j \geqq c$$

*is consistent, i.e., solvable for Hermitian matrices $X_j$, if and only if, for any $m$ positive semi-definite Hermitian matrices $D_i$ and any non-negative number $d$, the relations*

$$\text{(3)}\qquad \sum_{i=1}^{m}(D_iA_{ij} + A_{ij}^*D_i) + dC_j = 0 \qquad (1 \leqq j \leqq n)$$

*imply*

$$\text{(4)}\qquad \operatorname{tr}\sum_{i=1}^{m} D_iB_i + dc \leqq 0 .$$

PROOF. *Necessity.* Assume that $X_1, X_2, \cdots, X_n$ are Hermitian matrices satisfying (2). For any $m$ positive semi-definite Hermitian matrices $D_i$ and any non-negative number $d$, we have

$$\operatorname{tr}\sum_{i=1}^{m}\sum_{j=1}^{n} D_i(A_{ij}X_j + X_jA_{ij}^*) + d\cdot\operatorname{tr}\sum_{j=1}^{n} C_jX_j \geqq \operatorname{tr}\sum_{i=1}^{m} D_iB_i + dc ,$$

which can be written

$$\operatorname{tr}\sum_{j=1}^{n}\left\{\sum_{i=1}^{m}(D_iA_{ij} + A_{ij}^*D_i) + dC_j\right\}X_j \geqq \operatorname{tr}\sum_{i=1}^{m} D_iB_i + dc .$$

Hence relations (3) imply (4).

*Sufficiency.* Let $\mathscr{H}$ denote the real vector space of all Hermitian matrices (of the same order as the $A_{ij}$'s) and let $\mathscr{R}$ denote the one-dimensional vector space (i.e., the real line). Let $\mathscr{H}^m = \mathscr{H}\oplus\mathscr{H}\oplus\cdots\oplus\mathscr{H}$ be the direct sum with $m$ summands. In the vector space $\mathscr{H}^m\oplus\mathscr{R}$, a vector $\sigma = (S_1, S_2, \cdots, S_m, s)$ is determined by $m$ Hermitian matrices $S_i \in \mathscr{H}$ and a real number $s$. In $\mathscr{H}^m \oplus \mathscr{R}$ an inner product is defined in a natural way as follows. For two vectors $\sigma = (S_1, S_2, \cdots, S_m, s)$ and $\tau = (T_1, T_2, \cdots, T_m, t)$ in $\mathscr{H}^m\oplus\mathscr{R}$, the inner product $(\sigma, \tau)$ of $\sigma$, $\tau$ is

$$(\sigma, \tau) = \operatorname{tr}\sum_{i=1}^{m} S_iT_i + st .$$

We can therefore speak of orthogonality in $\mathscr{H}^m \oplus \mathscr{R}$.

Let $\mathscr{S}$ denote the linear subspace of $\mathscr{H}^m \oplus \mathscr{R}$ formed by all vectors of the form

$$\left(\sum_{j=1}^{n}(A_{1j}X_j + X_jA_{1j}^*), \cdots, \sum_{j=1}^{n}(A_{mj}X_j + X_jA_{mj}^*), \operatorname{tr}\sum_{j=1}^{n} C_jX_j\right),$$

where $X_j \in \mathscr{H}$. Let

$$\delta_k = (D_{k1}, D_{k2}, \cdots, D_{km}, d_k) \qquad (1 \leqq k \leqq p)$$

(where $D_{ki} \in \mathscr{H}$, $d_k \in \mathscr{R}$) be $p$ vectors which span the orthogonal complement of $\mathscr{S}$ in $\mathscr{H}^m \oplus \mathscr{R}$. Let $\beta = (B_1, B_2, \cdots, B_m, c)$ and $b_k = -(\delta_k, \beta)$ $(1 \leqq k \leqq p)$. Then the linear variety $\mathscr{S} - \beta$ is formed by all vectors $\sigma = (S_1, S_2, \cdots, S_m, s)$ in $\mathscr{H}^m \oplus \mathscr{R}$ satisfying $(\delta_k, \sigma) = b_k$ $(1 \leqq k \leqq p)$, i.e.,

$$\text{tr} \sum_{i=1}^{m} D_{ki} S_i + d_k s = b_k \qquad (1 \leqq k \leqq p). \tag{5}$$

Let $\mathscr{P}$ denote the set of all vectors $(Z_1, Z_2, \cdots, Z_m, z)$ in $\mathscr{H}^m \oplus \mathscr{R}$ such that each $Z_i$ is positive semi-definite and $z \geqq 0$. In the Euclidean $p$-space $\mathscr{R}^p$, let $\mathscr{Q}$ denote the convex cone formed by all points with coordinates of the form

$$\left( \text{tr} \sum_{i=1}^{m} D_{1i} Z_i + d_1 z, \cdots, \text{tr} \sum_{i=1}^{m} D_{pi} Z_i + d_p z \right),$$

when $(Z_1, Z_2, \cdots, Z_m, z)$ varies in $\mathscr{P}$. For $k = 1, 2, \cdots, p$, let

$$H_k = \text{diag.}\{D_{k1}, D_{k2}, \cdots, D_{km}, d_k\}$$

denote the Hermitian matrix such that $D_{k1}, D_{k2}, \cdots, D_{km}, d_k$ are its successive diagonal blocks and all elements outside these blocks are zero. It is clear that $\mathscr{Q}$ coincides with the set in $\mathscr{R}^p$ formed by all points with coordinates of the form $(\text{tr}\, H_1 Z, \text{tr}\, H_2 Z, \cdots, \text{tr}\, H_p Z)$, when $Z$ varies over all positive semi-definite Hermitian matrices (of the same order as the $H_k$'s). According to the above lemma, in order to show that $\mathscr{Q}$ is closed in $\mathscr{R}^p$, it suffices to find $p$ real numbers $c_1, c_2, \cdots, c_p$ such that $\sum_{k=1}^{p} c_k H_k$ is positive definite. Now, by hypothesis, there exist $m$ positive definite Hermitian matrices $Y_i$ satisfying (1). So we have

$$\text{tr}\left[\sum_{i=1}^{m} Y_i \sum_{j=1}^{n} (A_{ij} X_j + X_j A_{ij}^*)\right] + \text{tr} \sum_{j=1}^{n} C_j X_j$$
$$= \text{tr} \sum_{j=1}^{n} \left\{ \sum_{i=1}^{m} (Y_i A_{ij} + A_{ij}^* Y_i) + C_j \right\} X_j = 0$$

for any $X_j \in \mathscr{H}$. Thus in $\mathscr{H}^m \oplus \mathscr{R}$, the vector $\eta = (Y_1, Y_2, \cdots, Y_m, 1)$ is orthogonal to the linear subspace $\mathscr{S}$. Therefore $\eta$ is a linear combination of $\delta_1, \delta_2, \cdots, \delta_p$. We can find real numbers $c_1, c_2, \cdots, c_p$ such that $Y_i = \sum_{k=1}^{p} c_k D_{ki}$ $(1 \leqq i \leqq m)$ and $1 = \sum_{k=1}^{p} c_k d_k$. Then

$$\sum_{k=1}^{p} c_k H_k = \text{diag.}\{Y_1, Y_2, \cdots, Y_m, 1\},$$

i.e., $Y_1, Y_2, \cdots, Y_m, 1$ are the successive diagonal blocks in the matrix $\sum_{k=1}^{p} c_k H_k$, and all elements outside these blocks are zero. Hence $\sum_{k=1}^{p} c_k H_k$ is positive definite and $\mathscr{Q}$ is closed in $\mathscr{R}^p$.

Assume now that system (2) is inconsistent, which means

$$(\mathscr{S} - \beta) \cap \mathscr{P} = \emptyset. \tag{6}$$

In view of the characterization (5) of $\mathscr{S} - \beta$ and our definition of $\mathscr{Q}$, (6) amounts to saying that in $\mathscr{R}^p$, the point $(b_1, b_2, \cdots, b_p)$ is not contained in the