AN ABSTRACT OF THE THESIS OF

Richard David Hill for the M. A. in Mathematics
(Name) (Degree) (Major)

Date thesis is presented September 4, 1964

Title GENERALIZATIONS OF THE OSTROWSKI-SCHNEIDER MAIN INERTIA THEOREM

Abstract approved Redacted for Privacy
(Major professor)

As indicated by the title, this thesis generalizes the Main Inertia Theorem of Ostrowski and Schneider [8] . The first three results concern the formation of a polynomial function $f(A, A^*, H)$ so that the existence of an hermitian H for which $f(A, A^*, H)$ is positive definite is a necessary and sufficient condition that the matrix A have no eigenvalues on an arbitrary line, circle, and parabola (respectively) in the complex plane. The next two results are motivated by the existence of a hermitian H such that a certain polynomial function $f(A, A^*, H)$ being positive definite is necessary and sufficient that H have no eigenvalues on a certain point set in the complex plane. Finally, this thesis demonstrates how the methods previously described can generalize (1) the second part of the Main Inertia Theorem and (2) three theorems by Drazin and Haynsworth [5] concerning necessary and sufficient conditions for the existence of a set of m linearly independent eigenvectors of a complex matrix A all

corresponding to real eigenvalues, purely imaginary eigenvalues, and eigenvalues of absolute value one respectively.

GENERALIZATIONS OF THE OSTROWSKI-SCHNEIDER MAIN INERTIA THEOREM

by

RICHARD DAVID HILL

A THESIS

submitted to

OREGON STATE UNIVERSITY

in partial fulfillment of
the requirements for the
degree of

MASTER OF ARTS

June 1965

APPROVED:

Redacted for Privacy

Assistant Professor of Mathematics

Redacted for Privacy

Head of Department of Mathematics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented September 4, 1964

Typed by Muriel Davis

ACKNOWLEDGMENT

The author wishes to take this opportunity to thank his major professor Dr. David Carlson for his helpful comments, patience, and willingness to listen during the research for and writing of this thesis.

TABLE OF CONTENTS

Chapter		Page
1	INTRODUCTION	1
2	THE LINE	8
3	THE CIRCLE	15
4	GENERALIZATION OF A CONFORMAL MAPPING	19
5	POLYNOMIAL RESULTS	23
6	GENERALIZATION OF THE SECOND PART OF THE MAIN INERTIA THEOREM	26
7	GENERALIZATION OF THE DRAZIN-HAYNSWORTH THEOREM	29
	BIBLIOGRAPHY	31

GENERALIZATIONS OF THE OSTROWSKI-SCHNEIDER MAIN INERTIA THEOREM

CHAPTER 1. INTRODUCTION

A large amount of research in matrix theory in the past few years has concerned the location in the complex plane of the eigenvalues of an n by n complex valued matrix. The theory concerning certain types of special matrices has been developed for a long time. For example, diagonal and triangular matrices exhibit their eigenvalues on their main diagonal.

Basic to this thesis is the fundamental topic of the location of eigenvalues in half-planes, which leads us to the theory of matrix inertias. The inertia of a matrix A is defined to be the integer triple $\text{In } A = (\pi, \nu, \delta)$ where π and ν are the number of eigenvalues of A in the (open) right and left half-planes respectively, and δ is the number of eigenvalues on the imaginary axis. The only two classical theorems involving matrix inertias are those credited to Sylvester and Lyapunov.

<u>Sylvester's Theorem</u>: <u>If</u> P <u>is non-singular and</u> H <u>is hermitian, then</u> $\text{In } H = \text{In } (PHP^*)$ (where P^* is the conjugate-transpose of P).

<u>Lyapunov's Theorem</u>: <u>For a given</u> A <u>there exists a hermitian</u> $H>0$ (positive definite) <u>such that</u> $\text{Re}\{AH\} > 0$ iff $\text{In } A = (n, 0, 0)$.

Here the Toeplitz decomposition of a complex valued matrix A is used:

$$A = \mathrm{Re}\{A\} + i\,\mathrm{Im}\,\{A\}$$

where $\qquad \mathrm{Re}\{A\} = \dfrac{A + A^*}{2}\ , \quad \mathrm{Im}\,\{A\} = \dfrac{A - A^*}{2i}$

In 1962 Schneider and Ostrowski published a paper _Some Theorems on the Inertia of General Matrices_ [8] in which they generalized the classical theorems of Sylvester and Lyapunov to what they called the Main Inertia Theorem.

Main Inertia Theorem: Given a complex valued matrix A _there exists a hermitian_ H _such that_ $\mathrm{Re}\{AH\} > 0$ iff $\delta(A) = 0$ (i.e., iff A _has no eigenvalues which lie on the line_ $\mathrm{Re}\,\{z\} = 0$ _in the complex plane._) _Further, if_ $\mathrm{Re}\,\{AH\} > 0$, _then_ $\mathrm{In}\ A = \mathrm{In}\ H$.

We note in passing that Taussky [10, 11] independently achieved the above result.

This thesis generalizes the Ostrowski-Schneider theorem. We first derive a theorem which states that given a complex valued matrix A there exists a hermitian H such that $f(A, A^*, H) > 0$ iff A has no eigenvalues which lie on a given line in the complex plane for a polynomial function $f(A, A^*, H)$ specifically determined by the line. The next theorem in this thesis is a similar generalization of the Ostrowski-Schneider theorem corresponding to an arbitrary circle in the complex plane. Then we consider a method of deriving corresponding results for an arbitrary curve by a generalization of a conformal mapping. The next two results of this thesis start with the existence of a hermitian H such that a certain polynomial function $f(A, A^*, H) > 0$ and derive the set of points in the complex plane from which eigenvalues are excluded.

The final two chapters explain how the methods of this thesis can generalize (1) the remaining part of the Ostrowski-Schneider theorem and (2) a series of theorems by Drazin and Haynsworth regarding necessary and sufficient conditions for the existence of a set of m linearly independent eigenvectors of a complex matrix A all corresponding to real eigenvalues, purely imaginary eigenvalues, and eigenvalues of absolute value one respectively.

Motivation for this thesis has come from two sources. First, Householder and Varga noticed a connection between Lyapunov's theorem and Stein's theorem [9] which concerns matrices whose n^{th} power approaches zero as n approaches infinity. For our purposes it is important to note that the condition $\lim_{n\to\infty} C^n=0$ is equivalent to the condition that all eigenvalues of C lie within the unit circle.

Stein's Theorem: The matrix C _satisfies_ $\lim_{n\to\infty} C^n = 0$ iff _there exists a matrix_ $H>0$ _such that_ $H - CHC^* > 0$.

Taussky has discussed the equivalence of Lyapunov's and Stein's theorems in her paper Matrices C with $C^n \to 0$ [12].

Secondly, comparison of the three previously mentioned Drazin-Haynsworth theorems from their paper Criteria for the Reality of Matrix Eigenvalues suggests a similar generalization. These theorems are listed in Chapter 7 along with a discussion of the corresponding generalization.

Before we proceed to the body of this thesis, it seems appropriate to list the notation used and two propositions basic to the results.

Notation

A, B, C, D	n by n complex valued matrices
H	n by n hermitian matrix
A*	conjugate transpose of A
> 0	positive definite
≥ 0	positive semi-definite
α, β, γ, k	complex constants
a, b, c	real constants
χ	column vector
$w = u + iv$ $z = x + iy$	complex variables where u, v, x and y are real

Proposition I. Given A, polynomials p and q such that q(A) is non-singular, λ an eigenvalue of A, then $p(\lambda)/q(\lambda)$ is an eigenvalue of $[q(A)]^{-1} p(A)$.

Proof: By definition, a matrix A has an eigenvalue λ iff there exists a column vector $\chi \neq 0$ such that $A\chi = \lambda\chi$. We must show that $\{[q(A)]^{-1}p(A)\}\chi = \{p(\lambda)/q(\lambda)\}\chi$ under the hypotheses of this proposition. This proof is lengthy. It yields some intermediate results which will be called lemmas.

Let $p(A) = a_n A^n + a_{n-1}A^{n-1} + \ldots + a_1 A + a_0 I$.

<u>Lemma 0.1.</u> <u>If</u> A <u>has an eigenvalue</u> λ, <u>then</u> A^j <u>has an eigenvalue</u> λ^j $(j=1, 2, 3 \ldots)$. The proof will be by induction on j.

I. Lemma 0.1 is true for $j = 1$ by hypothesis.

II. Assume as our induction hypothesis that A^k has an eigenvalue λ^k.

$$A^{k+1}\chi = A(A^k\chi) = A(\lambda^k\chi) = \lambda^k(A\chi) = \lambda^k(\lambda\chi) = \lambda^{k+1}\chi .$$

Thus A^{k+1} has an eigenvalue λ^{k+1} and we have proven our first introductory lemma.

<u>Lemma 0.2.</u> <u>If</u> A <u>has an eigenvalue</u> λ, <u>then</u> bA <u>has an eigenvalue</u> $b\lambda$. (b a complex constant)

Proof: $(bA)\chi = b(A\chi) = b(\lambda\chi) = (b\lambda)\chi$.

<u>Lemma 0.3.</u> <u>If</u> A <u>has an eigenvalue</u> λ, <u>then</u> $p(A)$ <u>has an eigen-</u> <u>value</u> $p(\lambda)$.

Proof: $[p(A)]\chi = (a_n A^n + \ldots + a_1 A + a_0 I)\chi$

$$= a_n A^n\chi + \ldots + a_1 A\chi + a_0\chi$$

$$= a_n\lambda^n\chi + \ldots + a_1\lambda\chi + a_0\chi$$

(By Lemmas 0.1 and 0.2)

$$= (a_n\lambda^n + \ldots + a_1\lambda + a_0)\chi$$

$$= p(\lambda)\chi.$$

Lemma 0.4. *If* A *has an eigenvalue* λ *and* $q(A)$ *is non-singular, then* $q(\lambda) \neq 0$ *and* $[q(A)]^{-1}$ *has an eigenvalue* $1/q(\lambda)$.

Proof: We know that $q(\lambda)$ is an eigenvalue of $q(A)$ (Lemma 0.3). Since $q(A)$ is non-singular, $q(\lambda) \neq 0$ and $[q(A)]^{-1}$ exists. Now multiplying the matrix equation $[q(A)]\chi = q(\lambda)\chi$ on the left by $[q(\lambda)q(A)]^{-1}$ we get

$$\frac{1}{q(\lambda)}\chi = [q(A)]^{-1}\chi$$

Thus, by definition $[q(A)]^{-1}$ has an eigenvalue $1/q(\lambda)$.

Now for the proof of Proposition 1 using the lemmas developed.

$$\{[q(A)]^{-1}p(A)\}\chi = [q(A)]^{-1}\{[p(A)]\chi\}$$

$$= [q(A)]^{-1}\{p(\lambda)\chi\} \qquad \text{(Lemma 0.3)}$$

$$= p(\lambda)\{[q(A)]^{-1}\chi\}$$

$$= p(\lambda)\left\{\frac{1}{q(\lambda)}\chi\right\} \qquad \text{(Lemma 0.4)}$$

$$= \frac{p(\lambda)}{q(\lambda)}\chi \quad .$$

Proposition 2. Given A polynomials p and q such that q(A) is nonsingular, and μ an eigenvalue of $[q(A)]^{-1}p(A)$, then there exists an eigenvalue λ of A such that $\mu = p(\lambda)/q(\lambda)$.

Proof: Given a matrix A there always exists a Q such that $Q^{-1}AQ=B$ where B is triangular.

Let $B = \text{tri}(\alpha_1, \ldots, \alpha_n)$ represent a triangular matrix with $\alpha_1, \ldots, \alpha_n$ as its diagonal elements and elements which are of no concern above the main diagonal.

Since $[\text{tri}(\alpha_1, \ldots, \alpha_n)]\,[\text{tri}(\beta_1, \ldots, \beta_n)] = \text{tri}(\alpha_1\beta_1, \ldots, \alpha_n\beta_n)$ and $k[\text{tri}(\alpha_1, \ldots, \alpha_n)] = \text{tri}(k\alpha_1, \ldots, k\alpha_n)$, we know that $p[\text{tri}(\alpha_1, \ldots, \alpha_n)] = \text{tri}(p(\alpha_1), \ldots, p(\alpha_n))$ for any polynomial p. If $\alpha_1 \alpha_2 \cdots \alpha_n \neq 0$, then $\text{tri}(\alpha_1, \ldots, \alpha_n)$ is nonsingular and $[\text{tri}(\alpha_1, \ldots, \alpha_n)]^{-1} = \text{tri}(1/\alpha_1, \ldots, 1/\alpha_n)$. Thus, $[q(B)]^{-1}p(B) = \text{tri}(p(\alpha_1)/q(\alpha_1), \ldots, p(\alpha_n)/q(\alpha_n))$ where $B = \text{tri}(\alpha_1, \ldots, \alpha_n)$. As $[q(B)]^{-1}p(B) = Q^{-1}[q(A)]^{-1}p(A)Q$, the eigenvalues of $[q(B)]^{-1}p(B)$ are the eigenvalues of $[q(A)]^{-1}p(A)$. Thus, if μ is an eigenvalue of $[q(A)]^{-1}p(A)$, then there exists an eigenvalue α_j of A such that $\mu = p(\alpha_j)/q(\alpha_j)$.

CHAPTER 2. THE LINE

The result of this chapter is the creation of a polynomial function $f(A, A^*, H)$ so that the existence of an H for which $f(A, A^*, H) > 0$ is necessary and sufficient for a matrix A to have no eigenvalues on an arbitrary line in the complex plane. As stated in the Introduction, this theory and that of the three chapters to follow is proven from the first part of the Ostrowski-Schneider Main Inertia Theorem. The general theorem of this chapter yields several corollaries which were initially developed directly from the Ostrowski-Schneider theorem. These corollaries correspond to three categories of lines in the complex plane: the lines parallel to the real and imaginary axes, i. e., the lines $y = b$ and $x = a$ (a, b real), and the line which intersects the real axis at a and the imaginary axis at b, i. e. the line $x/a + y/b = 1$. We remark that this line could also be written $\mathrm{Re}\{z\}/a + \mathrm{Im}\{z\}/b = 1$, but it is more easily visualized in terms of its analogue in the Cartesian plane.

As a lemma in proving our general line theorem, we have a result corresponding to the category of lines through the origin. We shall later generalize this lemma to our first theorem by translating the line through the origin by k (k complex).

We remark that in this chapter, we consider r such that $-\infty < r < \infty$ contrary to $0 \leq r < \infty$ as is conventional. This allows us to consider the "line" $\arg z = \theta_0$ which would otherwise be a ray.

<u>Lemma 1.</u> <u>Given</u> B <u>there exists an</u> H <u>such that</u> $\mathrm{Re}\{e^{-i\theta}BH\} > 0$ iff B <u>has no eigenvalues on the line</u> $\arg z = \theta + \pi/2$.

Proof: Let $A = e^{-i\theta}B$ in the Ostrowski-Schneider theorem. Under this substitution, we shall show that the Ostrowski-Schneider theorem implies Lemma 1.

By the Ostrowski-Schneider theorem there exists an H for which $\mathrm{Re}\{e^{-i\theta}BH\} > 0$ (since $\mathrm{Re}\{e^{-i\theta}BH\} = \mathrm{Re}\{AH\}$) iff $A = e^{-i\theta}B$ has no eigenvalue z such that $\mathrm{Re}\{z\} = 0$. Now since the imaginary axis ($\mathrm{Re}\{z\} = 0$) can be written in polar form as $\arg z = \pi/2$ ($-\infty < r < \infty$), we have the following:

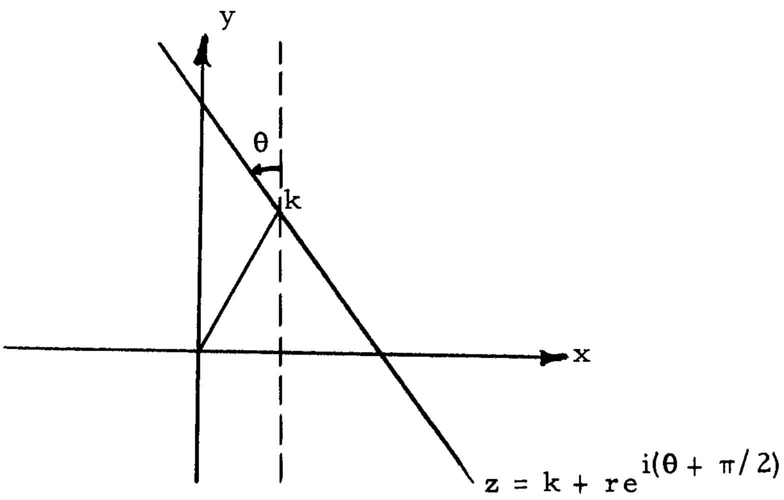
$e^{-i\theta}B$ has no eigenvalues z such that $\mathrm{Re}\{z\} = 0$

iff $e^{-i\theta}B$ has no eigenvalues z such that $\arg z = \pi/2$

iff B has no eigenvalues $ze^{i\theta}$ such that $\arg z = \pi/2$ (Propositions 1 and 2)

iff B has no eigenvalues w such that $\arg w = \theta + \pi/2$. (The argument of a product is the sum of the arguments.)

We are now ready to prove our theorem for the general line in the complex plane. The proof of Theorem 1 uses a matrix substitution $B = C - kI$ which transforms eigenvalues on any line through the origin (which may be expressed as $z = re^{i(\theta + \pi/2)}$) into any line in the complex plane, viz., $z = k + re^{i(\theta+\pi/2)}$. Here k is a complex constant, θ a real constant, and $-\infty < r < \infty$. Geometrically, we might say that the line $z = k + re^{i(\theta+\pi/2)}$ goes through the point k in the complex plane and has a "slope" $\tan^{-1}(\theta + \pi/2)$. The line is pictured below.



<u>Theorem 1.</u> <u>Given an</u> n <u>by</u> n <u>complex matrix</u> C <u>there exists a hermitian</u> H <u>such that</u> $\mathrm{Re}\{e^{-i\theta}CH - e^{-i\theta}kH\} > 0$ iff C <u>has no eigenvalues</u> z <u>such that</u> $z = k + re^{i(\theta + \pi/2)}$ <u>where</u> $-\infty < r < \infty$ k <u>complex,</u> θ <u>real.</u>

Proof: Let $C - kI = B$.

There exists an H such that $\mathrm{Re}\{e^{-i\theta}CH - e^{i\theta}kH\} > 0$

iff there exists an H such that $\mathrm{Re}\{e^{-i\theta}(C - kI)H\} > 0$

iff there exists an H such that $\mathrm{Re}\{e^{-i\theta}BH\} > 0$

iff B has no eigenvalues on the line $\arg z = \theta + \pi/2$ (Lemma 1) iff $\quad C - kI$ has no eigenvalues z such that $\arg z = \theta + \pi/2$

iff $\quad C - kI$ has no eigenvalues z such that $z = re^{i(\theta + \pi/2)}$ $(-\infty < r < \infty)$

iff $\quad C$ has no eigenvalues z such that $z = k + r^{i(\theta + \pi/2)}$ $(-\infty < r < \infty)$

Although the following three results were developed independently by the author, the overall unity of the theory is better expressed when they are derived from Theorem 1 as corollaries. To specialize the general form of the line to a line parallel to the real axis (i.e., $y = ci$, c real) we take $\theta = \pi/2$ and $k = ci$ (c real). For a line parallel to the imaginary axis, we choose $k = c$ and $\theta = 0$. Finally, for a line intersecting the real axis at a and the imaginary axis at b, we let $k = a$ and $\theta = \arctan a/b$.

<u>Corollary 1. Given</u> C <u>there exists an</u> H <u>such that</u> $\operatorname{Im}\{CH\} - cH > 0$ iff C <u>has no eigenvalues on the line</u> $y = ci$. (c real)

Proof: Let $\theta = \pi/2$ and $k = ci$ (c real) in Theorem 1.

Now C has no eigenvalues on the line $y = ci$

iff $\quad C$ has no eigenvalues z such that $z = ci + r$ $(-\infty < r < \infty)$

iff $\quad C$ has no eigenvalues z such that $z = k + re^{i(\pi/2 - \theta)}$

(Here we use Theorem 1 with θ replaced by $-\theta$ which yields a little nicer result.)

iff $\quad$ there exists an H such that $\operatorname{Re}\{e^{i\theta}(C - ciI)H\} > 0$. (Theorem 1) iff there exists an H such that $\mathrm{Re}\{e^{i\theta}(C-kI)H\} > 0$

iff there exists an H such that $\mathrm{Im}\{(C - ciI)H\} > 0$

iff there exists an H such that $\mathrm{Im}\{CH\} - \mathrm{Im}\{ciH\} > 0$

iff there exists an H such that $\mathrm{Im}\{CH\} - cH > 0$.

Corollary 2. Given C there exists an H such that $\mathrm{Re}\{CH\} - cH > 0$ iff C has no eigenvalues on the line $x = c$. (c real)

Proof: Let $k = c$ and $\theta = 0$ in Theorem 1 with θ again replaced by $-\theta$. Then $\mathrm{Re}\{e^{i\theta}CH - e^{i\theta}kH\} = \mathrm{Re}\{CH - cH\} = \mathrm{Re}\{CH\} - cH$

and there exists an H such that $\mathrm{Re}\{e^{i\theta}CH - e^{i\theta}kH\} > 0$

iff there exists an H such that $\mathrm{Re}\{CH\} - cH > 0$

iff C has no eigenvalues z such that $z = k + re^{i(\pi/2-\theta)}$ (Theorem 1)

iff C has no eigenvalues z such that $z = c + re^{\pi i/2}$

iff C has no eigenvalues z such that $z = c + ri$

iff C has no eigenvalues on the line $x = c$.

We now recall that the line in the following corollary intersects the real axis at $\mathrm{Re}\{z\} = a$ and the imaginary axis at $\mathrm{Im}\{z\} = b$.

Corollary 3. Given C there exists an H such that $\mathrm{Re}\{(b-ai)CH\} - abH > 0$ iff C has no eigenvalues on the line $\frac{x}{a} + \frac{y}{b} = 1$.

Proof: Let $k = a$ and $\theta = \tan^{-1} a/b$ in Theorem 1.

$$\mathrm{Re}\{e^{-i\theta}CH - e^{-i\theta}kH\}$$

$$= \mathrm{Re}\{e^{i\tan^{-1}-\frac{a}{b}}CH - e^{i\tan^{-1}-\frac{a}{b}}AH\}$$

$$= \mathrm{Re}\{[\cos(\tan^{-1}-\frac{a}{b}) + i\sin(\tan^{-1}-\frac{a}{b})](C-aI)H\}$$

$$= \mathrm{Re}\{\frac{b-ia}{\sqrt{a^2+b^2}}(C-aI)H\}$$

$$= \frac{1}{\sqrt{a^2+b^2}}\ \mathrm{Re}\{(b-ia)(C-aI)H\}$$

$$= \frac{1}{\sqrt{a^2+b^2}}\ (\mathrm{Re}\{(b-ia)CH\} - \mathrm{Re}\{(b-ia)aH\})$$

$$= \frac{1}{\sqrt{a^2+b^2}}\ (\mathrm{Re}\{b-ia)CH\} - abH)$$

Therefore there exists an H such that $\mathrm{Re}\{e^{-i\theta}CH - e^{-i\theta}kH\} > 0$

iff there exists an H such that $\mathrm{Re}\{(b-ia)CH\} - abH > 0$

iff C has no eigenvalues z such that $z = a + re^{i\tan^{-1}(-\frac{b}{a})}$ (Theorem 1)

iff C has no eigenvalues z such that $z = a + r[\cos(\tan^{-1}(-\frac{b}{a})) + i\sin(\tan^{-1}(-\frac{b}{a}))]$

iff C has no eigenvalues z such that $z = a + r\,\frac{-a+ib}{\sqrt{a^2+b^2}}$

iff C has no eigenvalues on the line $\frac{x}{a} + \frac{y}{b} = 1$.

The last equivalence is verified by showing the following:

1) $z = a + r\frac{-a + ib}{\sqrt{a^2+b^2}}$ is a line

2) the intercepts of the above line are (a, 0) and (0, b).

To show 1): $z = a + r\frac{-a + ib}{\sqrt{a^2+b^2}} = a - \frac{r\ a}{\sqrt{a^2+b^2}} + i\frac{rb}{\sqrt{a^2+b^2}}$

$= a - rc_1 + irc_2$ which is a straight line.

(c_1, c_2 real constants)

To show 2): $\mathrm{Im}\{z\} = 0$ iff $\frac{rb}{\sqrt{a^2+b^2}} = 0$ iff $r\,b = 0$

iff $r = 0$ iff $z = a$.

(We note that $ab \neq 0$ since a and b were assumed to be distinct intercepts.)

$\mathrm{Re}\{z\} = 0$ iff $a - \frac{a\ r}{\sqrt{a^2+b^2}} = 0$ iff $a\sqrt{a^2+b^2} = a\ r$

iff $\sqrt{a^2+b^2} = r$ iff $z = a + \sqrt{a^2+b^2}\ \frac{-a + ib}{\sqrt{a^2+b^2}}$ iff $z = ib$.

CHAPTER 3. THE CIRCLE

The initial result of this chapter is the development of an equivalent theorem (Lemma 2) to the Main Inertia Theorem. That this result is a consequence of the Main Inertia Theorem is sufficient for the purpose of this thesis. However, the stronger result of equivalence is valid. We might add that Lemma 2 is a generalization of Stein's theorem.

Lemma 2. *Given an* n *by* n *complex matrix* B *there exists a hermitian* H *such that* $BHB^* - H > 0$ *iff* B *has no eigenvalues* z *such that* $|z| = 1$.

Proof: Let $A = (B + e^{i\beta}I)^{-1}(B - e^{i\beta}I)$ in the Main Inertia Theorem where β is chosen real such that $-e^{i\beta}$ is not an eigenvalue of B. This implies $(B + e^{i\beta}I)$ is non-singular and hence $(B + e^{i\beta}I)^{-1}$ exists.

For any H, $BHB^* - H > 0$ iff $\operatorname{Re}\{AH\} > 0$ by the following argument:

$$BHB^* - H > 0 \text{ iff } 2\,BHB^* - 2H > 0$$

$$\text{iff} \quad (B + e^{i\beta}I)(AH + HA^*)(B^* + e^{-i\beta}I) > 0.$$

This is verified by the following calculation:

$$(B+e^{i\beta}I)(AH+HA^*)(B^*+e^{-i\beta}I)$$

$$= (B+e^{i\beta}I)\,[(B+e^{i\beta}I)^{-1}(B-e^{i\beta}I)H$$

$$+ H(B^*-\bar{e}^{i\beta}I)(B^*+e^{-i\beta}I)^{-1}]\,(B^*+e^{-i\beta}I)$$

$$= (BH-e^{i\beta}H)(B^*+e^{-i\beta}I) + (B+e^{i\beta}I)(HB^*-e^{i\beta}H)$$

$$= [BHB^*+e^{-i\beta}BH-e^{i\beta}HB^*-H] + [BHB^*-e^{-i\beta}BH+e^{i\beta}HB^*-H]$$

$$= 2\,BHB^* - 2H$$

iff $AH + AH^* > 0$ (Sylvester's theorem since $B+e^{i\beta}I$ is non-singular)

iff $\operatorname{Re}\{AH\} > 0$.

Now there exists an H such that $BHB^* - H > 0$

iff there exists an H such that $\operatorname{Re}\{AH\} > 0$

iff A has no eigenvalues z such that $z = ci$ (c real) (Main Inertia Theorem)

iff B has no eigenvalues w such that

$$ci = (w-e^{i\beta})/(w + e^{i\beta}) \qquad \text{(Propositions 1 and 2)}$$

$$wci + cie^{i\beta} = w-e^{i\beta}$$

$$cie^{i\beta} + e^{i\beta} = w(1-ci)$$

$$w = e^{i\beta}\,(1+ci)/(1-ci)$$

iff B has no eigenvalues w such that $|w| = 1$.

The last equivalence is verified by the following computations:

$$w = e^{i\beta}\;\frac{1+ci}{1-ci}\;\left(\frac{ci+1}{ci+1}\right) = e^{i\beta}\left(\frac{-c^2+2ci+1}{c^2+1}\right)$$

$$w\,\overline{w} = e^{i\beta}\left(\frac{-c^2+2ci+1}{c^2+1}\right)\; e^{-i\beta}\left(\frac{-c^2+1-2ci}{c^2+1}\right)$$

$$= \frac{c^4-2c^2+1+4c^2}{(c^2+1)^2} = \frac{(c^2+1)^2}{(c^2+1)^2} = 1$$

and $\qquad w\overline{w} = 1$ iff $|w| = 1$.

To prove that Lemma 2 implies the Main Inertia Theorem, we let $B = (I-A)^{-1}(I+A)$. Since we assume B to have no eigenvalues of unit modulus, $(I-A)$ will be non-singular. This implies that $(I-A)^{-1}$ exists. This renders the $e^{i\beta}$ factor unnecessary.

The proof follows in a similar manner.

We now generalize Lemma 2.which concerns the unit circle to Lemma 3 which concerns a circle with center at the origin and arbitrary radius r. Lemma 3 in turn is generalized to Theorem 2 which relates to an arbitrary circle in the complex plane. We express this arbitrary circle in terms of its center at the complex number k and its radius r.

<u>Lemma 3</u>. <u>Given</u> C <u>and</u> $r \neq 0$ (real) <u>there exists an</u> H <u>such that</u> $\frac{1}{r^2} CHC^*-H > 0$ iff C <u>has no eigenvalues</u> z <u>such that</u> $|z| = r$.

Proof: Let $B = \frac{1}{r} C$ in Lemma 2.

$$BHB^*-H = \left(\frac{1}{r} C\right) H \left(\frac{1}{r}C\right)^* -H = \frac{1}{r^2}\; CHC^* -H$$

Thus there exists an H such that $BHB^* - H > 0$

iff there exists an H such that $\frac{1}{r^2} CHC^* - H > 0$

and by Lemma 2 there exists an H such that $BHB^* - H > 0$

iff B has no eigenvalues z such that $|z| = 1$

iff $\frac{1}{r}C$ has no eigenvalues z such that $|z| = 1$

iff C has no eigenvalues z such that $|z| = r$.

<u>Theorem 2. Given an</u> n <u>by</u> n <u>complex matrix</u> D, $r \neq 0$ (real), <u>and</u> k <u>complex, there exists a hermitian</u> H <u>such that</u> $\frac{1}{r^2}(D+kI)H(D^*+\bar{k}I) - H > 0$ iff D <u>has no eigenvalues</u> z <u>such that</u> $|z-k| = r$.

Proof: Let $C = D + kI$ in Lemma 3.

$$\frac{1}{r^2} CHC^* - H = \frac{1}{r^2}(D + kI)\, H\, (D + kI)^* - H$$

$$= \frac{1}{r^2}(D + kI)\, H\, (D^* + \bar{k}I) - H$$

Thus there exists an H such that $\frac{1}{r^2} CHC^* - H > 0$ iff there exists an H such that $\frac{1}{r^2}(D + kI)\, H\, (D^* + kI) - H > 0$. This condition holds (Lemma 3)

iff C has no eigenvalues z such that $|z| = r$

iff $(D + kI)$ has no eigenvalues z such that $|z| = r$

iff D had no eigenvalues z such that $|z-k| = r$.

CHAPTER 4. GENERALIZATION OF A CONFORMAL MAPPING

A more general problem may be stated as follows: Given the equation of a well-defined curve in the complex plane, is it possible to find some polynomial function $f(A, A^*, H)$ such that the existence of an H for which $f(A, A^*, H) > 0$ is necessary and sufficient that the matrix A have no eigenvalues on the given curve?

One general approach to this problem is through a matrix generalization of a conformal mapping. This was the approach used in the proof of the equivalence between the Schneider-Ostrowski Main Inertia Theorem and Lemma 2.

To use this method of attack, we start with the given curve. A mapping of a curve for which we have a theorem onto the given curve is developed. (In particular, we might map the imaginary axis or the unit circle onto the given curve.) Next, matrices are substituted for the mapping variables to generalize their analogy as one by one matrices. An attempted proof follows, indicating any difficulties which may arise from the generalized mapping.

It may be noted that this method is not as workable as could be hoped. However, it is certainly useful enough to justify its inclusion in this thesis. Let us consider the following example in which we derive a theorem corresponding to a parabola.

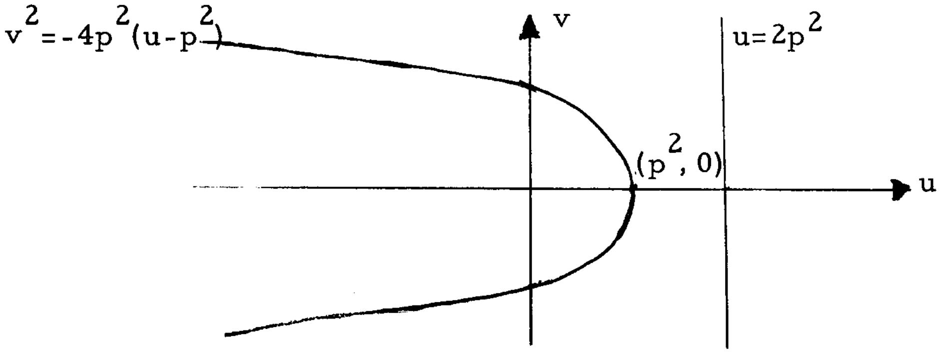
From page 36 of Kober [7] we have $w = z^2$ or $z = \sqrt{w}$ mapping the straight lines $x = \pm\ p\ (p \neq 0)$ onto the parabola $v^2 = -4p^2(u-p^2)$. This may be computationally verified in a straightforward manner.

We now consider $A^{\frac{1}{2}}$, one "square root" of the matrix A. The matrix $A^{\frac{1}{2}}$ always exists but it is not necessarily unique. Gantmacher [6] thoroughly discusses this problem. For our purposes, let $A^{\frac{1}{2}} = B$ be considered to be synonomous with $A = B^2$.

From Corollary 2 we have that given C there exists an H such that $\mathrm{Re}\{CH\} - 2pH > 0$ iff C has no eigenvalues on the line $x = p$. Letting $C = A^{\frac{1}{2}}$ or $A = C^2$ we would like to have the following theorem: Given A there exists an H such that $\mathrm{Re}\{A^{\frac{1}{2}}H\} - 2pH > 0$ iff A has no eigenvalues on the parabola $v^2 = -4p^2(u-p^2)$. However, C having no eigenvalues on the line $x = p\ (p \neq 0)$ does <u>not</u> imply that A has no eigenvalues on the parabola $v^2 = -4p^2(u-p^2)$, since C could have eigenvalues on the line $x = -p$ which would imply that A has one or more eigenvalues on the given parabola. Nevertheless, we can prove the following theorem:

<u>Theorem 3.</u> <u>Given matrices</u> A <u>and</u> B <u>such that</u> $A = B^2$, (i. e., A is some square root of B), <u>there exist hermitian matrices</u> H_1 <u>and</u> H_2 <u>such that</u> $\mathrm{Re}\{BH_1\} + 2pH_1 > 0$ <u>and</u> $\mathrm{Re}\{BH_2\} - 2pH_2 > 0$ iff A <u>has no eigenvalues</u> $w = u + iv$ <u>such that</u> $v^2 = -4p^2(u-p^2)$.

Before proceeding to our proof, we note that the parabola of Theorem 3 has its vertex at $(p^2, 0)$ in the w-plane and opens to the left along the u-axis. The focus of the parabola is at the origin and the directrix is the line $u = 2p^2$ as is shown in the diagram below.



We might also note that the constant "p" used in this discussion is not the usual one in the sense of analytic geometry, namely the distance between the focus and directrix of the parabola.

Proof: Given matrices A and B such that $A = B^2$, there exist hermitian matrices H_1 and H_2 such that $\mathrm{Re}\{BH_1\} + 2pH_1 > 0$ and $\mathrm{Re}\{BH_2\} - 2pH_2 > 0$.

iff B has no eigenvalues on the lines $x = \pm p$ (Corollary 2)

iff $B^2 = A$ has no eigenvalues on the parabola $v^2 = -4p^2(u-p^2)$.

(Propositions 1 and 2)

That the last equivalence follows by Propositions 1 and 2 may be more easily visualized when we consider that (by the contrapositive) the "only if" ($\Rightarrow$) is logically equivalent to "B^2 has eigenvalues on the parabola $v^2 = -4p^2(u-p^2)$ implies that B has eigenvalues on the lines $x = \pm p$," whereas the "if" part ($\Leftarrow$) is logically equivalent to "B has eigenvalues on the lines $x = \pm p$ implies that B^2 has eigenvalues on the parabola $v^2 = -4p^2(u-p^2)$."

We note here that a theorem corresponding to an arbitrary parabola in the complex plane can be derived by a rotation and translation of the parabola of Theorem 3. The method is the same as that used in the development of Theorem 1 from the Main Inertia Theorem.

CHAPTER 5. POLYNOMIAL RESULTS

Previously we have structured our generalizations of the Ostrowski-Schneider Main Inertia Theorem according to curves in the complex plane from which we could exclude eigenvalues. This chapter gives two results which were motivated by specifying a polynomial function $f(A, A^*, H)$, namely the n^{th} power of a given matrix A and the general quadratic in A. Our results are derived from the conditions of $\operatorname{Im}\{A^n H\} > 0$ and $\operatorname{Im}\{(\alpha A^2 + \beta A + \gamma I)H\} > 0$. In the first case eigenvalues are excluded from the family of n rays through the n roots of unity and in the second case from an hyperbola or two intersecting lines in the degenerate case.

Lemma 4. *Given C there exists an H such that $\operatorname{Im}\{CH\} > 0$ iff C has no real eigenvalues.*

Proof: Let $c = 0$ in Corollary 1. The proof is immediate.

Theorem 4. *Given an n by n complex matrix A there exists a hermitian H such that $\operatorname{Im}\{A^n H\} > 0$ iff A has no eigenvalues z on the n rays* $\arg z = \operatorname{cis}\ 2\pi \ell / n \quad (\ell = 0, 1, \ldots, n-1)$.

Proof: Let $C = A^n$ in Lemma 4.

A^n has no real eigenvalues

iff $\quad A$ has no eigenvalues z such that $z = \sqrt[n]{k}$ (k real) (Propositions 1 and 2)

iff $\quad A$ has no eigenvalues z such that $\arg z = \operatorname{cis} 2\pi\ell/n$ ($\ell = 0, 1, \ldots, n-1$ and $0 \leq r \leq \infty$) (Churchill [4, p. 14]).

<u>Theorem 5. Given an n by n complex matrix A and complex constants</u> $\alpha = a + bi \neq 0$, $\beta = c + di$, <u>and</u> $\gamma = e + fi$, <u>there exists a hermitian H such that</u> $\operatorname{Im}\{\alpha A^2 + \beta A + \gamma I)H\} > 0$ iff A <u>has no eigenvalues</u> $z = x + iy$ <u>on the hyperbola</u> $bx^2 + 2axy - by^2 + dx + cy + f = 0$.

Proof: There exists an H such that $\operatorname{Im}\{(\alpha A^2 + \beta A + \gamma I)H\} > 0$

iff $\quad \alpha A^2 + \beta A + \gamma I$ has no real eigenvalues (Lemma 4)

iff $\quad \alpha z^2 + \beta a + \gamma \neq r$ for all eigenvalues z of A (Propositions 1 and 2)

iff $\quad z$ does not lie on the hyperbola $bx^2 + 2axy - by + dx + cy + f = 0$.

The last equivalence is verified as follows: Since $\alpha = a + bi$, $\beta = c + di$, and $\gamma = e + fi$, we may express $\alpha z^2 + \beta z + \gamma \neq r$ as

$$(a + bi)(x^2 - y^2 + 2xyi) + (c + di)(x + yi) + (e + fi) \neq r.$$

Consider the related pair of equations:

(i) $ax^2 - 2bxy - ay^2 + cx + dy + e = r$

(ii) $bx^2 + 2axy = by^2 + dx + cy + f = 0.$

The first equation (i) above is always satisfied by some real r. Thus we are left with (ii) which is a hyperbola since

$B^2 - 4AC = 4a^2 + 4b^2 > 0$. Conversely, for any $z = x + iy$ lying on the hyperbola (ii) an r may be found so that (i) is satisfied. Thus the locus of points from which eigenvalues are excluded is the hyperbola (ii).

CHAPTER 6. GENERALIZATION OF THE SECOND PART OF THE MAIN INERTIA THEOREM

This chapter generalizes the second part of the Ostrowski-Schneider Main Inertia Theorem which may be stated as follows: If $\operatorname{Re}\{AH\} > 0$, then $\operatorname{In} A = \operatorname{In} H$. The inertia of an n by n complex matrix A has been defined as $\operatorname{In} A = (\pi, \nu, \delta)$ where π and ν are the number of eigenvalues of A in the (open) right and left half planes respectively and δ is the number of eigenvalues on the imaginary axis.

Now, since H is hermitian, all its eigenvalues are real. Thus if given A there exists an H such that $\operatorname{Re}\{AH\} > 0$, we have the number of eigenvalues of A in the (open) right half-plane equal to the number of positive (real) eigenvalues of H and the number of eigenvalues of A in the (open) left half-plane equal to the number of negative (real) eigenvalues of H. Moreover, the first part of the Ostrowski-Schneider theorem stating that A has no eigenvalues on the imaginary axis implies that H has no zero eigenvalues.

For any curve dividing the complex plane into two disjoint regions for which there exists a theorem of the type developed in the body of this thesis, we now have a theorem analogous to the second part of the Main Inertia Theorem. The conformal mapping which takes the imaginary axis onto the given curve guarantees that the (open) right half-plane maps onto one region and that the (open) left half-plane maps onto the other region. It will have been previously proven that there are no eigenvalues on the given curve. Thus the matrix in our "derived" theorem will have the same number of eigenvalues in one region as H has positive eigenvalues and the same numbers in the other region as H has negative eigenvalues.

Let us consider Lemma 2 as an example of a way in which all the lemmas, theorems, and corollaries of this thesis could have been written up so as to "completely" generalize the Schneider-Ostrowski theorem.

<u>Lemma 2'</u> (i) <u>Given an</u> n <u>by</u> n <u>complex matrix</u> B <u>there exists a hermetian</u> H <u>such that</u> $BHB^* - H > 0$ iff B <u>has no eigenvalues</u> z <u>such that</u> $|z| = 1$. (ii) <u>Further, the number of eigenvalues</u> z <u>of</u> B <u>such that</u> $|z| > 1$ (i. e., <u>those eigenvalues of</u> B <u>lying outside the unit circle</u>) <u>is equal to the number of positive eigenvalues of</u> H <u>and the number of eigenvalues</u> z <u>of</u> B <u>such that</u> $|z| < 1$ (i. e., <u>those lying within the unit circle</u>) <u>is equal to the number of negative eigenvalues of</u> H.

Our result is immediate under the mapping $w = \frac{z - e^{i\beta}}{z + e^{i\beta}}$ or $z = e^{i\beta}\frac{w+1}{1-w}$ where A consists of elements from the w-plane and B consists of elements from the z-plane. It is easily shown that the mapping $z = e^{i\beta}\frac{w+1}{1-w}$ maps the (open) left half w-plane inside the unit circle in the z-plane and the (open) right half of the w-plane outside the unit circle in the z-plane. Consider the point $w = 2$. (We could consider any other point not on the imaginary axis.) Since $w = 2$ is in the right half w-plane and it maps outside the unit circle in the z-plane under the conformal mapping $z = e^{i\beta}\,\frac{w+1}{1-w}$, our assertion is proven.

CHAPTER 7. GENERALIZATIONS OF THE DRAZIN-HAYNSWORTH THEOREMS

In their paper Criteria for the Reality of Matrix Eigenvalues [5], Drazin and Haynsworth initially prove the following theorem:

D-H Theorem 1. A set of m linearly independent eigenvectors of A all corresponding to real eigenvalues exists iff there exists a positive semi-definite hermitian matrix H of rank m such that $AH = HA^*$.

By replacing A by iA in their first theorem, they have the following result:

D-H Theorem 2. A set of m linearly independent eigenvectors of A all corresponding to purely imaginary eigenvalues exists iff there exists a positive semi-definite hermitian matrix H of rank m such that $AH = -HA^*$.

Later in their paper they state: "The following somewhat analogous result does not seem to be deducible directly from (D-H) Theorem 1 but can be proved by closely parallel arguments."

D-H Theorem 3. A set of m linearly independent eigenvectors of A all corresponding to eigenvalues of absolute value one exists iff there exists a positive semi-definite hermitian matrix H of rank m such that $AHA^* = H$.

A proof analogous to the one in this thesis of the equivalence of Lemma 2 and the Ostrowski-Schneider Main Inertia Theorem may be used to prove the equivalence of the Drazin-Haynsworth Theorems 2 and 3. We include it here.

Proof: We first assume D-H Theorem 2. Let B be given. As in

Chapter 2, we choose β such that $B + e^{i\beta}$ is non-singular and we define

(2) $$A = (B + e^{i\beta} I)^{-1} (B - e^{-i\beta} I).$$

Now by the calculation (1) in Chapter 2, $BHB^* - H = 0$ for some $H \geq 0$ of rank m

iff $AH + HA^* = 0$ for some $H \geq 0$ of rank m

iff A has at least m elementary divisors corresponding to imaginary eigenvalues (D-H Theorem 2)

iff B has at least m elementary divisors corresponding to eigenvalues of absolute value one (from (2), since A and B have the same eigenvectors).

Similarly D-H Theorem 3 may be used to prove D-H Theorem 2 which demonstrates their equivalence.

As an example of what another theorem would look like, let us just state the Drazin-Haynsworth analogue of Theorem 1 of this thesis.

<u>Theorem 1'.</u> <u>A set of</u> m <u>linearly independent eigenvectors of</u> A <u>all corresponding to eigenvalues</u> z <u>such that</u> $z = k + re^{i(\theta + \pi/2)}$ <u>exists iff there exists a positive semi-definite hermitian matrix</u> H <u>of rank</u> m <u>such that</u> $\mathrm{Re}\{e^{-i\theta} AH\} = \mathrm{Re}\{e^{-i\theta} kH\}$.

BIBLIOGRAPHY

1. Ahlfors, Lars V. Complex analysis. New York, McGraw-Hill, 1953. 247p.

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

3. Carlson, David and Hans Schneider. Inertia theorems for matrices: the semi-definite case. Journal of Mathematical Analysis and Applications 6:430-446. 1963.

4. Churchill, Ruel V. Complex variables and applications. New York, McGraw-Hill, 297p.

5. Drazin, Michael P. and Emilie V. Haynsworth. Criteria for the reality of martix eigenvalues. Mathematische Zeitschrift 78:449-452. 1962.

6. Gantmacher, F.R. Theory of matrices. vol. 1. New York, Chelsea, 1959. 374p.

7. Kober, H. Dictionary of conformal representations. New York, Dover, 1957. 208p.

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

9. Stein, P. Some general theorems on iterants. Journal of Research of the National Bureau of Standards 48:82-83. 1952.

10. Taussky, Olga. A remark on a theorem of Lyapunov. Journal of Mathematical Analysis and Applications 2:105-107. 1961.

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

12. Taussky, Olga. Matrices C with $C^n \to 0$. Journal of Algebra 1:5-10. 1964.