ANL-6456

ANL-6456

Argonne National Laboratory

ELEMENTARY DIVISORS AND SOME PROPERTIES OF THE LYAPUNOV MAPPING $X \to AX + XA^*$

by

Wallace Givens

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

ANL-6456
Mathematics and
Computers
(TID-4500, 16th Ed.,
Amended)
AEC Research and
Development Report

ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Argonne, Illinois

ELEMENTARY DIVISORS AND SOME PROPERTIES OF THE LYAPUNOV MAPPING $X \longrightarrow AX + XA^*$

by

Wallace Givens

Professor of Mathematics and Engineering Sciences,
Northwestern University
Consultant, Argonne National Laboratory

November 1961

Operated by The University of Chicago
under
Contract W-31-109-eng-38

ACKNOWLEDGEMENT

The preparation of this paper was stimulated and aided by the participation of the author in the Matrix Computation Conference held at Gatlinburg, Tennessee, in April, 1961, under the sponsorship of the Society for Industrial and Applied Mathematics with financial support from the National Science Foundation and the Atomic Energy Commission.

TABLE OF CONTENTS

ELEMENTARY DIVISORS AND SOME PROPERTIES OF THE LYAPUNOV MAPPING

$X \to AX + XA^*$

Wallace Givens

INTRODUCTION

In his monograph "The General Problem of Stability of Motion," Lyapunov established a connection between the location of the eigenvalues of a general n by n real matrix and the signature of a quadratic form. A clear treatment of these results is given by Gantmacher ([18], Ch. XV, § 5) but there does not appear to be available in the literature a full discussion which relies wholly on algebraic techniques and does not employ the connection with the solution of a system of linear differential equations. In section 12 we explore, using the simplest algebraic methods, the Lyapunov result and some of the recent interesting extensions of it by Taussky (cf. [60] and [61]).

The matrix correspondence which is the subject of this paper arises if one differentiates the hermitian (or, in the real case, the quadratic) form uGu^*, where $G = G^*$ = the conjugate transpose of G and u is a row vector, and makes use of the differential equation

$$\frac{du}{dt} = u\,A \quad , \tag{1}$$

for A an n by n matrix, to get

$$\frac{d(uGu^*)}{dt} = u\,(AG + GA^*)u^* \equiv uHu^* \quad . \tag{2}$$

We shall refer to the mapping of hermitian matrices

$$G \to H = AG + GA^* \tag{3}$$

as the Lyapunov mapping determined by A and the central result is that

> if H is negative definite for some positive definite G, then every eigenvalue of A has negative real part (A is stable) and, conversely, if A is stable the correspondence is one-to-one and for H negative definite G is positive definite.

For the complex field there is little gain in extending (3) by replacing G by a matrix X which is not assumed to be hermitian since this merely extends the real linear mapping (3) of a real n^2 dimensional vector space (of matrices G) to a complex space of the same dimension. This is not true if A is real and G is symmetric, for then we are extending a linear mapping of a real vector space of $1/2\ n\ (n+1)$ dimensions to one of n^2 dimensions which, however, decomposes into the direct sum of mappings of symmetric matrices and of skewsymmetric ones. Thus, over both the real and the complex fields it is interesting to study

$$X \to Y = AX + XA^T \quad , \tag{4}$$

where A^T is the transpose of A, and this is done in section 9.

The inadequacy of the algebraic treatment of the Lyapunov result in the literature is to some extent paralleled (and possibly caused) by the fact that the standard works on linear algebra (cf. Jacobson [32] or Bourbaki [8]) do not include the elementary divisor structure of so fundamental a mapping as

$$\begin{aligned} X \to Y &= AX + XB^T \\ &\equiv (A \otimes 1 + 1 \otimes B)\,\vec{X} \quad , \end{aligned} \tag{5}$$

where the tensor product of matrices is denoted by the symbol $\otimes$ and permits us to regard the n by n matrix X as an n^2 component vector $\vec{X}$ (cf. MacDuffee [39], Chapters VII and VIII). This is despite the fact that the main results have been known since they were established by Littlewood, Roth, Rutherford and Williamson in the years 1931 to 1937. The essential simplicity of the situation is to some extent obscured by the special devices those authors employed. In sections 4 and 5 we give an independent treatment, again choosing to depend only on elementary matrix arguments. A good case could be made for using the concepts of quotient spaces in this discussion since the arguments hinge on

1. a simple combinatoric count which establishes the dimension of the quotient of the null space of the r^{th} power of a nilpotent matrix relative to that of the (r+1)st power, and

2. a proof that natural mappings between such spaces have in each case the maximum possible dimension (that is, is either onto or injective).

The expert should have no difficulty rephrasing our results in this language but we have refrained from using it since simpler concepts are sufficient.

The expert as well as the relative newcomer to matrix algebra may wish to omit sections 1 and 2 where we establish the relation of the matrices used to mappings of vector spaces and thus avoid the naive treatment of the matrices as elements of a ring.

The primary emphasis in this paper is on establishing easy and convenient techniques for a further study of the Lyapunov mapping and completeness has uniformly been preferred to brevity. Since it is the inverse map (Y → X which appears to be most useful for the development of new techniques for computing eigenvalues, it seems probable that one should restrict A to be of some special type (say, tridiagonal) and make full use of the resulting special properties of the Lyapunov mapping, and this is intended as a future application of the present work.

Another mapping which is of importance in the differential equations development is the so-called Lyapunov transformation (cf. [18], Chapter XIV, § 2):

$$A(t) \to B(t) = L^{-1}AL - L^{-1}\frac{dL}{dt} \quad , \tag{6}$$

where L(t), when subjected to certain continuity and boundedness conditions, is called a Lyapunov matrix ([18], vol. II, p. 117). That an algebraic treatment of such a mapping is possible is not obvious but it may be possible by applying results of Jacobson [33]. Results in this direction could be of much interest in indicating genuinely novel methods for the computation of eigenvalues.

Since the mapping (3) can be written in the form

$$G \to H = \mathrm{Re}(AG) \quad , \tag{7}$$

where

$$AG = \mathrm{Re}(AG) + i\mathcal{I}m\,(AG) \tag{8}$$

is the decomposition of AG into the sum of its hermitian and skewhermitian parts, there is an obvious connection of the signature of H with the field of values of A relative to a metric given by G and this is studied in section 13.

Finally, in sections 14 and 15, two explicit solutions for Y in terms of X are proved, one generalizing a formula to be found in Bellman ([6], p. 243) and the other generalizing (and correcting a factor 2) one given by Hahn ([24], p. 23). In each case a restriction is imposed on A which guarantees the uniqueness of the solution for X. The original papers establishing the solution in section 15 were not readily available (cf. Bedel'baev [5] and Malkin [40]) and the proof given is independent of them.

1. The Product of A and M

Let A be an n by n matrix defining (with respect to some given basis) a linear transformation of a vector space (of column vectors) $\mathcal{V}$ to itself:

$$x \rightarrow y = Ax \quad . \tag{1.1}$$

With respect to a change of basis in $\mathcal{V}$ under which x is replaced by $\tilde{x}$ with

$$\tilde{x} = Tx \quad , \tag{1.2}$$

the matrix A is replaced by

$$\tilde{A} = TAT^{-1} \quad , \tag{1.3}$$

which describes the same linear transformation relative to the new basis:

$$\tilde{y} = \tilde{A}\tilde{x} \quad . \tag{1.4}$$

Introducing the space $\mathcal{V}^{\cup}$ of vectors dual to $\mathcal{V}$ and designating its elements by row vectors u,v,..., we have from the definition of the dual space that there is a unique scalar associated with a u in $\mathcal{V}^{\cup}$ and an x in $\mathcal{V}$. Choosing the basis in $\mathcal{V}^{\cup}$ canonical with respect to the basis in use in $\mathcal{V}$ means that the scalar associated with u and x is

$$ux = u_1x^1 + u_2x^2 + \ldots + u_nx^n \quad , \tag{1.5}$$

where

$$u = (u_1, u_2, \ldots, u_n) \tag{1.6}$$

and

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \quad . \tag{1.7}$$

If we continue to associate a canonical basis in $\mathcal{V}^{\cup}$ with the new basis in $\mathcal{V}$, we must have

$$ux = \tilde{u}\,\tilde{x} \text{ , for every } x, \tag{1.8}$$

and hence

$$u = \tilde{u}\,T \tag{1.9}$$

describes the corresponding change of basis in $\mathcal{V}^{\cup}$.

To obtain an explicitly given mapping of $\mathcal{V}^{U}$ to $\mathcal{V}$ we need to have some way of converting a column vector to a row vector and do this simply by introducing the transpose operation in case the field of scalars is the real numbers and the operation of the conjugate transpose when the complex numbers are the coefficient field. Thus

$$u \to x = u^* \tag{1.10}$$

means $x_i = u_i$ in the first case and $x_i = \bar{u}_i$ in the second case. [We could also consider $x_i = u_i$ in the complex number case but then the needed concept of positive definiteness (cf. infra) would be lacking.] Where we wish to emphasize the real number case we write $x = u^T$ and use A^* for the conjugate transpose (and, in the real number case, also for the transpose) and A^T for the transpose, as appropriate.

An arbitrary n by n matrix M defines, together with the "star" operation, a mapping of $\mathcal{V}^{U}$ to $\mathcal{V}$ by the formula

$$u \to x = Mu^* \quad . \tag{1.11}$$

[Since the mapping $u \to x = u^*$ is not canonical (that is, requires an arbitrary choice of bases before it is defined), it is more perspicuous to introduce the class of inner products

$$(v,u) \equiv v\ M\ u^* \tag{1.12}$$

in $\mathcal{V}^{U}$ and to note that the linearity properties of the inner product, axiomatized in a familiar way (cf. Bourbaki [9]), implies that Mu^* may be regarded as a vector of $(\mathcal{V}^{U})^{U} = \mathcal{V}$ (where the equality sign requires the identification of isomorphic spaces).]

If we change the basis in $\mathcal{V}$ and make the induced change (1.9) in $\mathcal{V}^{U}$, we observe that

$$\tilde{y} = \tilde{M}\,\tilde{u}^* \quad , \tag{1.13}$$

provided

$$\tilde{M} = TMT^* \quad . \tag{1.14}$$

Although it is meaningless to multiply two matrices M describing mappings (1.11) since two different spaces are involved, the mapping $u \to x$ can be followed by the mapping $x \to y$ given by (1.1):

$$u \to x = Mu^* \to y = Ax = AMu^* \quad , \tag{1.15}$$

or

$$u \to y = N_1 u^* \text{ with } N_1 = AM \quad . \tag{1.16}$$

After a change of basis in $\mathcal{V}$ and $\mathcal{V}^{U}$,

$$\begin{aligned} \tilde{u} \to \tilde{y} = \tilde{N}_1 \tilde{u}^* \text{ with } \tilde{N}_1 &= \tilde{A}\,\tilde{M} = (TAT^{-1})(TMT^*) \\ &= TN_1T^* \quad . \end{aligned} \tag{1.17}$$

We refer to (1.15) as the <u>product</u> of A and M and note that "M" is followed by "A."

Shifting the focus of attention from the vector spaces being mapped to the mappings themselves, we observe that for a fixed matrix A

$$M \to N_1 = AM \tag{1.18}$$

describes a transformation of the class of mappings (1.11) of $\mathcal{V}^{U}$ to $\mathcal{V}$ into itself and that this linear transformation has an inverse if and only if A does. [The last formula appears to describe the "left regular" representation of a ring by its "left multiplication" transformation but it does not since the "sum" of A and M has no invariant sense under change of bases and so is undefined; hence, A and M are not members of the same ring.] The matrices M constitute an n^2 dimensional vector space since the addition of matrices M_1 and M_2 and the multiplication of an M by a scalar has invariant sense under (1.14). From this point of view (1.18) becomes $N_1 = AM\,1_n$, or

$$\vec{M} \to \vec{N}_1 = (A \otimes 1_n)\,\vec{M} \quad , \tag{1.19}$$

where $A \otimes 1_n$ is the tensor product (cf. N. Bourbaki [8]) of A and the identity matrix and has n^2 rows and n^2 columns. Introducing indices,

$$n^{(1)}_{ij} = \sum_{p,q=1}^{n} (a_{ip}\ \delta_{jq})\ m_{pq} \tag{1.20}$$

so

$$(A \otimes 1)_{(i,j),(p,q)} = a_{ip}\ \delta_{jq}, \tag{1.21}$$

where

$$\delta_{jq} = 1 \text{ if } j = q \text{ and } 0 \text{ otherwise} \tag{1.22}$$

is the Kronecker delta and $A \otimes 1$ has the pair (i,j) as row index and (p,q) as column index. If (i,j) and (p,q) are to be linearly ordered, the lexicographic order with j regarded as the "first" letter would lead to the representation of $A \otimes 1$ in the convenient form (illustrated in detail for n = 3) of (1.23). (A blank space is to be filled by a zero.) This enumeration of the elements of $A \otimes 1$ has been called (cf. MacDuffee [39], p. 81) the "left direct product" and a "right direct product" is obtained if the lexicographic ordering assumes i as the first letter of the "word" (i,j). Then the enumeration of the elements of $A \otimes 1$ yields (1.24).

$$
\begin{array}{cc||ccccccccc||}
 & & & q=1 & & & q=2 & & & q=3 & \\
 & & p=1 & p=2 & p=3 & p=1 & p=2 & p=3 & p=1 & p=2 & p=3 \\
\hline
 & i=1 & a_{11} & a_{12} & a_{13} & & & & & & \\
j=1 & i=2 & a_{21} & a_{22} & a_{23} & & & & & & \\
 & i=3 & a_{31} & a_{32} & a_{33} & & & & & & \\
 & i=1 & & & & a_{11} & a_{12} & a_{13} & & & \\
j=2 & i=2 & & & & a_{21} & a_{22} & a_{23} & & & \\
 & i=3 & & & & a_{31} & a_{32} & a_{33} & & & \\
 & i=1 & & & & & & & a_{11} & a_{12} & a_{13} \\
j=3 & i=2 & & & & & & & a_{21} & a_{22} & a_{23} \\
 & i=3 & & & & & & & a_{31} & a_{32} & a_{33}
\end{array}
$$

$A \otimes 1 = \| a_{ip}\, \delta_{jq} \|$ ordered by $(i', j') < (i, j)$ if $j' < j$ or $j' = j$ and $i' < i$.

(1.23)

$$
\begin{array}{cc||ccccccccc||}
 & & & p=1 & & & p=2 & & & p=3 & \\
 & & q=1 & q=2 & q=3 & q=1 & q=2 & q=3 & q=1 & q=2 & q=3 \\
\hline
 & j=1 & a_{11} & & & a_{12} & & & a_{13} & & \\
i=1 & j=2 & & a_{11} & & & a_{12} & & & a_{13} & \\
 & j=3 & & & a_{11} & & & a_{12} & & & a_{13} \\
 & j=1 & a_{21} & & & a_{22} & & & a_{23} & & \\
i=2 & j=2 & & a_{21} & & & a_{22} & & & a_{23} & \\
 & j=3 & & & a_{21} & & & a_{22} & & & a_{23} \\
 & j=1 & a_{31} & & & a_{32} & & & a_{33} & & \\
i=3 & j=2 & & a_{31} & & & a_{32} & & & a_{33} & \\
 & j=3 & & & a_{31} & & & a_{32} & & & a_{33}
\end{array}
$$

$A \otimes 1 = \| a_{ip}\, \delta_{jq} \|$ ordered by $(i', j') < (i, j)$ if $i' < i$ or $i' = i$ and $j' < j$.

(Standard lexicographic ordering.)

(1.24)

For most purposes it leads merely to obfuscation to impose an arbitrary linear order (of which only two of a possible factorial n^2 have been illustrated) but it is sometimes convenient in considering numerical problems to inspect the displayed matrices given.

We refer to the last ordering as the "standard lexicographic ordering": $(i',j') < (i,j)$ if $i' < j'$ or if $i' = i$ and $j' < j$.

2. The Product of M and A

The mapping $x \to y = Ax$ of $\mathcal{V}$ to $\mathcal{V}$ induces the mapping

$$v \to u = vA \tag{2.1}$$

of $\mathcal{V}^{\upsilon}$ to $\mathcal{V}^{\upsilon}$ since

$$vy = vAx = ux \tag{2.2}$$

for arbitrary v and x. Applying the star mapping of $\mathcal{V}^{\upsilon}$ to $\mathcal{V}$ introduced in (1.10) yields

$$u^* = A^* \, v^* \tag{2.3}$$

which is a mapping $v^* \to u^*$ of a vector v^* of $\mathcal{V}$ to u^* of $\mathcal{V}$ with matrix A^*. It is important to note that since (1.10) is not invariant under change of bases,

$$(\tilde{A})^* = (TAT^{-1})^* \neq TA^*T^{-1} = \widetilde{(A^*)} \tag{2.4}$$

for an unrestricted (nonsingular) T. Indeed, equality would hold only if

$$A^*(T^*T) = (T^*T)\, A^* \tag{2.5}$$

and, using only real and symmetric A, one easily proves that this will hold for all A if and only if

$$T^*T = \rho \, 1_n \tag{2.6}$$

is a scalar matrix. Taking the trace shows ρ to be real and positive so to within a factor T is <u>unitary</u> (or, in the real case, <u>orthogonal</u>):

$$(\rho^{-1/2}\, T)^* \, (\rho^{-1/2}\, T) = 1_n \quad . \tag{2.7}$$

If we follow the mapping (2.1) of $\mathcal{V}^{\upsilon}$ to $\mathcal{V}^{\upsilon}$ by the mapping (1.11) of $\mathcal{V}^{\upsilon}$ to $\mathcal{V}$, we get the mapping

$$v \to u = vA \to x = Mu^* = MA^* \, v^* \quad , \tag{2.8}$$

or

$$v \to x = (MA^*)\, v^* \equiv N_2 \, v^* \text{ with } N_2 = MA^* \quad . \tag{2.9}$$

This product (2.8) of M and A is the result of following "A" by "M" in contradistinction to the first product where "M" was followed by "A."

Just as in the case of the first product, we observe that

$$M \to N_2 = MA^* \tag{2.10}$$

is a linear mapping of the n^2 dimensional vector space of matrices M to itself and that the inverse mapping exists if and only if A (and hence A*) is non-singular. Introducing indices,

$$n^{(2)}_{ij} = \sum_{p,q=1}^{n} (\delta_{ip}\, \bar{a}_{jq})\, m_{pq} \tag{2.11}$$

or, in terms of the tensor product* of matrices

$$\vec{N}_2 = (1 \otimes \bar{A})\, \vec{M} \ . \tag{2.12}$$

Writing out the elements of the matrix $1 \otimes \bar{A}$ for n=3 in the <u>second</u> lexicographic ordering given above (that of (1.24), <u>not</u> that of (1.23)) gives:

$$\begin{array}{cc|ccc|ccc|ccc|}
 & & & p=1 & & & p=2 & & & p=3 & \\
 & & q=1 & q=2 & q=3 & q=1 & q=2 & q=3 & q=1 & q=2 & q=3 \\
\hline
 & j=1 & \bar{a}_{11} & \bar{a}_{12} & \bar{a}_{13} & & & & & & \\
i=1 & j=2 & \bar{a}_{21} & \bar{a}_{22} & \bar{a}_{23} & & & & & & \\
 & j=3 & \bar{a}_{31} & \bar{a}_{32} & \bar{a}_{33} & & & & & & \\
\hline
 & j=1 & & & & \bar{a}_{11} & \bar{a}_{12} & \bar{a}_{13} & & & \\
i=2 & j=2 & & & & \bar{a}_{21} & \bar{a}_{22} & \bar{a}_{23} & & & \\
 & j=3 & & & & \bar{a}_{31} & \bar{a}_{32} & \bar{a}_{33} & & & \\
\hline
 & j=1 & & & & & & & \bar{a}_{11} & \bar{a}_{12} & \bar{a}_{13} \\
i=3 & j=2 & & & & & & & \bar{a}_{21} & \bar{a}_{22} & \bar{a}_{23} \\
 & j=3 & & & & & & & \bar{a}_{31} & \bar{a}_{32} & \bar{a}_{33} \\
\end{array} \tag{2.13}$$

---

*For definitions and properties of the tensor product see MacDuffee ([39], Ch.VII) or Bourbaki [8]. It is to be noted that the mapping (2.10) has matrix $1 \otimes \bar{A}$ and not $1 \otimes A^*$ under the usual notational conventions. In agreement with this, we have $M \to A_1 M B_1^* \to A_2(A_1 M B_1^*)B_2^* = (A_2A_1)\, M (B_2B_1)^*$, which becomes $\vec{M} \to (A_1 \otimes \bar{B}_1)\, \vec{M} \to (A_2 \otimes \bar{B}_2)\, (A_1 \otimes \bar{B}_1)\, \vec{M} = [A_2A_1) \otimes (\bar{B}_2\bar{B}_1)]\, \vec{M}$ in the tensor notation.

3. Eigenvalues of the Sum and Product

Since both $M \to N_1 = AM$ and $M \to N_2 = MA^*$ are linear mappings on the same vector space to itself, their sum and their product are defined.

The associative law,

$$(AM)A^* = A(MA^*) \quad , \tag{3.1}$$

holding for all M, when stated in the language of n^2 by n^2 matrices becomes the commutative law

$$(A \otimes 1)\ (1 \otimes \bar{A}) = (1 \otimes \bar{A})\ (A \otimes 1) \tag{3.2}$$

and their common value is denoted by

$$A \otimes \bar{A} \quad . \tag{3.3}$$

This can also be verified (unnecessarily and tediously) by multiplying the matrices of (1.24) and (2.13) in the two orders, getting in both cases the result (for n = 3):

$$\begin{array}{ll|c|c|c|c|c|c|c|c|c|}
 & & & p=1 & & & p=2 & & & p=3 & \\
 & & q=1 & q=2 & q=3 & q=1 & q=2 & q=3 & q=1 & q=2 & q=3 \\
 & j=1 & a_{11}\bar{a}_{11} & a_{11}\bar{a}_{12} & a_{11}\bar{a}_{13} & a_{12}\bar{a}_{11} & a_{12}\bar{a}_{12} & a_{12}\bar{a}_{13} & a_{13}\bar{a}_{11} & a_{13}\bar{a}_{12} & a_{13}\bar{a}_{13} \\
i=1 & j=2 & a_{11}\bar{a}_{21} & a_{11}\bar{a}_{22} & a_{11}\bar{a}_{23} & & & \text{etc.} & & & \\
 & j=3 & a_{11}\bar{a}_{31} & a_{11}\bar{a}_{32} & a_{11}\bar{a}_{33} & & & & & & \\
 & j=2 & & & & & & & & & \\
i=2 & j=2 & & (a_{21}\bar{A}) & & & (a_{22}\bar{A}) & & & (a_{23}\bar{A}) & \\
 & j=3 & & & & & & & & & \\
 & j=1 & & & & & & & & & \\
i=3 & j=2 & & (a_{31}\bar{A}) & & & (a_{32}\bar{A}) & & & (a_{33}\bar{A}) & \\
 & j=3 & & & & & & & & &
\end{array} \tag{3.4}$$

The commutativity of $A \otimes 1$ and $1 \otimes \bar{A}$ imply that the eigenvalues of any polynomial in them is the same polynomial in their eigenvalues. In particular, if

$$\alpha_i \text{ are the eigenvalues of } A \tag{3.5}$$

then

$$\bar{\alpha}_i \text{ are the eigenvalues of } \bar{A} \ , \tag{3.6}$$

$\alpha_i\bar{\alpha}_j$ are the eigenvalues of $A \otimes \bar{A}$ (3.7)

and

$\alpha_i + \bar{\alpha}_j$ are the eigenvalues of $A \otimes 1 + 1 \otimes \bar{A}$ (3.8)

where $i,j = 1,2,\ldots,n$ independently, since it is obvious from (1.23) and (2.13) that $A \otimes 1$ has the α_i, each repeated n times, for its eigenvalues and $1 \otimes \bar{A}$ has $\bar{\alpha}_i$, each of multiplicity n. [A better proof is given by observing that if $Ax = \alpha x$ then $[A \otimes 1]\,(xu) = \alpha\,(xu)$, where u is an arbitrary n-component row vector so that the xu constitute an n dimensional vector subspace of eigenvectors, each with eigenvalue α, in the n^2 dimensional space of matrices M.]

We shall sometimes be interested in a quite explicit form of $A \otimes 1 + 1 \otimes \bar{A}$ in cases in which A is taken in some special form by a suitable choice of bases in $\mathcal{V}$ and $\mathcal{V}^U$. Three such choices are

the triple diagonal form in which $a_{ij} = 0$ for $|j-i| > 1$ which can be achieved over the real or complex field; (3.9)

the "two diagonal form," of which the Jordan canonical form is a special case, defined by $a_{ij} = 0$ unless $j-i = 0$ or 1 (or, $i-j = 0$ or 1) and which, since the eigenvalues are the diagonal elements, requires the complex field; and (3.10)

a direct sum of blocks, each of which is the __companion__ matrix of an invariant polynomial of the polynomial matrix $A - \lambda 1$ (cf. Gantmacher [18], v. 1, p. 149), each block having the form (illustrated for n = 5): (3.11)

$$C = \left\|\begin{matrix} 0 & 0 & 0 & 0 & -\alpha_5 \\ 1 & 0 & 0 & 0 & -\alpha_4 \\ 0 & 1 & 0 & 0 & -\alpha_3 \\ 0 & 0 & 1 & 0 & -\alpha_2 \\ 0 & 0 & 0 & 1 & -\alpha_1 \end{matrix}\right\| , \tag{3.12}$$

where the associated invariant polynomial is

$$\det(\lambda 1 - C) = \lambda^5 + \alpha_1\lambda^4 + \alpha_2\lambda^3 + \alpha_3\lambda^2 + \alpha_4\lambda + \alpha_5. \tag{3.13}$$

In case (3.9), with A triple diagonal, $A \otimes 1 + 1 \otimes \bar{A}$ take the form (for n=3) and with the standard lexicographic ordering):

$$\begin{matrix}
a_{11}+\bar{a}_{11} & \bar{a}_{12} & & a_{12} & & & & & \\
\bar{a}_{21} & a_{11}+\bar{a}_{22} & \bar{a}_{23} & & a_{12} & & & & \\
& \bar{a}_{32} & a_{11}+\bar{a}_{33} & & & a_{12} & & & \\
a_{21} & & & a_{22}+\bar{a}_{11} & \bar{a}_{12} & & a_{23} & & \\
& a_{21} & & \bar{a}_{21} & a_{22}+\bar{a}_{22} & \bar{a}_{23} & & a_{23} & \\
& & a_{21} & & \bar{a}_{32} & a_{22}+\bar{a}_{33} & & & a_{23} \\
& & & a_{32} & & & a_{33}+\bar{a}_{11} & \bar{a}_{12} & \\
& & & & a_{32} & & \bar{a}_{21} & a_{33}+\bar{a}_{22} & \bar{a}_{23} \\
& & & & & a_{32} & & \bar{a}_{32} & a_{33}+\bar{a}_{33}
\end{matrix} \tag{3.14}$$

The case (3.10) is of course obtained from (3.14) by replacing all elements below the main diagonal by zero. The resulting triangular matrix exhibits the eigenvalues of $A \otimes 1 + 1 \otimes \bar{A}$ as $a_{ii} + \bar{a}_{jj}$ on the diagonal. In case A is in Jordan form with an elementary divisor of order three with eigenvalue a and one of order two with eigenvalue b (so, n=5), we have

$$\begin{aligned}
& a_{11} = a_{22} = a_{33} = a, \quad a_{44} = a_{55} = b, \\
& a_{12} = e_1 \neq 0, \quad a_{23} = e_2 \neq 0, \quad a_{45} = e \neq 0 \quad \text{and all other} \\
& a_{ij} = 0.
\end{aligned} \tag{3.15}$$

[Here we have not chosen to make $e_1 = e_2 = e = 1$ as is customary since the more general form exhibits the relation of A to $A \otimes 1 + 1 \otimes \bar{A}$ better and may be more suitable for numerical analysis.] Exhibiting the matrix $A \otimes 1 + 1 \otimes \bar{A}$, which is now of order 25, we have:

	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)
(1,1)	$a+\bar{a}$	$\bar{e}_1$				e_1																			
(1,2)		$a+\bar{a}$	$\bar{e}_2$				e_1																		
(1,3)			$a+\bar{a}$					e_1																	
(1,4)				$a+\bar{b}$	$\bar{e}$				e_1																
(1,5)					$a+\bar{b}$					e_1															
(2,1)						$a+\bar{a}$	$\bar{e}_1$				e_2														
(2,2)							$a+\bar{a}$	$\bar{e}_2$				e_2													
(2,3)								$a+\bar{a}$					e_2												
(2,4)									$a+\bar{b}$	$\bar{e}$				e_2											
(2,5)										$a+\bar{b}$					e_2										
(3,1)											$a+\bar{a}$	$\bar{e}_1$													
(3,2)												$a+\bar{a}$	$\bar{e}_2$												
(3,3)													$a+\bar{a}$												
(3,4)														$a+\bar{b}$	$\bar{e}$										
(3,5)															$a+\bar{b}$										
(4,1)																$b+\bar{a}$	$\bar{e}_1$				e				
(4,2)																	$b+\bar{a}$	$\bar{e}_2$				e			
(4,3)																		$b+\bar{a}$					e		
(4,4)																			$b+\bar{b}$	$\bar{e}$				e	
(4,5)																				$b+\bar{b}$					e
(5,1)																					$b+\bar{a}$	$\bar{e}_1$			
(5,2)																						$b+\bar{a}$	$\bar{e}_2$		
(5,3)																							$b+\bar{a}$		
(5,4)																								$b+\bar{b}$	$\bar{e}$
(5,5)																									$b+\bar{b}$

$A \otimes 1 + 1 \otimes \bar{A}$ for A in Jordan form with two elementary divisors ($n = 5 = 3 + 2$). (3.16)

$$
\begin{array}{c|ccccccccccccccccccccccccc}
 & (1,1) & (1,2) & (1,3) & (2,1) & (2,2) & (2,3) & (3,1) & (3,2) & (3,3) & (1,4) & (1,5) & (2,4) & (2,5) & (3,4) & (3,5) & (4,1) & (4,2) & (4,3) & (5,1) & (5,2) & (5,3) & (4,4) & (4,5) & (5,4) & (5,5) \\
\hline
(1,1) & a+\bar{a} & \bar{e}_1 & & e_1 \\
(1,2) & & a+\bar{a} & \bar{e}_2 & & e_1 \\
(1,3) & & & a+\bar{a} & & & e_1 \\
(2,1) & & & & a+\bar{a} & \bar{e}_1 & & e_2 \\
(2,2) & & & & & a+\bar{a} & \bar{e}_2 & & e_2 \\
(2,3) & & & & & & a+\bar{a} & & & e_2 \\
(3,1) & & & & & & & a+\bar{a} & \bar{e}_1 \\
(3,2) & & & & & & & & a+\bar{a} & \bar{e}_2 \\
(3,3) & & & & & & & & & a+\bar{a} \\
(1,4) & & & & & & & & & & a+\bar{b} & \bar{e} & e_1 \\
(1,5) & & & & & & & & & & & a+\bar{b} & & e_1 \\
(2,4) & & & & & & & & & & & & a+\bar{b} & \bar{e} & e_2 \\
(2,5) & & & & & & & & & & & & & a+\bar{b} & & e_2 \\
(3,4) & & & & & & & & & & & & & & a+\bar{b} & \bar{e} \\
(3,5) & & & & & & & & & & & & & & & a+\bar{b} \\
(4,1) & & & & & & & & & & & & & & & & b+\bar{a} & \bar{e}_1 & & e \\
(4,2) & & & & & & & & & & & & & & & & & b+\bar{a} & \bar{e}_2 & & e \\
(4,3) & & & & & & & & & & & & & & & & & & b+\bar{a} & & & e \\
(5,1) & & & & & & & & & & & & & & & & & & & b+\bar{a} & \bar{e}_1 \\
(5,2) & & & & & & & & & & & & & & & & & & & & b+\bar{a} & \bar{e}_2 \\
(5,3) & & & & & & & & & & & & & & & & & & & & & b+\bar{a} \\
(4,4) & & & & & & & & & & & & & & & & & & & & & & b+\bar{b} & \bar{e} & e \\
(4,5) & & & & & & & & & & & & & & & & & & & & & & & b+\bar{b} & & e \\
(5,4) & & & & & & & & & & & & & & & & & & & & & & & & b+\bar{b} & \bar{e} \\
(5,5) & & & & & & & & & & & & & & & & & & & & & & & & & b+\bar{b}
\end{array}
$$

$A \otimes 1 + 1 \otimes \bar{A}$ for A in Jordan form with elementary divisors $(\lambda - a)^3$ and $(\lambda - b)^2$.

(3.17)

The lexiographic ordering used in (3.16), while it is so widely employed as to seem "natural," in fact conceals the actual structure of the matrix. Selecting out those rows and columns for which $a+\bar{a}$ is the diagonal element, we get a 9 by 9 matrix which, after $(a+\bar{a})1_9$ has been subtracted, is exhibited in (3.18) as B_9. To determine the elementary divisors with eigenvalue $a+\bar{a}$, the powers of B_9 are shown and the ranks needed may be obtained by inspection.

Block of order 9 with eigenvalue $a+\bar{a}$ from (3.17): $(a+\bar{a})1_9$ subtracted.

$$\begin{array}{c||ccccccccc||l}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & \\
\hline
1 & & \bar{e}_1 & & e_1 & & & & & & \\
2 & & & \bar{e}_2 & & e_1 & & & & & \\
3 & & & & & & e_1 & & & & \\
4 & & & & & \bar{e}_1 & & e_2 & & & \\
5 & & & & & & \bar{e}_2 & & e_2 & & = B_9 \\
6 & & & & & & & & & e_2 & \\
7 & & & & & & & & \bar{e}_1 & & \\
8 & & & & & & & & & \bar{e}_2 & \\
9 & & & & & & & & & & \\
\hline
1 & & & \bar{e}_1\bar{e}_2 & & 2\bar{e}_1e_1 & & e_1e_2 & & & \\
2 & & & & & & 2\bar{e}_2e_1 & & e_1e_2 & & \\
3 & & & & & & & & & e_1e_2 & \\
4 & & & & & & \bar{e}_1\bar{e}_2 & & 2\bar{e}_1e_2 & & \\
5 & & & & & & & & & 2\bar{e}_2e_2 & = (B_9)^2 \\
6 & & & & & & & & & & \\
7 & & & & & & & & & \bar{e}_1\bar{e}_2 & \\
8 & & & & & & & & & & \\
9 & & & & & & & & & & \\
\hline
1 & & & & & & 3\bar{e}_2|e_1|^2 & & 3e_2|e_1|^2 & & \\
2 & & & & & & & & & 3e_1|e_2|^2 & \\
3 & & & & & & & & & & \\
4 & & & & & & & & & 3\bar{e}_1|e_2|^2 & \\
5 & & & & & & & & & & = (B_9)^3 \\
6 & & & & & & & & & & \\
7 & & & & & & & & & & \\
8 & & & & & & & & & & \\
9 & & & & & & & & & & \\
\end{array} \tag{3.18}$$

Rank $B_9=6$, rank $(B_9)^2=4$, rank $(B_9)^3=2$, rank $(B_9)^4=1$, rank $(B_9)^5=0$.

Hence, the elementary divisors with eigenvalue $a+\bar{a}$ in (3.16) are of orders 5, 3 and 1. For, if there are n_i elementary divisors of order i, $n_1+n_2+\dots+pn_p + pn_{p+1}+\dots = n$-rank Bp and $n_1 = 1$, $n_2 = 0$, $n_3 = 1$, $n_4 = 0$, $n_5 = 1$, $n_6 = n_7 = \dots = 0$ is the solution of these equations.

4. Elementary Divisors of a Matrix of the Form $A \otimes 1 + 1 \otimes B$.

We wish to determine the eigenvalues and associated elementary divisors of the matrix $A \otimes 1 + 1 \otimes \bar{A}$ but find it equally convenient to consider the more general case of two square matrices A, m by m, and B, n by n, and determine the similarity invariants of $A \otimes 1_n + 1_m \otimes B$, which is a square matrix of order mn.

If $\mathcal{S}$ is a linear subspace invariant under A, so that $x \in \mathcal{S}$ implies $(Ax) \in \mathcal{S}_1$, and $\mathcal{S}_2$ is invariant under B, $(y \in \mathcal{S}_2 \Rightarrow By \in \mathcal{S}_2)$, we have immediately that

$$(A \otimes 1 + 1 \otimes B)\ (x \otimes y) = (Ax) \otimes y + x \otimes By \in \mathcal{S}_1 \otimes \mathcal{S}_2, \tag{4.1}$$

where by $\mathcal{S}_1 \otimes \mathcal{S}_2$ we mean the linear subspace of $\mathcal{V}^{(1)}_m \otimes \mathcal{V}^{(2)}_n = \mathcal{V}^{(3)}_{mn}$ spanned by the vectors of $\mathcal{V}^{(3)}_{mn}$ of the form $x_1 \otimes x_2$ where $x_1 \in \mathcal{S}_1$ and $x_2 \in \mathcal{S}_2$. Hence, if $\mathcal{V}^{(1)}_m$ is written as a direct sum

$$\mathcal{V}^{(1)} = \mathcal{E}_1 \oplus \mathcal{E}_2 \oplus \cdots \oplus \mathcal{E}_p \tag{4.2}$$

of subspaces each invariant under A and, similarly,

$$\mathcal{V}^{(2)} = \mathcal{F}_1 \oplus \mathcal{F}_2 \oplus \cdots \oplus \mathcal{F}_q \quad , \tag{4.3}$$

where each $\mathcal{F}_j$ is invariant under B, then

$$\mathcal{V}^{(3)} = \sum_{\substack{i=1,\ldots,p \\ j=1,\ldots,q}} \oplus\ (\mathcal{E}_i \otimes \mathcal{F}_j) \equiv \sum_{k=1}^{pq} \oplus\ \mathcal{G}_k \tag{4.4}$$

is a direct sum decomposition of $\mathcal{V}^{(3)}$ into subspaces invariant under

$$C = A \otimes 1 + 1 \otimes B. \tag{4.5}$$

To obtain the elementary divisors of C it therefore suffices to take A and B in Jordan canonical (block) form and to consider only the case in which A and B each have a single elementary divisor.

We shall prove

Theorem 4.1. If A is m by m and has a single elementary divisor with eigenvalue a and B is n by n with a single divisor of eigenvalue b, then $A \otimes 1 + 1 \otimes B$ has the eigenvalue a + b and elementary divisors of orders

$$m + n - 1, \quad m + n - 3, \quad m + n - 5, \ldots, \quad |m-n| + 1 \quad .$$

When this has been proved, the above remarks establish the general result:

Theorem 4.2. If A has eigenvalues a_i with associated elementary divisors $(\lambda - a_i)^{p_i}$ and B has elementary divisors $(\lambda - b_j)^{q_j}$, where neither the a_i nor the b_j need be distinct, then $A \otimes 1 + 1 \otimes B$ has elementary divisors $(\lambda - a_i - b_j)^{r_{ij}}$, where

$$r_{ij} = p_i + q_j - 1,\ p_i + q_j - 3, \ldots, |p_i - q_j| + 1 \quad . \tag{4.6}$$

To prove Thm. 4.1, we choose a basis with respect to which both A and B are in Jordan canonical form. This is possible from the basic construction of a tensor product of vector spaces and also follows from the formula

$$(T \otimes U)(A \otimes 1 + 1 \otimes B)(T \otimes U)^{-1} = (TAT^{-1}) \otimes 1 + 1 \otimes (UBU^{-1}) \tag{4.7}$$

Thus we choose bases so that

$$Ae_i = ae_i + e_{i-1}, \text{ with } e_0 = 0 \text{ and } i = 1, \ldots, m, \tag{4.8}$$

and

$$Bf_j = bf_j + f_{j-1}, \text{ with } f_0 = 0 \text{ and } j = 1, \ldots, n. \tag{4.9}$$

A basis for the tensor product space consists of the $e_i \otimes f_j$, which we denote simply by (i,j). Then,

$$\begin{aligned} C(i,j) &\equiv (A \otimes 1 + 1 \otimes B)(e_i \otimes f_j) \\ &= (Ae_i) \otimes f_j + e_i \otimes (Bf_j) \\ &= (ae_i + e_{i-1}) \otimes f_j + e_i \otimes (bf_j + f_{j-1}) \\ &= (a+b)(e_i \otimes f_j) + e_{i-1} \otimes f_j + e_i \otimes f_{j-1} \\ &\equiv (a+b)(i,j) + (i-1,j) + (i,j-1). \end{aligned} \tag{4.10}$$

Writing

$$N = C - (a+b) 1_{mn}, \tag{4.11}$$

the matrix N is nilpotent and the ranks of its powers will determine the degrees of the elementary divisors of C according to the formulas:

if there are p_i elementary divisors of degree (= size of block) i, then

$$\begin{aligned} p_i + p_2 + p_3 + p_4 + \ldots &= mn - \text{rank } N \\ p_1 + 2p_2 + 2p_3 + 2p_4 + \ldots &= mn - \text{rank } N^2 \\ p_1 + 2p_2 + 3p_3 + 3p_4 + \ldots &= mn - \text{rank } N^3 \\ &\vdots \\ p_1 + 2p_2 + 3p_3 + 4p_4 + \ldots + kp_k + kp_{k+1} + \ldots &= mn - \text{rank } N^k. \end{aligned} \tag{4.12}$$

These formulas are "well known" and are trivially proved by the observation that if a Jordan block is written in the form

$$\begin{Vmatrix} a & 1 & & & \\ & a & 1 & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ & & & & a \end{Vmatrix} = a\,1 + U\ , \tag{4.13}$$

then the nullity of U^k = order of U - rank of U^k = k until k = order of U, after which $U^k = 0$ so the nullity can no longer increase, establishing the general formula

$$\sum_{i=1}^{mn} (\min (i,k))\ p_i = mn - \text{rank } N^k \ . \tag{4.14}$$

Note that $k = 0$ is admissible since $N^0 = 1$ has rank mn.

Differencing the above formulas gives

$$\begin{aligned} p_1 + p_2 + p_3 + p_4 + \cdots &= \text{rank } N^0 - \text{rank } N \\ p_2 + p_3 + p_4 + \cdots &= \text{rank } N - \text{rank } N^2 \\ p_3 + p_4 + \cdots &= \text{rank } N^2 - \text{rank } N^3 \\ p_4 + \cdots &= \text{rank } N^3 - \text{rank } N^4, \text{ etc.} \end{aligned} \tag{4.15}$$

Hence

$$\begin{aligned} p_1 &= (\text{rank } N^0 - \text{rank } N) - (\text{rank } N - \text{rank } N^2)\ , \\ p_2 &= (\text{rank } N - \text{rank } N^2) - (\text{rank } N^2 - \text{rank } N^3)\ , \\ &\ \ \vdots \\ p_r &= (\text{rank } N^{r-1} - \text{rank } N^r) - (\text{rank } N^r - \text{rank } N^{r+1})\ . \end{aligned} \tag{4.16}$$

To calculate the rank of N^r neither of the lexicographic orders of the last section is convenient. Instead we write (4.10) in the form

$$N(i,j) = (i-1,j) + (i,j-1) \tag{4.17}$$

and, defining,

$$\text{grade } (i,j) = i + j-1\ , \tag{4.18}$$

observe that N transforms a basis vector of grade $i + j-1$ into the sum of two vectors each having grade $i + j - 2$. The structure of the matrix N will then be exhibited more clearly if we order the basis vector 1) by grade, and 2) lexicographically within each grade. Thus if m=5, n=3, N is 15 by 15 and with the prescribed ordering has the appearance shown in (4.19).

Grades		1	2		3			4			5			6		7
		(1,1)	(1,2)	(2,1)	(1,3)	(2,2)	(3,1)	(2,3)	(3,2)	(4,1)	(3,3)	(4,2)	(5,1)	(4,3)	(5,2)	(5,3)
1	(1,1)	0	1	1												
2	(1,2)				1	1	0									
	(2.1)				0	1	1									
	(1.3)							1								
3	(2,2)							1	1							
	(3,1)								1	1						
	(2,3)										1					
4	(3,2)										1	1				
	(4,1)											1	1			
	(3,3)													1		
5	(4,2)													1	1	
	(5,1)														1	
6	(4,3)															1
	(5,2)															1
7	(5,3)															0

$$A \otimes 1_3 + 1_5 \otimes B - (a + b)\, 1_{15} = N \text{ where } (A - a1)^5 = 0,\ (B - b1)^3 = 0. \tag{4.19}$$

For convenience of notation we suppose $m \geqslant n$ and set $m = n+p-1$ for $p \geqslant 1$. [If $n > m$, ordering the (i,j) on j followed by i would exhibit the same structure as would interchanging m and n.]

Listing the (i,j) by grades, we find the table:

	Grade (g)	Basis Vectors	Number (n_g) of Basis Vectors
	1	$(1,1)$	1
	2	$(1,2), (2,1)$	2
	⋮	⋮	⋮
	$g \leqslant n$	$(1,g), (2,g-1), \ldots, (g,1)$	g
	⋮		⋮
	$n-1$	$(1,n-1), (2,n-2), \ldots, (n-1,1)$	$n-1$
	n	$(1,n), (2,n-1), \ldots, (n,1)$	n
if $m>n$,	$n+1$	$(2,n), (3,n-1), \ldots, (n,2), (n+1,1)$	n
if $m>n+1$,	$n+2$	$(3,n), (4,n-1), \ldots, (n+1,2), (n+2,1)$	n
	⋮	⋮	⋮
$m=n+p-1$	$n+p-1$	$(p,n), (p+1,n-1), \ldots, (n+p-2,2), (n+p-1,1)$	n
	$n+p(=m+1)$	$(p+1,n), (p+2,m-1), \ldots, (n+p-1,2)$	$n-1$
	⋮	⋮	⋮
	$m+n-g'$	$(m+1-g',n), (m+2-g',n-1), \ldots, (m,n+1-g')$	g'
	⋮	⋮	⋮
	$m+n-2$	$(m-1,n), (m,n-1)$	2
	$m+n-1$	(m,n)	1

(4.20)

Note that there are p sets of n basis vectors for grades n to $n+p-1$.

The general formula for the number n_g of basis vectors of degree g is then

$$n_g = \min\,(g,\ m,\ n,\ m+n-g). \tag{4.21}$$

Hence the matrix N will appear in block form (as displayed in (4.19)) if its rows and columns are partitioned into subsets $\mathcal{N}_1, \mathcal{N}_2, \ldots, \mathcal{N}_{m+n-1}$ conforming with the partition

$$n_1 + n_2 + \ldots + n_{m+n-1} = mn \tag{4.22}$$

of the order of N.

This partitioning is symmetric in the sense that

$$n_g = n_{m+n-g} \tag{4.23}$$

as follows from replacing g by m+n-g in (4.21).

To show that additional details of the structure observed in (4.19) are in fact general, we write N^r in block form

$$N^r = \left(N_{gh}^{(r)}\right), \tag{4.24}$$

where

$$N_{gh}^{(r)} \text{ has } n_g \text{ rows and } n_h \text{ columns} . \tag{4.25}$$

The remark following (4.18) shows that N always has the form shown in (4.26); that is, $N_{ij}^{(1)} = 0$ unless $j = i+1$:

$$\left\| \begin{matrix} 0 & N_{12}^{(1)} & & & & & \\ & 0 & N_{22}^{(1)} & & & & \\ & & 0 & N_{34}^{(1)} & & & \\ & & & 0 & \cdot & \cdot & \\ & & & & \cdot & \cdot & \\ & & & & & \cdot & \cdot \\ & & & & & & N_{q-1,q}^{(1)} \\ & & & & & & 0 \end{matrix} \right\| = N \tag{4.26}$$

where q = m+n-1 and blocks not shown are zero.

By inspection, or by observing that if $i \in \eta_g$ and $j \in \eta_h$ a non-zero term in the expression

$$(N^r)_{ij} = \sum_{k_1,\ldots,k_{r-1}} n_{ik_1} n_{k_1k_2} n_{k_2k_3} \ldots n_{k_{r-1}j} \tag{4.27}$$

must have $k_1 \in \eta_{g+1}$, $k_2 \in \eta_{g+2}, \ldots, k_{r-1} \in \eta_{g+r-1}$, and $j \in \eta_{g+r}$, where η_g is the index set for basis vectors of grade g, we have

$$N_{gh}^{(r)} = 0 \text{ unless } h = g+r \tag{4.28}$$

and then

$$N_{g,g+r}^{(r)} = N_{g,g+1}^{(1)} N_{g+1,g+2}^{(1)} \ldots N_{g+r-1,g+r}^{(1)} \tag{4.29}$$

5. Ranks of the Transformations of the Quotient Spaces

In this section we shall determine explicitly the ranks of each of these (block) matrices. To do so we first observe that each of the non-zero blocks in $N^{(1)}$ has maximum possible rank:

$$\text{rank}\ N^{(1)}_{g,g+1} = \min\ (n_g, n_{g+1}). \tag{5.1}$$

This is trivially established by exhibiting the three possible cases:

$$n_g = n_{g+1} - 1, \quad n_g = n_{g+1} \quad \text{and} \quad n_g = n_{g+1} + 1\ , \tag{5.2}$$

which occur, respectively, when

$$g < n\ , \quad n \leqslant g \leqslant n+p-2 = m-1, \quad \text{and } m \leqslant g. \tag{5.3}$$

Note that for m = n no "square" block matrices occur and in general there is one matrix with i rows and (i+1) columns (for i=1,...,n-1), p-1 matrices with n rows and n columns, and one matrix with (j+1) rows and j columns (for j=n-1,...,1). The matrices of these three types are:

$$\begin{array}{c c}
 & \begin{array}{cccccc} (1,g+1) & (2,g) & (3,g-1) & \cdots & (g,2) & (g+1,1) \end{array} \\
\begin{array}{c} (1,g) \\ (2,g-1) \\ (3,g-2) \\ \vdots \\ (g-1,2) \\ (g,1) \end{array} &
\left\| \begin{array}{cccccc} 1 & 1 & & & & \\ & 1 & 1 & & & \\ & & 1 & \cdots & & \\ & & \vdots & & & \\ & & & & 1 & \\ & & & & 1 & 1 \end{array} \right\| \equiv N^{(1)}_{g,g+1}\ ;
\end{array}$$

$$(g \text{ by } g+1) \tag{5.4}$$

and,

$$\begin{array}{c c}
 & \begin{array}{cccccc} (k+1,n) & (k+2,n-1) & (k+3,n-2) & \cdots & (n+k-1,2) & (n+k,1) \end{array} \\
\begin{array}{c} (k,n) \\ (k+1,n-1) \\ (k+2,n-2) \\ \vdots \\ (n+k-2,2) \\ (n+k-1,1) \end{array} &
\left\| \begin{array}{cccccc} 1 & & & & & \\ 1 & 1 & & & & \\ & 1 & 1 & \cdots & & \\ & \vdots & \vdots & & & \\ & & & & 1 & \\ & & & & 1 & 1 \end{array} \right\| \equiv 1_n + V_n\ .
\end{array}$$

$$(n \text{ by } n) \tag{5.5}$$

There is no need to display the third case since

$$N^{(1)}_{m+n-g-1,m+n-g} = \left(N^{(1)}_{g,g+1} \right)^T \quad \text{for } g \leqslant n-1\ . \tag{5.6}$$

It is to be noted that this "symmetry" does not extend to the n by n blocks, which are all of the form $1 + V$ with V a nilpotent matrix with ones on the subdiagonal and zeros elsewhere. That rank $N^{(1)}_{g,g+1} = g$ for $g \leqslant n-1$ is obvious by inspection since the g rows (or, the first or last g columns) are linearly independent.

We now proceed to extend (5.1) to the general rule

$$\text{rank } N^{(r)}_{g,g+r} = \min (n_g, n_{g+r}) \quad . \tag{5.7}$$

Where the matrix is square ($n_g = n_{g+r}$), the case $n_g = n_{g+r} = n$ is settled by observing that

$$N^{(r)}_{g,g+r} = (1 + V)^r \tag{5.8}$$

is nonsingular. For $n_g = n_{g+r} < n$, (4.20) gives

$$g = n_g = n_{g+r} \quad \text{and} \quad g + r = m+n-g \quad . \tag{5.9}$$

Hence,

$$r = m+n-2g, \text{ with } 0 < g < n \leqslant m. \tag{5.10}$$

We can obtain an explicit form for the elements of an arbitrary one of the blocks by expressing N as the sum of commuting nilpotents,

$$N = N_1+N_2 \quad , \quad N_1N_2 = N_2N_1 \quad , \tag{5.11}$$

where $N_1(i,j) = (i-1,j)$ and $N_2(i,j) = (i,j-1)$ and using the binomial expansion to get

$$N^r = \sum_{s=0}^{r} \binom{r}{s} N_1^s \, N_2^{r-s} \quad . \tag{5.12}$$

Hence, if N^r is applied to (i,j), it yields

$$N^r(i,j) = \sum_{s=0}^{r} \binom{r}{s} (i-s, \ j-r+s) \quad , \tag{5.13}$$

where, as usual $(p,q) = 0$ for $p \lesseqgtr 0$ or $q \lesseqgtr 0$. Thus the element in column (i,j) and row (p,q) of N^r is, taking $i-s = p$ and $j-r+s = q$,

$$n^{(r)}_{(p,q),(i,j)} = \binom{r}{i-p} \quad , \text{ if } r = (i+j) - (p+q) \tag{5.14}$$

and is zero otherwise.

For the special case of a square block of size less than n, we have

$$n^{m+n-2g}_{(p,g+1-p),(m-g+i,n+1-i)} = \binom{m+n-2g}{(m-g)+(i-p)} \quad , \tag{5.15}$$

which is the element in row p and column i for $p,i = 1,2,\ldots,g$ as "single index" enumeration of the rows and columns.

The determinant of a matrix of this form was calculated in 1865 by von Zeipel [70] and his result is given by Muir ([43], vol. III, p. 449) and by Muir and Metzler ([44], p. 682). For completeness, we give an evaluation of this determinant which does not include the extraneous detail needed by these authors for their more general cases.

Thus, let D(a,b,g) be the determinant of the g by g matrix $M^{(0)} = \| M_{ij}^{(0)} \| \equiv M^{(0)}(a,b,g)$, with

$$m_{ij}^{(0)} = \binom{a}{b+j-i} \tag{5.16}$$

as the element in its i-th row and j-th column. Then we seek the value of D(m+n-2g, m-g,g) (and want to know that it is different from zero). Using the combinatorial relation

$$\binom{K-1}{L} + \binom{K-1}{L-1} = \binom{K}{L} \tag{5.17}$$

repeatedly, we get, on adding sequentially the i-th row to the (i+1)st row for i = 1,2,...,g, a new matrix $M^{(1)}$ with elements

$$m_{ij}^{(1)} = \begin{cases} \dbinom{a}{b+j-i} & \text{for } i = 1, \\[2ex] \dbinom{a+1}{b+1+j-i} & \text{for } i > 1 \quad . \end{cases} \tag{5.18}$$

Adding the i-th row to the (i+1)st for i = 2,3,...,g (thus leaving the first and second row unchanged), we get the special case k=2 of

$$m_{ij}^{(k)} = \begin{cases} \dbinom{a+i-1}{b+i-1+j-i} & \text{for } i \leqq k \ , \\[2ex] \dbinom{a+k}{b+k+j-i} & \text{for } i > k \ . \end{cases} \tag{5.19}$$

The induction is easily established and D(a,b,g) is also the value of the determinant of $M^{(n)}$, where

$$m_{ij}^{(n)} = \binom{a+i-1}{b+j-1} \tag{5.20}$$

Factoring a+i-1 out of the i-th row for i=1,...,g and b+j-1 out of the j-th column for j=1,...,g has the effect of replacing a by a-1 and b by b-1 in the formula for the (i,j)-th element. Repeating the process until a has been reduced to a-b and b to zero,

$$D(a,b,g) = \frac{(a)(a-1)\ldots(a-b+1)}{(b)(b-1)\ldots(1)} \cdot \frac{(a+1)\ldots(a-b+2)}{(b+1)\ldots(2)} \cdots \frac{(a+g-1)\ldots(a-b+g)}{(b+g-1)\ldots(g)} \cdot D$$

$$= \left[\prod_{i=1}^{g} \binom{a+i-1}{b}\right] \left[\prod_{j=1}^{g} \binom{b+j-1}{b}\right]^{-1} \cdot D \quad , \tag{5.21}$$

where

$$D = \det \left\| \begin{matrix} \binom{a-b}{0} & \binom{a-b}{1} & \cdots & \binom{a-b}{g-1} \\ 0 & \binom{a-b}{0} & \cdots & \binom{a-b}{g-2} \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & \binom{a-b}{0} \end{matrix} \right\| \tag{5.22}$$

has the value one since it is triangular with diagonal elements all one. We therefore have

$$D(m+n-2g,\ m-g,g) = \left[\prod_{i=1}^{g} \binom{m+n-2g+i-1}{m-g}\right] \left[\prod_{j=1}^{g} \binom{m-g+j-1}{m-g}\right]^{-1} \tag{5.23}$$

is different from zero.

Returning to complete the proof of (5.7), we observe that when $N^{(r)}_{g,g+r}$ is not square, it is always of the form

$$A_0A_1\ldots A_kS \text{ if } n_g < n_{g+r} \tag{5.24}$$

or of the form

$$S\,A_k^T\,A_{k-1}^T\ldots A_0^T\ ,\ \text{if } n_g > n_{g+r} \quad , \tag{5.25}$$

where S = 1 or S is one of the square nonsingular matrices studied above. Hence it suffices to show that

$$\operatorname{rank}\,(A_0A_1\ldots A_k) = n_g \quad , \tag{5.26}$$

where A_i has n_g+i rows and n_g+i+1 columns and its rows are linearly independent by (5.1). The argument is completed by observing that for a row vector $u_0 \neq 0$, the linear independence of the rows of A_0 implies that $u_0A_0 \neq 0$, and hence, inductively,

$$u_i A_i = u_0 A_0 A_1 \ldots A_i \neq 0 \text{ for } i = 0,1,\ldots,k \tag{5.27}$$

and the rows of $A_0 A_1 \ldots A_k$ are linearly independent.

Remembering the fact that N^r is, with respect to the block subdivision, zero except for blocks along its r-th diagonal, we have immediately

$$\text{rank } N^r = \text{rank } N^{(r)}_{1,1+r} + \text{rank } N^{(r)}_{2,2+r} + \ldots + \text{rank } N^{(r)}_{q-r,q} \quad . \tag{5.28}$$

Hence,

$$\text{rank } N^r = \min(n_1, n_{1+r}) + \min(n_2, n_{2+r}) + \ldots + \min(n_{q-r}, n_q) \quad , \tag{5.29}$$

where there are $q = m+n-1$ grades (so, $n_1+n_2+\ldots+n_q = mn$).

To clarify the structure of these formulas, consider the example of (4.19), in which $m=5$, $n=3$, and

$$mn = 1 + 2 + 3 + 3 + 3 + 2 + 1 \tag{5.30}$$

is the partition of mn. Then

$$n_1 = 1,\ n_2 = 2,\ n_3 = 3,\ n_4 = 3,\ n_5 = 3,\ n_6 = 2,\ n_7 = 1 \text{ and } q = 7. \tag{5.31}$$

Hence,

$$\begin{aligned}
\text{rank } N &= \min(1,2) + \min(2,3) + \min(3,3) + \min(3,3) + \min(3,2) + \min(2,1) \\
&= 1 + 2 + 3 + 3 + 2 + 1 \\
&= 12, \\
\text{rank } N^2 &= \min(1,3) + \min(2,3) + \min(3,3) + \min(3,2) + \min(3,1) \\
&= 1 + 2 + 3 + 2 + 1 \\
&= 9\ , \\
\text{rank } N^3 &= \min(1,3) + \min(2,3) + \min(3,2) + \min(3,1) \\
&= 1 + 2 + 2 + 1 \\
&= 6\ , \\
\text{rank } N^4 &= \min(1,3) + \min(2,2) + \min(3,1) \\
&= 1 + 2 + 1 \\
&= 4\ , \\
\text{rank } N^5 &= \min(1,2) + \min(2,1) \\
&= 1 + 1 \\
&= 2\ , \\
\text{rank } N^6 &= \min(n_1, n_7) = \min(1,1) = 1.
\end{aligned} \tag{5.32}$$

Finally,

$$N^{m+n-1} = 0 \tag{5.33}$$

is true as a general result since each application of N reduces the grade of a basis vector by one and the maximum grade is $m+n-1$.

With the usual convention that $N^o = 1$, therefore of rank mn, (5.29) will be true for $r = 0,1,2,\dots,m+n-1$, provided we put $n_{m+n} = 0$. From (4.16) we obtain the degrees of the elementary divisors by calculating the second differences of the ranks of the N^r. In the example,

$$\text{rank } N^r\text{:}\quad 15,12,9,6,4,2,1,0,0 \qquad (\text{for } r=1,\dots,8), \tag{5.34}$$

$$\Delta_r \equiv \text{rank } N^{r-1} - \text{rank } N^r\text{:}\quad 3,3,3,2,2,1,1,0 \qquad (\text{for } r=1,\dots,8), \tag{5.35}$$

$$p_r = \Delta_r - \Delta_{r+1}\text{:}\quad 0,0,1,0,1,0,1 \qquad (\text{for } r=1,\dots,7), \tag{5.36}$$

so that

$$p_1=0,\ p_2=0,\ p_3=1,\ p_4=0,\ p_5=1,\ p_6=0,\ p_7=1 \tag{5.37}$$

and there are elementary divisors of orders

$$m+n-1 = 5+3-1=7, \quad m+n-3 = 5 \quad \text{and} \quad m-n+1 = 5-3+1=3. \tag{5.38}$$

Carrying through the formalism in the general case by using (5.29) and (4.21), we have

$$\begin{aligned} \text{rank } N^r &= \sum_{i=1}^{i=q-r} \min\Big(\min(i,m,n,m+n-i),\ \min(i+r,m,n,m+n-i-r)\Big) \\ &= \sum_{i=1}^{i=q-r} \min(i,m,n,m+n-i-r) \quad , \end{aligned} \tag{5.39}$$

for $r=0,1,\dots,q-1$ and rank $N^q=0$. Differencing, and observing that for $i=q-r$ and $r=t-1$

$$\begin{aligned} \min(i,m,n,m+n-i-r) &= \min(q-t+1,m,n,m+n-q) \\ &= \min(m+n-t,m,n,1) \\ &= 1 \text{ for } t=1,2,\dots,m+n-1=q-1 \quad , \end{aligned} \tag{5.40}$$

we have

$$
\begin{aligned}
\triangle_r &= \text{rank } N^{r-1} - \text{rank } N^r \\
&= [\min(1,m,n,m+n-1-r+1) - \min(1,m,n,m+n-1-r)] \\
&\quad + [\min(2,m,n,m+n-2-r+1) - \min(2,m,n,m+n-2-r)] \\
&\quad + [\min(3,m,n,m+n-3-r+1) - \min(3,m,n,m+n-3-r)] \\
&\quad + \cdots \\
&\quad + [(s_r-1) - (s_r-1) \text{ or } m-m \text{ or } n-n] \\
&\quad + [(m+n-s_r-r+1) - (m+n-s_r-r)] \\
&\quad + \cdots \\
&\quad + [(m+n-q+r-r+1) - (m+n-q+r-r)] \\
&\quad + 1 \quad ,
\end{aligned}
\tag{5.41}
$$

or

$$
\begin{aligned}
\triangle_r &= \underbrace{0+0+\cdots+0+}_{s_r-1}\ \underbrace{1+1+\cdots+1}_{q-r-(s_r-1)}\ +1 \\
&= q-r-s_r+2 \quad ,
\end{aligned}
\tag{5.42}
$$

and s_r is defined to be the least integer for which

$$m+n-s_r-r < \min(s_r,m,n) \tag{5.43}$$

and $r=1,2,\ldots,q+1$. (Note that $\triangle_q$ = rank N^{q-1} - rank $N^q = 1-0 = 1$ is verified since $m+n-q=1$ so $1-s_q<\min(s_q,m,n)$ gives $s_q=1$ and $\triangle_q=q-q-s_q+2=1$, correctly. Also, $s_{q+1}=1$ yields $\triangle_{q+1}=0$, again correctly. To obtain s_r more conveniently we can take advantage of the symmetry in m and n and set $p = |m-n| + 1$, $k = \min(m,n)$ to get s_r as the integer defined by

$$
\begin{aligned}
s_r &= \min\left\{s;\ 2k-1+p-r < s+\min(s,k)\right\} \\
&= \min\left\{s;\ 2k-1+p-r < \min(2s,s+k)\right\} \quad .
\end{aligned}
\tag{5.44}
$$

Clearly,

$$
\text{if } r \leq p-1,\ s_r \geqq k \text{ and } s_r = \min\left\{s;\ k+p-1-r < s\right\} = k+p-r; \quad \text{and}
\tag{5.45}
$$

$$
\begin{aligned}
&\text{if } r = p+2t \text{ or } p+2t+1 \text{ for } t=0,1,\ldots,\max(m,n)-1, \\
&\quad s_r = \min\left\{s;\ (2k-1-2t \text{ or } 2k-1-2t-1) < 2s\right\} \\
&\quad\quad = k-t \ .
\end{aligned}
\tag{5.46}
$$

The sequence of values of s_r is then

$$k+p-1,\ k+p-2,\ldots,k+1,\ \underbrace{k,k},\ \underbrace{k-1,k-1},\ldots,\underbrace{2,2},\underbrace{1,1}. \tag{5.47}$$

and there are $p-1+2k = |m-n|+1-1+2\min(m,n) = m+n=q+1$ terms in the sequence.

Differencing the Δ_r from (5.42) gives, by (4.16),

$$p_r = \Delta_r - \Delta_{r+1} = (q-r-s_r+2) - (q-r-1-s_{r+1}+2) \tag{5.48}$$
$$= s_{r+1} - s_r+1$$

and the sequence of values of the p_r for $r=1,2,\ldots,q$, is

$$\underbrace{0,0,\ldots,0}_{p-1}\ ,\ \underbrace{1,0}_{2}\ ,\ \underbrace{1,0}_{2}\ ,\ldots,\ \underbrace{1,0}_{2}\ ,\ \underbrace{1}_{1}\quad , \tag{5.48}$$

and, since $N^{q+2} = N^{q+3} = \ldots = 0$, additional zeros may be used to extend the sequence. (There are no elementary divisors of degree greater than q.) The result of Theorem 4.1 now is established:

N, and hence $C = (a+b)\ 1_{mn}+N$, has a single elementary divisor of each of the orders

$$p,\ p+2,\ p+4,\ldots,q \tag{5.50}$$
$$= |m-n|+1,\ |m-n|+3,\ |m-n|+5,\ldots,m+n-1.$$

6. Restriction to the Hermitian Subspace

So long as we deal with the general operator $A \otimes 1 + 1 \otimes B$, the invariant subspaces are dependent on the structure of both A and B and have been determined in the last sections.

In the particular case in which we are primarily interested, a real linear subspace* of the n^2 dimensional vector space is invariant for all matrices A since

$$M = M^* \text{ implies } (AM + MA^*)^* = AM + MA^* \qquad (6.1)$$

and the hermitian matrices therefore constitute an invariant real vector subspace (since $r_1M_1 + r_2M_2$ is Hermitian for all real numbers r_1 and r_2 if M_1 and M_2 are).

Generally, we have not needed to distinguish the real and complex cases but here there is a quite fundamental distinction since when the number field is complex the effect of the operator $A \otimes 1 + 1 \otimes \bar{A}$ on the hermitian matrices alone uniquely determines the operator since, over the real numbers, there are n^2 linearly independent hermitian matrices while a basis for the space of real symmetric matrices contains $\frac{1}{2}n(n+1)$ matrices and a full description of the real matrix $A \otimes 1 + 1 \otimes A$ requires a study of its effect on skewsymmetric matrices also. Another way of stating the distinction is to observe that the mapping $M \to \sqrt{-1}\, M$ establishes an (additive) isomorphism of hermitian matrices onto the skew (or, anti-) hermitian matrices (for which $M^* = -M$) and there is no similar isomorphism of real symmetric and skewsymmetric matrices. In section 9 we study the real case but here consider the complex one.

Given any n linearly independent vectors v_j, we can form the n^2 hermitian matrices:

$$H_{jj} = 2\, v_j v_j^*, \quad j = 1,2,\ldots,n \qquad (6.2)$$

$$H_{jk} = v_j v_k^* + v_k v_j^*, \quad j < k \qquad (6.3)$$

and

$$K_{jk} = i\,(v_j v_k^* - v_k v_j^*), \quad j < k. \qquad (6.4)$$

That they are linearly independent follows immediately from their form when the v_j are expressed relative to themselves as a basis, for then

H_{jj} = a matrix with 2 in the (j,j) position and zeros elsewhere, (6.5)

H_{jk} = a matrix with 1 in the (j,k) and (k,j) position and zeros elsewhere, (6.6)

and

K_{jk} = a matrix with i in the (j,k) position, -i in the (k,j) position and zeros elsewhere. (6.7)

---

*In classical projective geometry, such subspaces were called "chains."

Just as in section 4, we can consider a decomposition of the space $\mathcal{V}$ in which A acts into a direct sum of subspaces each invariant under A. Since $B = \bar{A}$, an independent decomposition of the vector space on which B acts is no longer possible and if

$$\mathcal{V} = \mathcal{E}_\alpha \oplus \mathcal{E}_\beta \oplus \ldots, \tag{6.8}$$

we can indicate the fixed mapping of $\mathcal{V}$ which carries each vector to its complex conjugate by denoting the space $\mathcal{W}$ in which $\bar{A}$ acts by $\bar{\mathcal{V}}$ and splitting it in corresponding fashion into

$$\bar{\mathcal{V}} = \bar{\mathcal{E}}_\alpha \oplus \bar{\mathcal{E}}_\beta \oplus \ldots .$$

We then need to consider

$$\mathcal{V} \otimes \bar{\mathcal{V}} = \sum \oplus (\mathcal{E}_\alpha \otimes \bar{\mathcal{E}}_\alpha) \oplus \sum_{\alpha \neq \beta} \oplus \left[\left(\mathcal{E}_\alpha \otimes \bar{\mathcal{E}}_\beta \oplus \mathcal{E}_\beta \otimes \bar{\mathcal{E}}_\alpha\right)\right] \tag{6.9}$$

since the spaces $\mathcal{E}_\alpha \otimes \bar{\mathcal{E}}_\beta$, while invariant, will not in general contain any hermitian matrices whereas $(\mathcal{E}_\alpha \otimes \bar{\mathcal{E}}_\beta) \oplus (\mathcal{E}_\beta \otimes \bar{\mathcal{E}}_\alpha)$ is spanned by its real subspace of hermitian matrices. [Strictly, the notation is inadequate, since we might have two elementary divisors with the same eigenvalue: $\alpha = \beta$ yet $\mathcal{V}_\alpha \neq \mathcal{V}_\beta$.]

Choosing a Jordan basis for one of the subspaces $\mathcal{E}_\alpha$ associated with an elementary divisor $(\lambda - \alpha)^m$ we have

$$Av_j = \alpha v_j + v_{j-1} \quad , \quad v_0 = 0 \text{ and } j=1,2,\ldots,m. \tag{6.10}$$

Then defining the H_{jj} by (6.2),

$$\begin{aligned} AH_{jj} + H_{jj}A^* &= 2(Av_j)\, v_j^* + 2v_j(v_j^*A^*) \\ &= 2(\alpha v_j + v_{j-1})v_j^* + 2v_j\,(\alpha v_j + v_{j-1})^* \\ &= 2(\alpha + \bar{\alpha})v_j\, v_j^* + 2(v_{j-1}\, v_j^* + v_j v_{j-1}^*) \\ &= 2a_1\, H_{jj} + 2H_{j-1,j} \; , \end{aligned} \tag{6.11}$$

where

$$\alpha = a_1 + ia_2 \tag{6.12}$$

and the $H_{j-1,j}$ are defined by (6.3) with $H_{01} = 0$.

Similarly,

$$AH_{jk} + H_{jk}A^* = 2a_1H_{jk} + H_{j-1,k} + H_{j,k-1} \; , \tag{6.13}$$

and

$$AK_{jk} + K_{jk}A^* = 2a_1K_{jk} + K_{j-1,k} + K_{j,k-1} \; , \tag{6.14}$$

where $j < k$ and for $j=k-1$, $H_{j,k-1}$ is properly defined as H_{jj} while

$$K_{jj} = 0 \tag{6.15}$$

for (6.14) to be correct. Also, $H_{0k} = 0$ and $K_{0k} = 0$.

For distinct elementary divisors $(\lambda - \alpha)^m$ and $(\lambda - \beta)^n$, we introduce new vectors w_r for which

$$Aw_r = \beta w_r + w_{r-1}, \quad w_0 = 0, \quad r = 1',2',\ldots,n' \tag{6.16}$$

and write $\beta = b_1 + ib_2$. The use of $1',2',\ldots,n'$ as the index set in the space $\mathcal{E}_\beta$ serves to remind us that the matrix defined by

$$H_{jr} = v_j w_r^* + w_r v_j^* \tag{6.17}$$

does not reduce to a matrix of the type (6.2) for any choice of j and r. (But, we do continue to set

$$H_{0r} = H_{j0} = 0.) \tag{6.18}$$

Calculating,

$$AH_{jr} + H_{jr}A^* = (a_1+b_1)H_{jr} + (a_2-b_2)K_{jr} + H_{j-1,r} + H_{j,r-1} \quad . \tag{6.19}$$

Similarly, from the definition

$$K_{jr} = i(v_j w_r^* - w_r v_j^*) \tag{6.20}$$

we obtain

$$AK_{jr} + K_{jr}A^* = -(a_2-b_2)\, H_{jr} + (a_1+b_1)K_{jr} + K_{j-1,r} + K_{j,r-1}. \tag{6.21}$$

For a single matrix A, some of the invariant spaces under the Lyapunov mapping

$$G \to H = AG + GA^* \equiv (A \otimes 1 + 1 \otimes \bar{A})\, \vec{G} \equiv \mathcal{L}_A(G) \tag{6.22}$$

may be exhibited in explicit form by a suitable ordering of the basis, with grade (so that H_{ij} has grade $i+j-1$, etc.) and lexicographic ordering within grade determining the ordering within the subspaces spanned by the H_{jk} and K_{jk}, respectively. An arbitrary ordering on the Jordan blocks permits an alternating ordering of the H_{jr} and K_{jr} with grade and lexicographic ordering on the index pairs (jr). The structure of the resulting Lyapunov matrix is exhibited for elementary divisors of orders 2 and 3 in (6.23).

$$
\begin{array}{l*{25}{c}}
 & H & & & K & H & & & & & & K & & & H & K & H & K & H & K & H & K & H & K & H & K \\
 & (11) & (12) & (22) & (12) & (1'1') & (1'2') & (1'3') & (2'2') & (2'3') & (3'3') & (1'2') & (1'3') & (2'3') & (11') & (11') & (12') & (12') & (21') & (21') & (13') & (13') & (22') & (22') & (23') & (23') \\
H\,(11) & 2a & 1 \\
(12) & & 2a & 1 \\
(22) & & & 2a \\
K\,(12) & & & & 2a \\
H\,(1'1') &&&&& 2b & 1 \\
(1'2') &&&&&& 2b & 1 & 2 \\
(1'3') &&&&&&& 2b && 1 \\
(2'2') &&&&&&&& 2b & 1 \\
(2'3') &&&&&&&&& 2b & 2 \\
(3'3') &&&&&&&&&& 2b \\
K\,(1'2') &&&&&&&&&&& 2b & 1 \\
(1'3') &&&&&&&&&&&& 2b & 1 \\
(2'3') &&&&&&&&&&&&& 2b \\
H\,(11') &&&&&&&&&&&&&& c & -d & 1 && 1 \\
K\,(11') &&&&&&&&&&&&&& d & c && 1 && 1 \\
H\,(12') &&&&&&&&&&&&&&&& c & -d &&& 1 && 1 \\
K\,(12') &&&&&&&&&&&&&&&& d & c &&&& 1 && 1 \\
H\,(21') &&&&&&&&&&&&&&&&&& c & -d &&& 1 \\
K\,(21') &&&&&&&&&&&&&&&&&& d & c &&&& 1 \\
H\,(13') &&&&&&&&&&&&&&&&&&&& c & -d &&& 1 \\
K\,(13') &&&&&&&&&&&&&&&&&&&& d & c &&&& 1 \\
H\,(22') &&&&&&&&&&&&&&&&&&&&&& c & -d & 1 \\
K\,(22') &&&&&&&&&&&&&&&&&&&&&& d & c && 1 \\
H\,(23') &&&&&&&&&&&&&&&&&&&&&&&& c & -d \\
K\,(23') &&&&&&&&&&&&&&&&&&&&&&&& d & c
\end{array}
$$

[$a = a_1$, $b = b_1$, $c = a_1 + b_1$ and $d = a_2 - b_2$.]

$A \otimes 1 + 1 \otimes \bar{A}$ when A has elementary divisors $(\lambda-a_1-ia_2)^2$ and $(\lambda-b_1-ib_2)^3$.

(6.23)

For the decomposition $\mathcal{V} = \mathcal{V}_\alpha \oplus \mathcal{V}_\beta$, with $\mathcal{V}_\alpha$ of dimension 2 and $\mathcal{V}_\beta$ of dimension 3, the decomposition is into

$\mathcal{V}_\alpha \otimes \bar{\mathcal{V}}_\alpha$, eigenvalue $\alpha + \bar{\alpha}$, elementary divisors of orders 1, 3; (6.24)

$\mathcal{V}_\beta \otimes \bar{\mathcal{V}}_\beta$, " $\beta + \bar{\beta}$, " " " " 1,3,5; (6.25)

$(\mathcal{V}_\alpha \otimes \bar{\mathcal{V}}_\beta) \times (\mathcal{V}_\beta \times \bar{\mathcal{V}}_\alpha)$, eigenvalue $\alpha + \bar{\beta}$ and $\beta + \bar{a}$, elementary divisors of orders 2,2,4,4. (6.26)

In the last space, the fact that we are dealing with a linear mapping of an n^2 dimensional real vector space into itself guarantees that each elementary divisors with eigenvalue $\alpha + \bar{\beta}$ is paired with one of the same order belonging to the complex conjugate eigenvalue $\bar{\alpha} + \beta$ and hence the allocation of the four orders is not in doubt, and the eigenvalue $\alpha + \bar{\beta}$ has elementary divisors of orders 2 and 4 and so does $\bar{\alpha} + \beta$.

7. Elementary Divisor Decomposition of $A \otimes 1 + 1 \otimes \bar{A}$

Equations (6.11) and (6.13) clearly imply that the H_{jj} and H_{jk} $(j < k)$ span a linear subspace of $\frac{1}{2}m(m+1)$ dimensions which is invariant under A. (But observe that this same subspace will not be invariant under $B \otimes 1 + 1 \otimes \bar{B}$ for all B.) Similarly, the $K_{jk}(j < k)$ span an invariant $\frac{1}{2}m(m-1)$ dimensional space and the Lyapunov transformation of hermitian matrices splits into a direct sum of transformations in these two spaces. The elementary divisors are, by Theorem 4.1, $(\lambda - 2a_1)^{2m-1}$, $(\lambda - 2a_1)^{2m-3}, \ldots, 3, 1$ and these must split into two sets since the elementary divisors of a direct sum are the aggregate of the elementary divisors of the summands. We now prove:

Theorem 7.1. The elementary divisors of $A \otimes 1 + 1 \otimes \bar{A}$, for A with a single elementary divisor $(\lambda - a_1 - ia_2)^m$, relative to the space spanned by the $H_{jk}(j \leqq k)$ are of degrees $2m-1$, $2m-5$, $2m-9, \ldots, 1$ if m is odd or 3 if m is even. The elementary divisors in the space spanned by the $K_{jk}(j \leqq k)$ are of degrees $2m-3$, $2m-7$, $2m-11, \ldots, 3$ if m is odd or 1 if m is even.

The proof parallels that of section 4.

Introducing the notation

$$(i,j) = K_{ij} \text{ and } Q = A \otimes 1 + 1 \otimes \bar{A} - 2a_1 \, 1 \otimes 1 \quad , \tag{7.1}$$

we have

$$Q(i,j) = (i-1,j) + (i,j-1) \quad . \tag{7.2}$$

The listing of the (i,j) by grade and lexicographically within grade gives for m = 2k-1:

Grade g	Basis Vectors	Number: n_g
1	-	0
2	(1,2)	1
3	(1,3)	1
4	(1,4), (2,3)	2
5	(1,5), (2,4)	2
6	(1,6), (2,5), (3,4)	3
7	(1,7), (2,6) (3,5)	3
.	.	.
.	.	.
.	.	.
m-3	(1,m-3), (2,m-4), (3,m-5),...,(k-2,k-1)	k-2
m-2	(1,m-2), (2,m-3), (3,m-4),...,(k-2,k)	k-2
m-1	(1,m-1), (2,m-2), (3,m-3),...,(k-1,k)	k-1
m	(1,m), (2,m-1), (3,m-2),...,(k-1,k+1)	k-1
m+1	(2,m), (3,m-1), (4,m-2),...,(k,k+1)	k-1
m+2	(3,m), (4,m-1), (5,m-2),...,(k,k+2)	k-2
m+3	(4,m), (5,m-1), (6,m-2),...,(k+1,k+2)	k-2
.	.	.
.	.	.
.	.	.
2m-5	(m-4,m), (m-3,m-1)	2
2m-4	(m-3,m), (m-2,m-1)	2
2m-3	(m-2,m)	1
2m-2	(m-1,m)	1
2m-1	-	0

(7.3)

For <u>m = 2k</u> the table is:

Grade g	Basis Vectors	Number: n_g
1	-	0
2	(1,2)	1
3	(1,3)	1
4	(1,4), (2,3)	2
5	(1,5), (2,4)	2
6	(1,6), (2,5), (3,4)	3
7	(1,7), (2,6), (3,5)	3
.	.	.
.	.	.
.	.	.
m-3	(1,m-3), (2,m-4),...,(k-2,k)	k-2
m-2	(1,m-2), (2,m-3),...,(k-1,k)	k-1
m-1	(1,m-1), (2,m-3),...,(k-1,k+1)	k-1
m	(1,m), (2,m-1),...,(k,k+1)	k
m+1	(2,m), (3,m-1),...,(k,k+2)	k-1
m+2	(3,m), (4,m-1),...,(k+1,k+2)	k-1
m+3	(4,m), (5,m-1),...,(k+1,k+3)	k-2
.	.	.
.	.	.
.	.	.
2m-5	(m-4,m), (m-3,m-1)	2
2m-4	(m-3,m), (m-2,m-1)	2
2m-3	(m-2,m)	1
2m-2	(m-1,m)	1
2m-1	-	0

(7.4)

Since Q transforms a basis vector of grade g into a sum of one or two of grade g-1, the block structure of Q is like that of N as displayed in (4.26) and we designate the block with rows referring to grade g and columns of grade g+r by $Q_{g,g+r}$. Then Q^r is zero except for the blocks $Q^{(r)}_{g,g+r}$.

There are 2m-3 grades, so g= 2,3,...,2m-2 and we set

$$q = 2m-2, \quad n_0 = n_1 = n_{2m-1} = n_{2m} = 0. \tag{7.5}$$

Also,

$$\text{for } g \text{ even, } n_g = \frac{1}{2}\min(g, 2m-g), \text{ and} \tag{7.6}$$

$$\text{for } g \text{ odd, } n_g = \frac{1}{2}\min(g-1, 2m-g-1) \quad . \tag{7.7}$$

We postpone establishing the necessary properties of the rank of the blocks, and proceed now to calculate rank N^r as in (5.29) and take second differences to get the degrees of the elementary divisors. Thus,

$$\begin{aligned} \text{rank } N^r &= \min(n_2, n_{2+r}) + \min(n_3, n_{3+r}) + \ldots + \min(n_{2m-2-r}, n_{2m-2}) \\ &= \sum_{j=0}^{2m-r} \min(n_j, n_{j+r}), \end{aligned} \tag{7.8}$$

where zero terms have been added to simplify the range of summation. For r even, say r=2s, j and j+r have the same parity and $\min(n_j, n_{j+r}) = \frac{1}{2}\min(j, 2m-j, j+r, 2m-j-r) = \frac{1}{2}\min(j, 2m-r-j)$ for j even and $\min(n_j, n_{j+r}) = \frac{1}{2}\min(j-1, 2m-r-1-j)$ for j odd.

Similarly, for r odd, say r=2s-1, $\min(n_j, n_{j+r}) = \frac{1}{2}\min(j, 2m-j, j+r-1, 2m-j-r-1) = \frac{1}{2}\min(j, 2m-r-1-j)$ for j even and $\min(n_j, n_{j+r}) = \frac{1}{2}\min(j-1, 2m-j-1, j+r, 2m-j-r) = \frac{1}{2}\min(j-1, 2m-j-r)$ for j odd.

Hence, for r=2s,

$$\begin{aligned} \text{rank } N^{2s} &= \sum_{i=0}^{m-s} \min(n_{2i}, n_{2i+2s}) + \sum_{i=1}^{m-s} \min(n_{2i-1}, n_{2i-1+2s}) \\ &= \frac{1}{2}\sum_{i=1}^{m-s} \min(2i, 2m-2s-2i) + \frac{1}{2}\sum_{i=1}^{m-s} \min(2i-2, 2m-2s-2i) \\ &= \sum_{i=1}^{m-s} \min(i, m-s-i) + \sum_{i=1}^{m-s} \min(i-1, m-s-i) \\ &= \frac{1}{2}(m-s-1)(m-s) = \frac{1}{2}[(m-s)^2 - (m-s)] \end{aligned} \tag{7.9}$$

in all cases, although the first summand is of the form

$1+2+\ldots+(p-1) + p + (p-1)+\ldots+1$ when $m-s=2p$ and of the form

$1+2+\ldots+(p-1) \quad +(p-1)+\ldots+1$ when $m-s=2p-1$ while the second

summand becomes

$1+2+\ldots+(p-1) \quad + \quad (p-1)+\ldots+1$ when $m-s=2p$

and it is $1+2+\ldots+(p-2) +(p-1) + (p-2)+\ldots+)$ when $m=2p-1$.

For $r=2s-1$,

$$\operatorname{rank} N^{2s-1} = \sum_{i=0}^{m-s-1} \min(n_{2i}, n_{2i+2s-1}) + \sum_{i=1}^{m-s+1} \min(n_{2i-1}, n_{2i+2s-2})$$

$$= \frac{1}{2} \sum_{i=0}^{m-s-1} \min(2i, 2m-2s-2i) + \frac{1}{2} \sum_{i=1}^{m-s+1} \min(2i-2, 2m-2s-2i+2)$$

$$= \sum_{i=1}^{m-s-1} \min(i, m-s-i) + \sum_{i=2}^{m-s} \min(i-1, m-s-i+1)$$

$$= 2 \sum_{i=1}^{m-s} \min(i, m-s-i) \tag{7.10}$$

$$= \frac{1}{2} (m-s)^2 \text{ for } m-s \text{ even and } \frac{1}{2} [(m-s)^2-1] \text{ for } m-s \text{ odd.}$$

We can summarize the last results in a table by introducing the congruence class of $2m-r$ (modulo 4), that is

$$2m-r = 4m_0+r_0, \qquad r_0 = 0,1,2 \text{ or } 3. \tag{7.11}$$

Then, for $r=2s$, $2m-r=2(m-s) \equiv 0 \pmod 4$ when $m-s$ is even and $2m-r \equiv 2 \pmod 4$ when $m-s$ is odd. For $r=2s-1$, $2m-r=2(m-s)+1$ is $\equiv 1 \pmod 4$ when $m-s$ is even and is $\equiv 3 \bmod 4$ when $m-s$ is odd. Hence,

$2m-r \equiv r_0 \pmod 4$	rank N^r
0	$\frac{1}{8}(2m-r-2)(2m-r)$
1	$\frac{1}{8}(2m-r-1)^2$
2	$\frac{1}{8}(2m-r-2)(2m-r)$
3	$\frac{1}{8}(2m-r-3)(2m-r+1)$

(7.12)

It follows that Δ_r = rank N^{r-1} - rank N^r has the values given:

$2m-r \equiv r_0 \pmod 4$	Δ_r
0	$\frac{1}{4}(2m-r)$
1	$\frac{1}{4}(2m-r-1)$
2	$\frac{1}{4}(2m-r-2)$
3	$\frac{1}{4}(2m-r+1)$

(7.13)

The elementary divisors are now obtained by differencing the last table:

$2m-r \equiv r_0 \pmod 4$	$p_r = \Delta_r - \Delta_{r+1}$
0	0
1	0
2	0
3	1

(7.14)

[Note: if Δ_r is read from $r_0=2$, then Δ_{r+1} is read from $r_0=1$ and not $r_0=3$. Also, r+1 replaces r in the expression in the table for Δ_r so that, for for $r_0=2$ we have $\Delta_r - \Delta_{r+1} = \frac{1}{4}(2m-r-2) - \frac{1}{4}(2m-r-2) = 0$ as given above.]

We thus conclude that in the space spanned by the $K_{jk}(j<k)$ there are single elementary divisors of degrees

$$r = 2m-3,\ 2m-7,\ 2m-11, \ldots,\ 3 \text{ if } m \text{ is odd or } 1 \text{ if } m \text{ is even} \qquad (7.15)$$

as stated in Theorem 7.1

For the elementary divisors in the space spanned by the $H_{jk}(j \lesseqgtr k)$, the fact that $A \otimes 1 + 1 \otimes \bar{A}$ splits into a direct sum guarantees that they are of orders 2m-1, 2m-5,... as required to complete the set specified in Thm. 4.1.

To complete the list of elementary divisors of $A \otimes 1 + 1 \otimes \bar{A}$, we need not only those of Thm. 7.1, which refer to the subspaces $\mathcal{E}_{\alpha_i} \otimes \mathcal{E}_{\bar{\alpha}_i}$, but also those referring to $(\mathcal{E}_{\alpha_i} \otimes \mathcal{E}_{\bar{\alpha}_j}) \oplus (\mathcal{E}_{\alpha_j} \otimes \mathcal{E}_{\bar{\alpha}_i})$.

These are given by Theorem 4.2 to be

$$\left(\lambda - (\alpha_i + \bar{\alpha}_j)\right)^{r_{ij}} \text{ and } \left(\lambda - (\bar{\alpha}_i + \alpha_j)\right)^{r_{ij}} \qquad (7.16)$$

for

$$r_{ij} = p_i+p_j-1,\ p_i+p_j-3,\ldots,|p_i-p_j| + 1\ , \tag{7.17}$$

where the Jordan blocks of A with eigenvalue α_i has order p_i.

With respect to the real space of hermitian matrices we would therefore have elementary divisors (over the reals) of

$$\left(\lambda^2 - (\alpha_i + \bar{\alpha}_i + \alpha_j + \bar{\alpha}_j)\lambda + |\alpha_i + \bar{\alpha}_j|^2\right)^{r_{ij}}. \tag{7.18}$$

If the complex decomposition is desired, it is only necessary to add $\pm$ i times equation (6.21) to (6.19) to get

$$A\,H^+_{jr} + H^+_{jr}\,A^* = (\alpha_i + \bar{\alpha}_j)\,H^+_{jr} + H^+_{j-1,r} + H^+_{j,r-1} \tag{7.19}$$

and

$$A\,H^-_{jr} + H^-_{jr}\,A^* = (\bar{\alpha}_i + \alpha_j)\,H^-_{jr} + H_{j-1,r} + H_{j,r-1} \tag{7.20}$$

where

$$H^{\pm}_{jr} = H_{jr} \mp i\,K_{jr}\ . \tag{7.21}$$

8. Rank of the $Q^{(r)}_{g,g+r}$

To supply the omitted parts of the proof of Thm. 7.1, we must show that each block submatrix on the r-th diagonal of N^r has maximum rank ($= \min(n_g, n_{g+r})$). This proves to be easier than in the general case of section 4 since, as exhibited in (8.3), N is symmetric about the secondary diagonal (from upper right to lower left corner) when its rows and columns are enumerated lexicographically within grade. Thus.

$$\left(Q^{(1)}_{g,g+1}\right)^T = Q^{(1)}_{2m-1-g,2m-g} \tag{8.1}$$

and, since there are 2m-3 grades and consequently 2m-4 of the $Q^{(1)}_{g,g+1}$, each of the blocks under consideration will either be of the form

$$Q = A_0 A_1 \ldots A_k [(B_1 B_2 \ldots B_h)\ (B_h^T \ldots B_2^T\ B_1^T)]$$

or the transpose of such a matrix ($h \geqslant 0$, $k \geqslant 0$, $0 < 2h+k \leqslant 2m-4$). The matrix in the square bracket is of the form MM^T and, since M is real, has rank equal to the rank of M. After proving that the rank of each A_i and B_j is the number of its rows, we will have MM^T nonsingular and, by the argument of (5.27), the rank of Q will be the number of rows in A_0 or, if the transpose of Q, the number of columns in A_0^T.

To establish (8.1) and

$$\text{rank } Q^{(1)}_{g,g+1} = n_g = \text{number of rows, for } g = 2,3,\ldots,m-1, \tag{8.2}$$

we simply exhibit the form of the matrices $Q^{(1)}_{g,g+1}$ and $Q^{(1)}_{2m-3-g,2m-2-g}$ for $g = 2,3,\ldots,m-1$ in (8.4) and (8.5). This completes the proof of the facts used in the last section to prove Thm. 7.1.

K \ K	(12)	(13)	(14)	(23)	(15)	(24)	(16)	(25)	(34)	(17)	(26)	(35)	(27)	(36)	(45)	(37)	(46)	(47)	(56)	(57)	(67)
(12)		1																			
(13)			1	1																	
(14)					1	1															
(23)						1															
(15)							1	1													
(24)								1	1												
(16)										1	1										
(25)											1	1									
(34)												1									
(17)													1								
(26)													1	1							
(35)														1	1						
(27)																1					
(36)																1	1				
(45)																	1				
(37)																		1			
(46)																		1	1		
(47)																				1	
(56)																				1	
(57)																					1
(67)																					

Component of $A \otimes 1 + 1 \otimes \bar{A} - 2a_1\, 1 \otimes 1$ in space of K_{jk} $(j < k)$ when A has a single elementary divisor of order $m = 7$: $(\lambda - a_1 - ia_2)^7$.

(8.3)

$Q^{(1)}_{g,g+1}$ is $\frac{g}{2}$ by $\frac{g}{2}$ for g even and $\frac{g-1}{2}$ by $\frac{g+1}{2}$ for odd, $(g=2,3,\ldots,m-1)$

$$
\begin{array}{ll|cccccccc|}
 & & (1,g+1) & (2,g) & (3,g-1) & \cdots & \overbrace{(h-1,h+3) \quad (h,h+2)}^{g=2h} & & \text{OR} & \overbrace{(h,h+3) \quad (h+1,h+2)}^{g=2h+1} \\
(1,g) & & 1 & 1 & & & & & & \\
(2,g-1) & & & 1 & 1 & & & & & \\
(3,g-2) & & & & 1 & & & & & \\
\vdots & & & & & \ddots & & & & \\
(h-2,h+3) & \multirow{3}{*}{$\left.\right\} g=2h$} & & & & & 1 & & & \\
(h-1,h+2) & & & & & & 1 \quad\quad 1 & & & \\
(h,h+1) & & & & & & \quad\quad\;\; 1 & & & \\
\text{OR} & & & & & & & & \text{OR} & \\
(h-2,h+4) & \multirow{3}{*}{$\left.\right\} g=2h+1$} & & & & & & & & \\
(h-1,h+3) & & & & & & & & & 1 \\
(h,h+2) & & & & & & & & & 1 \quad\quad 1 \\
\end{array}
$$

(8.4)

$Q^{(1)}_{2m-1-g,2m-g}$ is $\frac{g}{2}$ by $\frac{g}{2}$ for g even and is $\frac{g+1}{2}$ by $\frac{g-1}{2}$ for g odd, g=2,3,...,m-1

$$
\begin{array}{l|cccccccccc|}
 & (m-g+1,m) & (m-g+2,m-1) & (m-g+3,m-2) & \cdots & \overbrace{(m-h-1,m-h+2) \quad (m-h,m-k+1)}^{g=2h} & & \text{OR} & \overbrace{(m-h-2,m-h+2) \quad (m-h-1,m-h+1)}^{g=2h+1} & \\
(m-g,m) & 1 & & & & & & & & \\
(m-g+1,m-1) & 1 & 1 & & & & & & & \\
(m-g+2,m-2) & & 1 & 1 & & & & & & \\
\vdots & & & & \ddots & & & & & \\
\left.\begin{array}{l}(m-h-3,m-h+3)\\(m-h-2,m-h+2)\\(m-h-1,m-h+1)\end{array}\right\} g=2h & & & & & \begin{array}{cc} & \\ 1 & \\ 1 & 1\end{array} & & & & \\
\text{OR} & & & & & & & \text{OR} & & \\
\left.\begin{array}{l}(m-h-3,m-h+2)\\(m-h-2,m-h*1)\\(m-h-1,m-h)\end{array}\right\} g=2h+1 & & & & & & & & \begin{array}{cc}1 & \\ 1 & 1\\ & 1\end{array} & \\
\end{array}
\qquad (8.5)
$$

9. Elementary Divisors of $A \otimes 1 + 1 \otimes A$ in the Symmetric and Skewsymmetric Spaces

Whether or not A is real or complex, the restriction of the linear mapping

$$X \to Y = AX + XA^T \equiv (A \otimes 1 + 1 \otimes A)\, X \tag{9.1}$$

carries a symmetric matrix X into a symmetric matrix Y since

$$X^T = X \text{ implies } Y^T = X^T A^T + AX^T = XA^T + AX = Y \;. \tag{9.2}$$

Similarly,

$$X^T = -X \text{ implies } Y^T = -Y \tag{9.3}$$

so the n^2 dimensional vector space of matrices X is a direct sum of a "symmetric" linear subspace of dimension $\frac{1}{2}\, n(n+1)$ and a "skewsymmetric" space of dimension $\frac{1}{2}\, n(n-1)$. This decomposition is independent of the choice of A in contrast to the decomposition studied in the last two sections.

The technique already developed permits one to calculate straightforwardly the elementary divisor distribution in the two spaces. Specializing Thm. 4.2 to the case B = A, we have

if A has elementary divisors $(\lambda - \alpha_i)^{p_i}$, then $A \otimes 1 + 1 \otimes A$ has elementary divisors

$$(\lambda - 2\, a_i)^{r_i} \quad \text{for } r_i = 2p_i - 1,\ 2p_i - 3, \ldots, 1$$

and two of each of the elementary divisors

$$(\lambda - \alpha_i - \alpha_j)^{r_{ij}} \quad \text{for } i < j \text{ and } r_{ij} = p_i + p_j - 1, \\ p_i + p_j - 3, \ldots, |p_i - p_j| + 1. \tag{9.4}$$

For a matrix A of simple structure (that is, all $p_i = 1$), one easily sees that the eigenvectors associated with the $2\,\alpha_i$ are in the symmetric space and the two dimensional eigenspace associated with $\alpha_i + \alpha_j$, $i < j$, intersects both the symmetric and the skewsymmetric space in a one-dimensional space.

For a matrix with a single elementary divisor $(\lambda - \alpha)^n$, it appears plausible that the elementary divisors $(\lambda - 2\alpha)$, $(\lambda - 2\alpha)^3$, $(\lambda - 2\alpha)^5, \ldots, (\lambda - 2\alpha)^{2n-1}$ split into two sequences, with the sequence of orders

$$1, 5, 9, \ldots, \quad (2n-3) \text{ if n is even; } (2n-1) \text{ if n is odd)} \tag{9.5}$$

belonging to the skewsymmetric space if n is even and to the symmetric space when n is odd since the sum of these integers has the appropriate value $\frac{1}{2}\, n(n-1)$ when n is even and $\frac{1}{2}\, n(n+1)$ when n is odd. That this is indeed the case is a part of the following theorem.

Theorem 9.1. Let A have elementary divisors $(\lambda-\alpha_i)^{p_i}$. Then the symmetric component of $A \otimes 1 + 1 \otimes A$ has elementary divisors

$$(\lambda-2\alpha_i)^{s_i} \text{ for } s_i = (2p_i-1), (2p_i-5), \ldots, (3 \text{ for } p_i \text{ even}; 1 \text{ for } p_i \text{ odd})$$

and

$$(\lambda-\alpha_i-\alpha_j)^{r_{ij}} \text{ for } i < j \text{ and } r_{ij} = p_i+p_j-1,\ p_i+p_j-3, \ldots, |p_i-p_j|+1.$$

Replacing the range of s_i by

$$s_i = (2p_i-3), (2p_i-7), \ldots, (1 \text{ for } p_i \text{ even}; 3 \text{ for } p_i \text{ odd}),$$

the elementary divisors of the skewsymmetric component are $(\lambda-2\alpha_i)^{s_i}$ and $(\lambda-\alpha_i\lambda\alpha_j)^{r_{ij}}$, $i < j$.

It should be noted that when A is real and α_i and $\bar{\alpha}_i$ are distinct, the corresponding elementary divisors in the symmetric as well as in the skew-symmetric space occur in conjugate pairs as is necessary since a real n by n matrix is the sum of a real symmetric and a real skewsymmetric one.

The necessary modifications of sections 6, 7, and 8 to establish the theorem will now be sketched.

Let Jordan bases for elementary divisor subspaces be given as in (6.10) and (6.16). Defining H_{jk} for $j \leqslant k$ as in (6.2) and (6.3), with the star interpreted as the transpose,

$$AH_{jk} + H_{jk}A^T = 2\,\alpha\, H_{jk} + H_{j-1,k} + H_{j,k-1}, \quad j \leqslant k, \tag{9.6}$$

replacing (6.12) and (6.13). Clearly, the elementary divisors in the subspace spanned by the H_{jk} have the same degree as before, with the eigenvalue $\alpha + \bar{\alpha}$ replaced by 2α. Hence the elementary divisors in the symmetric component of the space are calculated for Thm. 9.1 by the same argument as for the first sentence in Thm. 7.1.

Similarly, defining

$$K_{ij} = v_i v_j^T - v_j v_i^T, \quad i<j, \tag{9.7}$$

gives a basis for the skewsymmetric subspace of $\mathcal{V}_\alpha \otimes \mathcal{V}_\alpha$ and

$$AK_{ij} + K_{ij}A^T = 2\,\alpha\, K_{ij} + K_{i-1,j} + K_{i,j-1}, \quad i<j, \tag{9.8}$$

so the remaining elementary divisors do indeed refer to the skewsymmetric component.

For the subspace

$$\mathcal{V}_\alpha \otimes \mathcal{V}_\beta \oplus \mathcal{V}_\beta \otimes \mathcal{V}_\alpha \tag{9.9}$$

analogous to the cross product terms in (6.9), we need a new discussion since in the complex case there was no invariant decomposition and the elementary divisors could be listed by excluding those already identified from the known complete list.

Let

$$H_{jr} = v_j w_r^T + w_r v_j^T \tag{9.10}$$

and

$$K_{jr} = v_j w_r^T - w_r v_j^T. \tag{9.11}$$

Then

$$AH_{jr} + H_{jr}A^T = (\alpha + \beta)H_{jr} + H_{j-1,r} + H_{j,r-1} \tag{9.12}$$

with $j=1,2,\ldots,m$, $r=1',2',\ldots,m'$, and

$$AK_{jr} + K_{jr}A^T = (\alpha + \beta)K_{jr} + K_{j-1,r} + K_{j,r-1} \tag{9.13}$$

in contradistinction to (6.19) and (6.21), where the subspaces spanned by the H_{jr} and by the K_{jr} were <u>not</u> separately invariant as here. Since (9.12) and (9.13) are identical in form, the elementary divisors of the linear transformations they describe must be the same and one of each of the pairs given by the general theory of section 4 can be assigned to the symmetric and the other to the skewsymmetric component.

Strictly speaking, the "assignment" of an elementary divisor "block" to the symmetric component and one of the same order to the skewsymmetric component expresses in a noninvariant way the following geometrical situation. Let $\alpha = \alpha_i$ and $\beta = \alpha_j$ $(j > i)$ be eigenvalues of A and assume that $\alpha + \beta \neq 2\alpha_k$ for any eigenvalue α_k. Then each of the linear subspaces

$$\mathcal{S}^{(r)} = \left\{X;\ L^r_{\alpha+\beta}(X) = 0\right\}, \tag{9.14}$$

is of even dimension, where

$$L_{\alpha+\beta}(X) = AX + XA^T - (\alpha + \beta)X \tag{9.15}$$

and r denotes the r-th iterate (power) of the linear operator $L_{\alpha+\beta}$. Moreover, each $\mathcal{S}^{(r)}$ is a direct sum of symmetric and skewsymmetric components of the same dimension. For r=1, our conclusion is that, with α and β distinct eigenvalues of A and $\alpha + \beta$ not equal to $2\alpha_k$, the number of linearly independent symmetric matrices X which satisfy

$$AX + XA^T = (\alpha + \beta)X \tag{9.16}$$

is the same as the number of linearly independent skewsymmatrices which satisfy it. (Of course, every solution is a sum of a symmetric and a skew-symmetric solution.)

The same statements are true for the matrices which satisfy

$$A^2X + 2AXA^T + X(A^T)^2 = 2(\alpha + \beta)\ (AX + XA^T) - (\alpha + \beta)^2\ X \tag{9.17}$$

for this is a defining equation for $\mathcal{S}^{(2)}$.

The special case in which there are eigenvalues α_i and α_j such that $\alpha_i + \alpha_j = 0$ yet A is nonsingular (so that $2\alpha_k \neq 0$ for any eigenvalue α_k) is of interest for then the much studied matrix equation

$$AX + XA = B \tag{9.18}$$

has either 1) no solution, or (for suitable B), 2) an infinite number of solutions. In the second case, the solutions of the homogeneous equation $AX + XA = 0$ have a basis consisting of an equal number of symmetric and skewsymmetric solutions. When $\alpha_i + \alpha_j = 0$ and A is singular, there are more symmetric than skewsymmetric matrices in a basis restricted to these two types since each elementary divisor of degree m with $\alpha = 0$ contributes $\frac{1}{2}\,m(m+1)$ symmetric but only $\frac{1}{2}\,m(m-1)$ skewsymmetric matrices.

[In passing, we note that the determination of canonical forms of a maximal set of matrices of order n satisfying

$$A_p\,A_q = -\,A_q\,A_p\ ,\quad \big(p,q=1,\ldots,N(n)\big) \tag{9.19}$$

often subject to the additional condition

$$(A_i)^2 = 1_n \text{ or } -1_n\ , \tag{9.20}$$

has been much studied in connection with the quantum theory and the spin representation of the Lorentz group. With $(A_i)^2 = 1$, the eigenvalues are +1 and -1, they occur in equal number if $(A_q)^{-1}\,A_p\,A_q = -A_p$ for $q \neq p$ (so n is necessarily even: $n = 2m$) and the elementary divisors are simple so that the linear manifold of matrices X such that

$$AX = -XA \tag{9.21}$$

is of dimension $2m^2$ and the symmetric subspace has dimension m^2 as does the skewsymmetric one. The conclusions are trivially verified by taking A_p in canonical form $1_m \oplus (-1_m)$ and observing that an X satisfying (9.21) has arbitrary off diagonal blocks but zero diagonal blocks.]

10. Results of Lewis and Taussky

In [36] D. C. Lewis, Jr. and Olga Taussky have established a number of results related to the present investigation. In this section we establish and slightly extend their main results in the present context.* Results of [36] have been re-phrased since we have been discussing the matrix mapping $G \rightarrow AG + GA^*$ for $G = G^*$, while Lewis and Taussky consider $G \rightarrow GA + A^*G$ for $G = G^*$.

To obtain information about the nonsingularity and signature of matrices expressed as linear combinations of the matrices used as a basis is in general just as difficult as obtaining this information from a matrix specified by its elements.

For the special cases we shall consider, it will be sufficient to observe that

$$\text{rank} \sum_{i=1}^{p} c_i v_i v_i^T = \text{rank} \sum_{i=1}^{p} c_i v_i v_i^* = p$$

under the hypotheses that 1) the vectors v_i $(i=1,\ldots,p)$ are linearly independent, and 2) all $c_i \neq 0$. When the c_i are all real and positive, $\sum_{i=1}^{p} c_i v_i v_i^*$ is positive semidefinite and is nonsingular and hence definite when $p=n$ and the v_i are therefore a basis.

We now take the v_i to be a linearly independent set of eigenvectors of A so that

$$A v_i = \alpha_i v_i \quad , \quad (i=1,\ldots,p) \tag{10.1}$$

where the α_i are not necessarily distinct but there are at least as many elementary divisors as there are vectors v_i with the same eigenvalue. Setting

$$G = \sum_{i=1}^{p} c_i v_i v_i^* \quad , \quad \bar{c}_i = c_i \neq 0 \quad , \tag{10.2}$$

G is hermitian and

$$AG = \Sigma\, c_i \alpha_i v_i v_i^* \tag{10.3}$$

so that, for an arbitrary complex number σ,

---

*The three principal points of view of matrix algebra regard a matrix as 1) describing a mapping of a vector space (which remains in the center of view), as 2) an element of the ring of (square) matrices, and 3) as determining mappings $X \rightarrow AX$ and $Y \rightarrow YA$ of vector spaces (perhaps of rings) of matrices X and Y of suitable orders. Lewis and Taussky rely mainly on the techniques of 2) while here we are using 3). For questions involving nonsingularity, 1) and 2) are frequently more convenient yet some questions (uniqueness, etc.) are best understood by 3).

$$\sigma AG + \bar{\sigma} GA^* - 2\lambda G = \Sigma c_i (\sigma \alpha_i + \bar{\sigma} \bar{\alpha}_i - 2\lambda) v_i v_i^*. \tag{10.4}$$

The right member is singular if and only if one of the coefficients is. Hence we have the following result.

Theorem 10.1 C. The matrices G given by (10.2), which depend on the eigenvector structure of A but not on σ, are such that

$$\text{rank } G = p \lesseqgtr \text{number of elementary divisors } (\lambda - \alpha_i)^{p_i} \text{ of } A, \tag{10.5}$$

and

$$\det(\sigma AG + \bar{\sigma} GA^* - 2\lambda G) = 0 \tag{10.6}$$

has the real parts of the $\sigma \alpha_i$ as its roots. When all the elementary divisors of A are simple, it is possible to take $p = n$ and G is definite for $c_i > 0$, $i=1,\ldots,n$.

For A real we can insist that G be real by 1) choosing v_i real for each real α_i used in the selection of eigenvectors and 2) whenever an α_i which is not real occurs, pairing it with $\alpha_j = \bar{\alpha}_i$ and taking $v_j = \bar{v}_i$ and $c_i = c_j$ so that (10.2) becomes

$$G = \sum_{\alpha \text{ real}} c_i v_i v_i^* + \sum_{\alpha \text{ not real}} c_j (v_j v_j^* + \bar{v}_j v_j^T), \tag{10.7}$$

where the c_i and c_j are real. For such G we have

Theorem 10.1 R. The matrices G given by (10.7) are real and continue to have the other properties stated in Theorem 10.1 C.

The results of Lewis and Taussky in their Theorem 1 have been generalized in the following respects: a class of matrices G, not necessarily definite, has been given by means of the parameters c_i; an arbitrary subset of the eigenvalues has been used; and, the role played by simple elementary divisors has perhaps been clarified. For the Lyapunov mapping

$$G \to \mathcal{L}_{\sigma A}(G) = \sigma AG + \bar{\sigma} GA^*, \tag{10.8}$$

a λ satisfying (10.6) is not necessarily an eigenvalue since we require there only the vanishing of the determinant rather than of the matrix. The matrices $v_i v_i^*$ are, however, eigenvectors under (10.8) as well as instances of matrices to which Thm. 10.1 C is applicable.

The impossibility of extending Thm. 10.1 R to obtain a definite G when A has a nonsimple elementary divisor is established by an example given by Lewis and Taussky which is trivially generalized as follows to show the impossibility of extending Thm. 10.1 C to obtain a definite G. Thus if

$$A = \begin{Vmatrix} 1 & 1 \\ 0 & 1 \end{Vmatrix} \text{ and } G = \begin{Vmatrix} a & \beta \\ \bar{\beta} & c \end{Vmatrix} \text{ with a and c real and } \beta \text{ complex,}$$

$$AG + GA^* - 2\lambda G = \left\| \begin{matrix} 2a(1-\lambda) + \beta+\bar{\beta} & c+2\beta(1-\lambda) \\ & \\ c+2\bar{\beta}(1-\lambda) & 2c(1-\lambda) \end{matrix} \right\| \tag{10.9}$$

Hence,

$$\det (AG + GA^* - 2\lambda G) = 4(ac-\beta\bar{\beta})\ (1-\lambda)^2 - c^2, \tag{10.10}$$

and $\lambda=1$ is a zero of this polynomial if and only if $c = 0$ and then it is a double zero. (When $c = 0$ and $\det G = 0$, so $\beta = 0$, the polynomial vanishes identically as is otherwise obvious by taking $G = a\, v_1\, v_1^*$, which we know makes the matrix of (10. 9) of rank $\leqslant 1$ for all values of λ.) Since for $\beta \neq 0$,

$$x^*gx = a\ \bar{x}_1\ x_1 + \beta\ \bar{x}_1\ x_2 + \bar{\beta}\ x_1\ \bar{x}_2 \tag{10.11}$$

assumes both positive and negative values, there is no hope of obtaining a semidefinite G for which the roots of (10.6) are the real parts of $\sigma\,\alpha_i$ with their proper multiplicities. The example, however, suggests that we may be able to prove the following theorem.

Theorem 10.2. Let A have eigenvalues α_i, $i=1,\ldots,n$ and let σ be an arbitrary complex number. Then, without restriction on the elementary divisors of A, there exists a nonsingular hermitian matrix G, which may be chosen independent of σ and is such that (10.6) has as its roots the real parts of the $\sigma\,\alpha_i$ for $i=1,2,\ldots,n$. When the elementary divisors of A are not simple it may be impossible to choose G definite. When A is real, G may be chosen real.

To construct an hermitian G satisfying the conditions of the theorem, we may consider only matrices which are a direct sum of matrices transforming the spaces $\mathcal{E}_\alpha \otimes \bar{\mathcal{E}}_\alpha$ thus gaining the nonsingularity of the entire matrix from that of its component summands. We now write the definitions (6.2) and (6.3) in the form

$$(j,k) = v_j\ v_k^* + v_k\ v_j^*, \tag{10.12}$$

where it is convenient to allow $j > k$ as well as $j \leqslant k$ and $j,k=1,\ldots,m$ with m the dimension of $\mathcal{E}_\alpha$ in which the v_j form a Jordan basis.

Setting

$$G = \sum_{j=1}^{p} c_j\ (j,m+1-j) \text{ if } m = 2p \tag{10.13}$$

and

$$G = \sum_{j=1}^{p+1} c_j(j,m+1-j) \text{ if } m = 2p+1, \tag{10.14}$$

gives a hermitian matrix for all real values of the c_j. To establish the nonsingularity of G, it is sufficient to refer the v_j to themselves as a basis* and then G has the form shown in (10.15).

$$\left\|\begin{array}{ccccccccc}
0 & 0 & \cdots & 0 & (0) & 0 & \cdots & 0 & c_1 \\
0 & 0 & \cdots & 0 & (0) & 0 & \cdots & c_2 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & 0 & (0) & c_p & & 0 & 0 \\
(0) & (0) & \cdots & (0) & \boxed{2c_{p+1}} & (0) & & (0) & (0) \\
0 & 0 & & c_p & (0) & 0 & & 0 & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & c_2 & \cdots & 0 & (0) & 0 & \cdots & 0 & 0 \\
c_1 & 0 & \cdots & 0 & (0) & 0 & \cdots & 0 & 0
\end{array}\right\| \tag{10.15}$$

if m is odd

[order m = 2p or 2p+1]

For all real choices of the $c_j \neq 0$, a matrix of the form (10.15) has minimum signature (cf. Jacobson [32], ch. V, § 11); that is, a totally isotropic space is defined by the equations $x_1 = x_2 = \ldots = x_p = 0$ so the number of plus and minus signs in a diagonal representation is either the same (when m is even) or differs by plus or minus one according to the sign of c_{p+1} (when m is odd). For the entire matrix G, which is a direct sum of matrices of the form (10.15), the signature varies in steps of two between the number of elementary divisors and its negative. We do not attempt to identify all possible signatures for G since for matrices with off-diagonal blocks (i.e., operating in $(\mathcal{E}_\alpha \otimes \overline{\mathcal{E}}_\beta) \oplus (\mathcal{E}_\beta \otimes \overline{\mathcal{E}}_\alpha)$ for $\mathcal{E}_\alpha \neq \mathcal{E}_\beta$) as well as diagonal blocks even the determination of nonsingularity is difficult.

Since Jordan bases for σA are identical with those of A and since the eigenvalues of σA are $\sigma\alpha_i$, it is a mere notational convenience to set $\sigma = 1$ for the following argument.

Alternatively, choose T so that $Tv_i = e_i$ = (the vector with the i-th component one and other components zero) and observe the form of TGT to be that tabulated.

Calculating, with (6.11) and (6.13) summarized by

$$A(j,k) + (j,k)A^* = 2a(j,k) + (j-1,k) + (j,k-1)\ , \tag{10.16}$$

where

$$a = \text{real part of } \alpha\ , \tag{10.17}$$

we get

$$\begin{aligned} AG + GA^* - 2\lambda G = {} & \sum_{j=1}^{p} 2c_j(a-\lambda)\,(j,m+1-j) \\ & + \sum_{k=1}^{p-1} (c_k + c_{k+1})\,(k,m-k) \\ & + c_p(p,p), \text{ for } m = 2p \end{aligned} \tag{10.18}$$

and

$$\begin{aligned} AG + GA^* - 2\lambda G = {} & \sum_{j=1}^{p+1} 2c_j(a-\lambda)\,(j,m+1-j) \\ & + \sum_{k=1}^{p-1} (c_k + c_{k+1})\,(k,m-k) \\ & + (c_p + c_{p+1})\,(p,p+1)\ , \text{ if } m = 2p+1\ . \end{aligned} \tag{10.19}$$

With the same choice of basis as in (10.15), the matrix $AG + GA^* - 2\lambda G$ is displayed in (10.20), the determinant clearly has $(a-\lambda)^m$ as a factor and the multiplicity of the root a is the same as that of α in A.

$$\left\|\begin{array}{ccccccccccc}
 & & & & & & & & & c_1+c_2 & 2c_1(a-\lambda) \\
 & & & & & & & & c_2+c_3 & 2c_2(a-\lambda) & \\
 & & & & & & & \iddots & \iddots & & \\
 & & & & & & c_{p-1}+c_p & \iddots & & & \\
 & & & & & \boxed{c_p+2c_{p+1}} & 2c_p(a-\lambda) & & & & \\
 & & & & \boxed{c_p+2c_{p+1}} & \boxed{4c_{p+1}(a-\lambda)} & & & & & \\
 & & & c_{p-1}+c_p & 2c_p(a-\lambda) & & & & & & \\
 & & \iddots & \iddots & & & & & & & \\
 & c_2+c_3 & \iddots & & & & & & & & \\
c_1+c_2 & 2c_2(a-\lambda) & & & & & & & & & \\
2c_1(a-\lambda) & & & & & & & & & &
\end{array}\right\|$$

[If $m = 2p$, $2c_p$ appears in the (p,p) position.] (10.20)
[If $m = 2p+1$, central elements are as shown in boxes.]

To complete the proof of Theorem 10.2 we observe that, for A real, each Jordan block with nonreal eigenvalue α is paired with one of the same size with eigenvalue $\bar{\alpha}$ so that we can always use $\bar{v}_j$ as a Jordan basis in $\mathcal{E}_{\bar{\alpha}}$ and the same c_j in $\mathcal{E}_{\bar{\alpha}} \otimes \bar{\mathcal{E}}_{\bar{\alpha}}$ as in $\mathcal{E}_{\alpha} \otimes \bar{\mathcal{E}}_{\alpha}$. Then either α is real and the v_j and hence G can be chosen real or G always occurs with $\bar{G}$ so only a sum of real terms $G + \bar{G}$ occurs.

11. Additional Results of Lewis and Taussky

Another direction in which Lewis and Taussky retain the definiteness of the matrix G while removing the condition that A has simple elementary divisors is to ask only that the zeros of (10.6) be "close" to the real parts of $\sigma\alpha_i$. We parallel and trivially generalize their argument leading to their Theorem 2.

Precisely, let $\epsilon > 0$ be any nonzero positive number and call w_j an ϵ-Jordan basis for a single elementary divisor of A if

$$A\, w_j = \alpha\, w_j + \epsilon\, w_{j-1} \quad , \quad j=1,2,\ldots,m \tag{11.1}$$

(To get such a basis from the vectors satisfying (6.10), it suffices to set

$$w_j = \epsilon^j\, v_j \quad ; \tag{11.2}$$

which is completely innocuous from a theoretical standpoint but which involves drastic scaling for ϵ small and is therefore non-trivial numerically.)

Taking

$$G = \sum_{j=1}^{m} c_j\, w_j\, w_j^* \text{ for real } \epsilon_j \neq 0 \quad , \tag{11.3}$$

we find that

$$\sigma\, A\, G + \bar{\sigma}\, G\, A^* - 2\,\lambda G = \sum_{j=1}^{m} (\sigma\alpha + \bar{\sigma}\,\bar{\alpha} - 2\lambda)\, c_j\, w_j\, w_j^* \tag{11.4}$$

$$+\ \epsilon \sum_{j=2}^{m} (\sigma\, c_j\, w_{j-1}\, w_j^* + \bar{\sigma}\, c_j\, w_j\, w_{j-1}^*) .$$

Hence, setting

$$x = \mathrm{Re}\,(\sigma\,\alpha) - \lambda\ , \tag{11.5}$$

and referring the w_j to themselves as a basis,

$$\det\,(\sigma AG + \bar{\sigma}GA^* - 2\lambda G)$$

$$= 2^m\, c_1 c_2 \cdots c_m \cdot \det \begin{Vmatrix} x & \frac{\epsilon\sigma}{2}\sqrt{\frac{c_2}{c_1}} & & & \\ \frac{\epsilon\bar{\sigma}}{2}\sqrt{\frac{c_2}{c_1}} & x & \frac{\epsilon\sigma}{2}\sqrt{\frac{c_3}{c_2}} & & \\ & \frac{\epsilon\bar{\sigma}}{2}\sqrt{\frac{c_3}{c_2}} & x & \ddots & \\ & & \ddots & \ddots & \end{Vmatrix} \tag{11.6}$$

The Gershgorin theorem states that all values of x which make this determinant equal to zero lie in the union of the m circles

$$\left| x \right| \leqslant \left| \frac{\varepsilon\,\sigma}{2} \sqrt{\frac{c_2}{c_1}} \right| , \tag{11.7}$$

$$\left| x \right| \leqslant \left| \frac{\varepsilon\,\bar{\sigma}}{2} \sqrt{\frac{c_j}{c_{j-1}}} \right| + \left| \frac{\varepsilon\,\sigma}{2} \sqrt{\frac{c_{j+1}}{c_j}} \right| , \; j=2,\ldots,m-1, \tag{11.8}$$

and

$$\left| x \right| \leqslant \left| \frac{\varepsilon\,\bar{\sigma}}{2} \sqrt{\frac{c_m}{c_{m-1}}} \right| . \tag{11.9}$$

Setting

$$c = \max \left| \frac{c_j}{c_{j-1}} \right|^{\frac{1}{2}} , \; j=2,\ldots,m , \tag{11.10}$$

we have the following theorem.

Theorem 11.1 For given A and $\varepsilon > 0$, with G any hermitian matrix obtained by summing matrices of the form (11.3) over the elementary divisors of A, the n zeros of

$$\det\,(\sigma A G + \bar{\sigma} G A^{*} - 2\lambda G) = 0 \tag{11.11}$$

will differ from the n real numbers $\mathrm{Re}(\sigma\alpha_i)$ by no more than $|\sigma| c\varepsilon$. When all the c_i are chosen positive, the matrix G is positive definite and the zeros are all real. For A real, summing conjugate matrices for conjugate eigenvalues will yield a real G and in any event the class of G used above depends on A and on ε (through the ε-Jordan basis) but not on σ. By choosing the signs of the c_i, the signature of the nonsingular G may be arbitrarily assigned.

Making use of the fact that a positive definite hermitian matrix is the square of a matrix of the same type, we can set $G = H^2$ with $H^{*}=H$ positive definite and obtain

$$\sigma A G + \bar{\sigma} G A^{*} - 2\lambda G = H[(\sigma H^{-1} A H) + (\sigma H^{-1} A H)^{*} - 2\lambda 1]\, H. \tag{11.12}$$

Hence, taking the c_j in (11.3) and (11.10) equal to one and choosing $\sigma = 1$ and $\sigma = -i$ we have, on taking determinants, a result essentially equivalent to corollaries 2.1, 2.2 and 2.3 of [36].

Theorem 11.2 Given A and $\varepsilon \geqslant 0$ with equality permitted only when the elementary divisors of A are simple, there exists a positive definite hermitian matrix H such that the similarity transform

$$A_H = H^{-1} A H \tag{11.13}$$

of A has for its hermitian and skewhermitian parts

$$A_H^+ = \frac{1}{2}(A_H + A_H^*) \text{ and } A_H^- = -\frac{i}{2}(A_H - A_H^*) \tag{11.14}$$

matrices with real eigenvalues

$$\alpha_i^+ = (\text{Re } \alpha_i) + \beta_i \quad \text{and} \quad \alpha_i^- = (\mathcal{I}m\ \alpha_i) + \gamma_i\ , \tag{11.15}$$

where $|\beta_i| \leqslant \varepsilon$ and $|\gamma_i| \leqslant \varepsilon$. For A nonsingular or with simple elementary divisors and ε chosen so that

$$\varepsilon < \min_i |\text{ Re } \alpha_i | \text{ and } \varepsilon < \min_i |\mathcal{I}m\ \alpha_i|\ , \tag{11.16}$$

A_H^+ and $HA_H^+H = \frac{1}{2}(AG + GA^*)$ have the same signature as $\|\text{ diag }(\text{Re } \alpha_i)\|$ and, simultaneously, A_H^- and $HA_H^- H = -\frac{i}{2}(AG - GA^*)$ have the same signature as $\|\text{ diag }(\mathcal{I}m\ \alpha_i)\|$, where $G = H^2$ is positive definite.

These results may be specialized to real matrices A and symmetric and skewsymmetric matrices A_H^+ and iA_H^- and, indeed, it is in this form that they are given in [36] with a complex number version indicated.

12. The Basic Lyapunov Stability Relationships

In his famous monograph, "The General Problem of Stability of Motion," [38] a profound connection between the signature of a hermitian matrix and the sign of the real parts of the eigenvalues of a general matrix was proved. The literature of the subject seems always to rely on arguments involving differential equations (as in Gantmacher [18], vol. II, pp. 185-190) or an integral representation of a matrix (as in Bellman [6], p. 243) and often uses a parameter which is allowed to approach infinity as a limit. The purpose of this section is to give an ab initio discussion of the Lyapunov mapping without this reliance on nonalgebraic methods. Ideally, the dependence on the complex number field should be avoided but this has not been attained.

Recently, Olga Taussky [58-61] extended the Lyapunov result by algebraic methods.* A. M. Ostrowski and Hans Schneider have made further extensions and the priority of these three authors is acknowledged. The lectures by Taussky and Ostrowski at the Gatlinburg Matrix Computation Conference in April 1961 were heard by the author of this paper and preprints of [60] and [61] were kindly made available by Taussky. While some of the results of this section may extend previous results, the extent of this will not be clear until the appearance of [47].

In the following discussion the attempt is made to reduce to exceedingly simple terms the main results, separating them into a number of lemmas which sometimes deliberately overlap one another.

Lemma 12.1. If $G = G^*$ and $AG + GA^*$ is definite, then both G and A are nonsingular.

If G is singular, $Gx = 0$ for an $x \neq 0$ so $x^*G^* = x^*G = 0$ and $x^*(AG + GA^*)x = 0$, contradicting the definiteness of $AG + GA^*$. Similarly, if A were singular, $uA = 0$ for some $u \neq 0$ and $u(AG + GA^*)u^* = 0$, again giving a contradiction.

This proof was suggested by that of Taussky in [60] and generalizes it by using a "coordinate-free" argument, by including the observation that A also cannot be singular and by avoiding the reduction of the general result to the case in which $AG + GA^* = \pm 1$, as could be done.

While we did not explicitly assume A nonsingular, it was a consequence that none of its eigenvalues α_i could be zero. It might be conjectured that even the stronger result $\alpha_i + \bar{\alpha}_k \neq 0$ but this is false as follows trivially by taking A diagonal real and $G = A$. If, however, both G and $AG + GA^*$ are definite, it is a consequence of the Lyapunov result that $\alpha_i + \bar{\alpha}_k \neq 0$. A simpler proof of this than is to be found in the literature (cf. [18] or [6]) is given in Lemma 12.8.

---

*It was the abstract [58] of Taussky's which interested the present author in this field and her work was also mentioned by Ostrowski in his Gatlinburg lecture as having stimulated the work to be reported in [47].

We recall the definition:

A is stable if its eigenvalues have negative real parts (12.1)

Lemma 12.2. If A is stable and $H = AG + GA^*$ is negative definite, then G is positive definite.

Let $A_0 = TAT^{-1}$ and $G_0 = TGT^*$, where T is chosen so that $THT^* = -1$. Then $A_0G_0 + G_0A_0^* = -1$. With a suitably chosen unitary matrix U (so $UU^* = 1$), we can make $A_1 = UA_0U^*$ upper triangular and have

$$A_1G_1 + G_1A_1^* = -1 \tag{12.2}$$

with $G_1 = UGU^*$. Now let $A_2(t)$ agree with A_1 on the diagonal but contain the parameter t as a multiplier of each off-diagonal element. Then $A_2(1) = A_1$, $A_2(t)$ has the same eigenvalues as A_1 (so is stable) and the sum of two complex numbers each having negative real part can never be zero. Hence for each t in the interval $0 \leqslant t \leqslant 1$ there is a unique and therefore necessarily hermitian $G_2(t)$ satisfying

$$A_2(t)\ G_2(t) + G_2(t)\ A_2^*(t) = -1\ . \tag{12.3}$$

By Lemma 12.1, $G_2(t)$ cannot be singular for any t. The (real) eigenvalues of $G_2(t)$ are continuous functions of t and never zero so that the number which are positive is independent of t. Hence $G_2(1)$ has the same signature as

$$G_2(0) = \|\ \text{diag}\ (-2\ \text{Re}\ \alpha_i)^{-1}\ \|\ , \tag{12.4}$$

which is positive definite. Since G has the same signature as G_0 and $G_1 = G_2(1)$, it is positive definite as was to be proved.

In the proof of the last lemma, we required A to be stable only 1) to guarantee the existence of a (continuous) solution $G_2(t)$ of (12.3) and 2) to identify the signature of $G_2(0)$ in (12.4). The same proof therefore gives the next lemma if we guarantee the existence (and uniqueness) of $G_2(t)$ by assuming $\alpha_i + \bar{\alpha}_k \neq 0$.

Lemma 12.3. (Taussky [61], unpublished; generalizing Lyapunov) Let A have eigenvalues α_i and suppose $\alpha_i + \bar{\alpha}_k \neq 0$ for any $i, k=1,\ldots,n$. Then if $G = G^*$ is any hermitian matrix for which $AG + GA^*$ is positive definite, G has the same signature as $\|\ \text{diag}\ (\text{Re}\ \alpha_i)\ \|$.

The last lemma is unsatisfactory in that we do not yet know (but, cf. Lemma 12.10) whether or not the hypothesis $\alpha_i + \bar{\alpha}_k \neq 0$ is needed for the conclusion to hold. Some insight is gained from the special case in which $k = i : \alpha_i + \bar{\alpha}_i = 0$.

Lemma 12.4. (Ostrowski) If A has a pure imaginary eigenvalue α ($\alpha + \bar{\alpha} = 0$ and $\alpha = 0$ is permitted), then AG + GA* cannot be definite for any hermitian G.

For, $uA = \alpha u$ for a nonzero u and hence $u(AG + GA^*)u^* = \alpha uGu^* + \bar{\alpha} uGu^* = 0$ so that AG + GA* is not definite for any $G = G^*$.

If A has no pure imaginary eigenvalues but $\alpha_i + \bar{\alpha}_k = 0$ for $i \neq k$, the mapping $G \to H = AG + GA^*$ is many-to-one and not all hermitian matrices are possible values for H. Since the mapping is linear from one real vector space to another, the set of image matrices H is a linear vector subspace. Such a subspace cannot contain all (nonsingular) hermitian matrices of any one fixed signature without being identical with the space of all hermitian matrices. This follows from the observation that any hermitian matrix can be written as a linear combination of two of a prescribed signature and this assertion, since it is independent of the choice of basis, follows from the equation

$$
\begin{aligned}
1_r \oplus (-1_s) &= [2\cdot 1_h \oplus (-1_k) \oplus (1_p) \oplus (-2\cdot 1_q)] \\
&\quad - [1_h \oplus (-2\cdot 1_k) \oplus (2\cdot 1_p) \oplus (-1_q)], \qquad (12.5)
\end{aligned}
$$

provided we have

$$h + k = r \quad \text{and} \quad p + q = n-r = s \quad . \qquad (12.6)$$

Each of the matrices in the square brackets has signature (h+p, k+q) and h+p can be made equal to an arbitrary integer c, $0 \leq c \leq n$, by taking $h = \min(r,c)$, $p = c-h$, $k = r-h$, $q = n-(h+p+k)$. Thus we have proved

Lemma 12.5. If H = AG + GA* can be made equal to an arbitrary hermitian matrix of any one signature, then the mapping $G \to H$ is one-to-one on hermitian matrices and the eigenvalues of A satisfy $\alpha_i + \bar{\alpha}_k \neq 0$.

An extension to a suitable class of singular H would be possible but awkward restrictions are required since the rank of a sum is never greater than the sum of the ranks.

Although not all positive definite hermitian matrices are attainable as images under $G \to H$ unless $\alpha_i + \bar{\alpha}_k \neq 0$, the next lemma gives the necessary and sufficient condition that some are.

Lemma 12.6. (Ostrowski) The matrix AG + GA* = H assumes definite hermitian values if and only if A has no pure imaginary eigenvalues.

The "only if" part was proved in Lemma 12.4. The "if" part is a special case of the next lemma.

Lemma 12.7. If A, of order n, has δ pure imaginary eigenvalues, there exists an hermitian matrix G of rank (at least) n-δ such that AG + GA* is positive semidefinite and of rank n-δ. If the elementary divisors associated with the pure imaginary eigenvalues are all simple, then G can be chosen nonsingular while AG + GA* remains positive semidefinite.

The proof follows if we take A in ε-Jordan block form with an ε in the nonzero off-diagonal positions and choose for G a diagonal matrix with +1 in those positions where A has an eigenvalue with positive real part, -1 where the corresponding eigenvalue has negative real part and zero where the eigenvalue is pure imaginary. Then AG + GA* is a direct sum of blocks of the form

$$\begin{Vmatrix} |\mathrm{Re}\ \alpha| & \eta & & & \\ \bar{\eta} & |\mathrm{Re}\ \alpha| & \eta & & \\ & \bar{\eta} & |\mathrm{Re}\ \alpha| & \eta & \\ & & \bar{\eta} & \ddots & \ddots \\ & & & \ddots & \ddots \end{Vmatrix} \tag{12.7}$$

where $\eta = \pm\,\varepsilon$, and (possibly) a zero block and is therefore positive semidefinite for a sufficiently small ε. If the Jordan blocks with pure imaginary eigenvalues are all of order one, the corresponding diagonal elements of G can be chosen as arbitrary real numbers without varying AG + GA* and so G may be chosen nonsingular.

The results of this section have usually depended on _a priori_ information about A and this is particularly true of Taussky's interesting generalization of a part of Lyapunov's result (cf. Lemma 12.3) where the assumption $\alpha_i + \bar{\alpha}_k \neq 0$ was essential.

For possible eventual use in computing eigenvalues, the crucial result of Lyapunov would appear to be that which gives information about the eigenvalues of a general matrix A from knowledge of the signatures of two hermitian matrices.

Lemma 12.8. (Lyapunov) If G and H = AG + GA* are both positive definite, then every eigenvalue of A has positive real part. One may, however, have G positive definite and A have only eigenvalues with positive real part and yet have AG + GA* indefinite and nonsingular.

Proof (using mild properties of the field of values of a matrix, studied more extensively in the next section): Let $uA = \alpha u$ for a nonzero u. Then $u(AG + GA^*)u^* = uHu^*$ gives $(\alpha + \bar{\alpha})(uGu^*) = uHu^*$ and both uGu^* and uHu^* are positive so $\alpha + \bar{\alpha} > 0$ as required. An example which proves the second part of the lemma is:

$$A = \begin{Vmatrix} a & 0 \\ c & b \end{Vmatrix}, \quad a > 0,\ b > 0; \quad G = \begin{Vmatrix} g & 0 \\ 0 & h \end{Vmatrix}, \quad g > 0,\ h > 0 ; \tag{12.8}$$

and $AG + GA^* = \begin{Vmatrix} 2ag & cg \\ cg & 2bh \end{Vmatrix}$ which

has for its determinant $4abg^2\left(\frac{h}{g} - \frac{c^2}{4ab}\right)$ and is therefore indefinite whenever $\frac{h}{g} < \frac{c^2}{4ab}$.

Ostrowski, in his Gatlinburg lecture, introduced the concept of the inertia of a matrix which is defined as:

Inertia of $A = (\pi, \nu, \delta)$, where π is the number of eigenvalues of A with positive real part, ν is the number with negative real part and δ is the number of pure imaginary eigenvalues (12.9)

We note that the inertia of an hermitian matrix is $(\pi, \nu, 0)$.

If we require of $G = G^*$ only that its image H is positive <u>semi</u>definite ($H = AG + GA^* \geqslant 0$), the following example shows that $\pi_A = \nu_A = 0$ permits π_G and ν_G each to be either zero or positive independently of one another. (More general results are presumably contained in [47]).

Let $A = A_1 \oplus (iI_p)$ and $G = G_1 \oplus D$, where A_1 has no pure imaginary eigenvalue, so $G_1 = G_1^*$ can be chosen by Lemma 12.7, to make $A_1G_1 + G_1A_1^* > 0$, and D is real and diagonal. Then $AG + GA^* \geqslant 0$ while the inertia of G depends on the signs of the diagonal elements of D. If we replace iI_p by a direct sum of Jordan blocks with pure imaginary eigenvalues and D by a corresponding direct sum of diagonal matrices of the form $\|\operatorname{diag}(d,0,0,\ldots,0)\|$, the inertia of G can again be varied much as in the initial case of simple eigenvalues.

We have therefore proved:

Lemma 12.9. If A has at least one pure imaginary eigenvalue, then $AG + GA^* \geqslant 0$ for hermitian G of differing signatures.

We now remove the hypothesis $\alpha_i + \alpha_k \neq 0$ in Lemma 12.3 and, using a technique suggested by a remark of Ostrowski in his Gatlinburg lecture, establish

Lemma 12.10. If $G = G^*$ is any hermitian matrix for which $AG + GA^*$ is positive definite, then G has the same number of positive (negative) eigenvalues as A has eigenvalues with positive (negative) real part.

Proof: Since by Lemma 12.3 the result is proved except when there are different hermitian matrices G_1 and G_2 such that

$$H \equiv AG_1 + G_1A^* = AG_2 + G_2A^* > 0 \quad , \tag{12.10}$$

the technique of allowing G to vary with A while holding H fixed is inappropriate. Instead we consider G fixed and observe that

$$(A + pI)G + G(A + pI)^* = H + 2pG \ , \tag{12.11}$$

if p is real. We can choose a positive p so that it does not satisfy any of the finite number of equations

$$(\alpha_i + p) + (\bar{\alpha}_k + p) = 0 \tag{12.12}$$

and also small enough so that

$$H + 2pG > 0 \quad . \tag{12.13}$$

For, $x(H + 2pG)\, x^* > 0$ for every unit vector x if

$$2p \; (\max \mid \text{eigenvalue of } G \mid) < \text{minimum eigenvalue of } H \tag{12.14}$$

Then Lemma 12.3 implies that the (unique) G satisfying (12.11) has the same number of positive eigenvalues as $A + pI$ has eigenvalues with positive real parts. If we also require of p that

$$0 < 2p < \min_i \mid \alpha_i + \bar{\alpha}_i \mid , \tag{12.15}$$

where we know $\alpha_i + \bar{\alpha}_i \neq 0$ by Lemma 12.4 (or 12.6), the sign of $\mathrm{Re}(\alpha_i + p)$ agrees with that of $\mathrm{Re}(\alpha_i)$ for $i = 1,\ldots,n$, completing the proof.

Application* of Lemma 12.10. Take for A the companion matrix (cf.(3.13)) of the polynomial $\lambda^4 + 4$:

$$A = \begin{Vmatrix} 0 & 0 & 0 & -4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{Vmatrix} \quad . \tag{12.16}$$

---

* Found in attempting to construct a counter example!

Then if we write an arbitrary real symmetric matrix in the form

$$G = \left\| \begin{matrix} a & p & t & -s \\ p & b & q & u \\ t & q & c & r \\ -s & u & r & d \end{matrix} \right\|, \tag{12.17}$$

we find that

$$AG+GA^T = \left\| \begin{matrix} 8s & a-4u & p-4r & t-4d \\ a-4u & 2p & b+t & -s+q \\ p-4r & b+t & 2q & u+c \\ t-4d & -s+q & u+c & 2r \end{matrix} \right\| . \tag{12.18}$$

Hence, $AG + GA^T$ will be diagonal if $p = 4r$, $q = s$, $a = -4c$, $u = -c$, $t = 4d$ and $b = -4d$. Since the roots of $\lambda^4 + 4 = 0$ are $\pm 1 \pm i$, the signature of the matrix

$$G = \left\| \begin{matrix} -4c & 4r & 4d & -s \\ 4r & -4d & s & -c \\ 4d & s & c & r \\ -s & -c & r & d \end{matrix} \right\| \tag{12.19}$$

will certainly be $(+ + - -)$ by Lemma 12.10 if $r > 0$ and $s > 0$. Calculating the determinant gives

$$\det G = (4r^2 - s^2)^2 + 4(4d^2 + c^2 + 2rs)^2 > 0 \tag{12.20}$$

and the value zero cannot be attained for $r > 0$ and $s > 0$. (But, $\det G = 0$ for, say, $r = 1$, $s = -2$, $d = 1$ and $c = 0$.)

Our goal of a computationally effective device for obtaining information about the (complex) eigenvalues of a general n by n matrix from information obtained from hermitian matrices is partly achieved in the next lemma although there remains the serious barrier of determining a G for which $AG + GA^*$ is positive definite in the absence of any special properties of A.

Lemma 12.11. Let the eigenvalues of A, G and H be α_i, γ_i and η_i, respectively, and assume that $G + G^*$ and that $H = AG + GA^*$ is positive definite. Then,

$$\min_i \mid \alpha_i + \bar{\alpha}_i \mid \geqslant \frac{\min_k \eta_k}{\max_j |\gamma_j|} . \tag{12.21}$$

Proof: For any one of the α_i, there is at least one unit vector u such that $uA = \alpha_i u$. Then, $uHu^* = u(AG + GA^*)u^* = (\alpha_i + \bar{\alpha}_i)(uGu^*)$. Also, $uHu^* \geqslant \min_k \eta_k > 0$ while $|uGu^*| \leqslant \max_j |\gamma_j|$. Hence,

$$\min_k \eta_k \lesssim uHu^* = |\alpha_i + \bar{\alpha}_i| \quad |uGu^*|$$

$$\lesseqgtr |\alpha_i + \bar{\alpha}_i| \max_j |\gamma_j| , \qquad (12.22)$$

and consequently

$$|\alpha_i + \bar{\alpha}_i| \geqslant \frac{\min_k \eta_k}{\max_j |\gamma_j|} , \qquad (12.23)$$

establishing (12.21).

The inequality of the lemma cannot be improved when A is diagonal since then we could take $G = \|\mathrm{diag}\ (\pm 1)\|$ with the signs chosen to make H positive definite and this would give $|\gamma_j| = 1$ and $|\alpha_i + \bar{\alpha}_i| = \eta_i$. Since we omitted the numerical value sign about the η_k in (12.21), the lemma holds trivially when H is not positive definite. Thus for given A, an arbitrary choice of G could be made and by already known methods (too long for inclusion here) an absolutely precise* lower bound could be calculated with the aid of a digital computer for the right member of (12.21). In the event that this lower bound were positive, the eigenvalues of A would be excluded from a vertical strip about the origin in the complex number plane. By considering $e^{i\theta}A - pI$ instead of A, strips centered on an arbitrary line could be obtained provided some additional information were available to permit a choice of G for which H would be positive definite.

The next section elucidates the difficulty of such a choice.

---

*Including all round-off effects.

13. Fields of Values

The field of values of A relative to a metric

$$\| u \|_G^2 = uGu^* \tag{13.1}$$

given by the positive definite hermitian matrix G is defined to be the set of complex numbers

$$F_G(A) = \left\{ \frac{(uA)Gu^*}{uGu^*} \quad ; \quad u \neq 0 \right\} \quad . \tag{13.2}$$

Writing, for $G = G^*$,

$$AG + GA^* = 2\,\mathrm{Re}(AG) \tag{13.3}$$

and

$$AG - GA^* = 2\,i\,\mathcal{I}m\,(AG) \tag{13.4}$$

so that

$$AG = \mathrm{Re}(AG) + i\,\mathcal{I}m\,(AG) \tag{13.5}$$

is the decomposition of AG into its hermitian and skewhermitian parts,

$$F_G(A) = \left\{ x + i\,y \right\} , \tag{13.6}$$

where

$$x = u(\mathrm{Re}(AG))u^* \tag{13.7}$$

and

$$y = u(\mathcal{I}m(AG))u^* \tag{13.8}$$

for

$$\| u \|_G = 1 \quad . \tag{13.9}$$

The salient facts about fields of values which we need here are contained in [22]* or follow immediately from the above equations, and we summarize them briefly:

1) $F_G(A)$ is a convex region;
2) the rectangle with sides

   x = minimum eigenvalue of Re(AG)
   x = maximum " " "
   y = minimum " " $\mathcal{I}m$(AG)
   y = maximum " " "

   circumscribes $F_G(A)$ and each edge actually contains at least one point of $F_G(A)$;

---

*We here use row vectors u instead of column vectors as in [22]. Had our discussion referred to $GA + A^*G$, column vectors would have been more natural.

3) the eigenvalues of A are in $F_G(A)$ and for each nonsimple elementary divisor the associated eigenvalue is an interior point;

4) there exists a metric G such that $F_G(A)$ is exactly the convex closure P(A) of the eigenvalues if and only if only eigenvalues with simple elementary divisors lie on the boundary of the polygonal region P(A);

5) the intersection of $F_G(A)$ for all $G > 0$ is P(A); and

6) $F_G(A) = F_1(T^{-1}AT)$ for $G = TT^*$.

It follows immediately from 5) that a positive definite G exists for which Re(AG) is negative definite if and only if A is stable and this is the Lyapunov result (cf. Lemmas 12.3 and 12.8).

The circumscribing rectangle of $F_G(A)$ given in 2) must contain P(A) and it will be impossible to take it to be as small as the rectangle of this orientation circumscribing P(A) if the rectangle has on its boundary an eigenvalue with nonsimple elementary divisor. For a suitable choice of G, $F_G(A)$ may, however, be made large except when A is scalar.

Lemma 13.1. If $A = \lambda 1$, $F_G(A) = \{\lambda\}$. For A nonscalar, a positive definite G may be chosen so that the rectangle $a \leqslant x \leqslant b$, $c \leqslant y \leqslant d$ circumscribing the field of values, $F_G(A)$ of A with respect to G has $a \leqslant a_1$, $b \geqslant a_2$, $c \leqslant a_3$ and $d \geqslant a_4$ for arbitrarily prescribed real numbers a_i.

Proof: By 6) it is sufficient to prove the desired result for $F_1(T^{-1}AT)$, where T may be chosen after the a_i have been assigned. Unless A is scalar, it is similar to a triangular matrix with an off-diagonal element which may be arbitrarily chosen; for, first (upper) triangularize A, then choose i and $j > i$ so that $a_{ii} \neq a_{jj}$ and effect the similarity which adds k times the i-th column to the j-th and subtracts k times the j-th row from the i-th to change the (i,j)-th element by $k(a_{ii}-a_{jj})$. Among the complex numbers in the field of values are those given by

$$\left\{a_{ii}|u|^2 + a_{jj}|v|^2 + Ku\bar{v},\ |u|^2 + |v|^2 = 1\right\},$$

where we may still choose K. Taking $u = e^{i\alpha}\cdot 2^{-1/2}$ and $v = e^{-i\alpha}\,2^{-1/2}$ gives as complex numbers in the field of values

$$\frac{1}{2}\,(a_{ii} + a_{jj} + e^{i2\alpha}K),\ \text{for } \alpha \text{ arbitrary} \tag{13.10}$$

and this is the circumference of an arbitrarily large circle.

The implication of this for the Lyapunov mapping is that, while a judicious choice of $G > 0$ may allow Re(AG) to give valuable information about the eigenvalues of A, a nonselective choice of G cannot be expected to give useful information.

14. Formal Solution of $AX + XA^* = Y$ in Integral Form

Generalizing a formula given by Bellman ([6], p. 243) we can obtain the following formal but explicit solution of the matrix equation

$$AX + XA^* = Y \tag{14.1}$$

in case A is stable. Let

$$X = -\int_0^\infty e^{At}\, Y\, e^{A^*t}\, dt \quad , \tag{14.2}$$

where the exponential function of a matrix is given by the usual power series and always converges. For a general matrix A the integral will not converge since for A diagonal and Y = 1 we have for the diagonal elements of X integrals of the form

$$-\int_0^\infty e^{(\alpha + \bar{\alpha})t}\, dt \ .$$

With the assumption that A is stable, so $\alpha + \bar{\alpha} < 0$, neither the presence of Y nor the matrix form, with the possibility that A has nonsimple elementary divisors, affects the convergence since a change of basis would allow A to be taken as a direct sum of α-Jordan blocks of the form

$$\alpha \left\| \begin{matrix} 1 & 1 & 0 & & & \\ 0 & 1 & 1 & & & \\ 0 & 0 & 1 & . & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & \cdot \end{matrix} \right\| \tag{14.3}$$

with $\alpha + \bar{\alpha} < 0$ and the argument could be completed by using the fact that $\int_0^\infty e^{(\alpha + \bar{\beta})t}\, dt$ exists when both $\alpha + \bar{\alpha} < 0$ and $\beta + \bar{\beta} < 0$.

Since we already know that the stability of A guarantees the uniqueness of X, it remains to show that the integral formula does in fact yield the solution. Calculating,

$$\begin{aligned} AX + XA^* &= -\int_0^\infty \left[A\, e^{At} Y\, e^{A^*t} + e^{At} Y\, e^{A+t} A^* \right] dt \\ &= -\int_0^\infty \frac{d}{dt} \left[e^{At} Y\, e^{A^*t} \right] dt \\ &= -\left[e^{At} Y\, e^{A^*t} \right]_0^\infty \quad = Y \quad . \end{aligned} \tag{14.4}$$

15. Formal Solution of $AX + XA^T = Y$ in Terms of Adjoints

In case A has no multiple eigenvalues α_i and $\alpha_i + \alpha_j \neq 0$ for every $i,j = 1,\ldots,n$, an explicit solution is given by Hahn* ([24], p. 23) with a reference to papers by Bedel'baev [5] and Malkin [40].

We define the adjoint of a matrix M to be the matrix with (i,j)-th element equal to the cofactor of the (j,i)-th element of M and thus obtain (cf. Wedderburn [65], pp. 6 and 65-66, or Bourbaki [8], § 6, No. 5, pp. 86-87).

$$M\,(\operatorname{adj} M) = (\operatorname{adj} M)\,M = (\det M)\,1. \tag{15.1}$$

Well known and easy consequences of Wedderburn's formulas are

$$\operatorname{transpose}(\operatorname{adj} M) = \operatorname{adj}(\operatorname{transpose} M)\,, \tag{15.2}$$

$$\operatorname{adj}(M_1M_2) = (\operatorname{adj} M_2)\,(\operatorname{adj} M_1)\,, \tag{15.3}$$

$$\operatorname{adj}(M^{-1}) = (\operatorname{adj} M)^{-1} \tag{15.4}$$

and

$$\operatorname{adj} \|\operatorname{diag}(d_1,d_2,\ldots,d_n)\| = \|\operatorname{diag}(e_1,e_2,\ldots,e_n)\| \tag{15.5}$$

where

$$e_i = d_1d_2\cdots d_{i-1}\,d_{i+1}\cdots d_n \tag{15.6}$$

Under our assumption that the eigenvalues α_i of A are distinct,

$$\operatorname{rank}(A - \alpha_i 1) = n-1 \tag{15.7}$$

and $\operatorname{adj}(A - \alpha_i 1) \neq 0$. Since

$$(A - \alpha_i 1)\,[\operatorname{adj}(A - \alpha_i 1)] = d(\alpha_i)1 = 0,$$

each column of $\operatorname{adj}(A - \alpha_i 1)$ is either zero or is a column eigenvector of A with eigenvalue α_i and similarly each row is zero or a row eigenvector. It follows that

$$\operatorname{rank}\operatorname{adj}(A - \alpha_i 1) = 1 \tag{15.8}$$

and

$$\operatorname{adj}(A - \alpha_i 1) = v_i\,u_i, \quad i=1,\ldots,n, \tag{15.9}$$

where the v_i are a linearly independent set of n column vectors, the u_i are a similar set of row vectors,

$$A\,v_i = \alpha\,v_i \quad \text{and} \quad u_iA = \alpha_i\,u_i\;. \tag{15.10}$$

---

*Hahn's reference to a paper by Vejvoda in the paragraph preceding his equation (8.10) is incorrect; the correct reference is [63]. This section generalizes the result stated by Hahn from Y = 1 (and not Y = 2·1 as Hahn states) to general Y.

Since $u_i(A\,v_j) = \alpha_j\,u_i\,v_j$ and $(u_iA)\,v_j = \alpha_i\,u_i\,v_j$, we have

$$u_i\,v_j = 0 \text{ for } i \neq j.$$

To calculate $u_i\,v_i$ we have recourse to the fact that A is similar to a diagonal matrix so that

$$T(A - \alpha_i\,1)\,T^{-1} = \|\,\mathrm{diag}\,(d_1, d_2, \ldots, d_n)\,\| \quad , \tag{15.11}$$

where

$$d_j = \alpha_j - \alpha_i \quad . \tag{15.12}$$

Using (15.3), (15.4) and (15.5),

$$(\mathrm{adj}\,T)^{-1}\,[\mathrm{adj}\,(A - \alpha_i\,1)]\,(\mathrm{adj}\,T) = \|\,\mathrm{diag}\,(0,0,\ldots,0,e_i,0,\ldots,0)\,\|, \tag{15.13}$$

where the i-th element on the diagonal is

$$\begin{aligned} e_i &= \prod_{j\neq i} (\alpha_j - \alpha_i) \\ &= - \left[\frac{d}{d\lambda} \prod_{j=1}^{n} (\alpha_j - \lambda) \right]_{\lambda = \alpha_i} \\ &= - d'(\alpha_i) \quad , \end{aligned} \tag{15.14}$$

and

$$d(\lambda) = \det\,(A - \lambda\,1) \quad . \tag{15.15}$$

Hence,

$$\mathrm{trace}\,[\mathrm{adj}\,(A - \alpha_i\,1)] = e_i = -d'(\alpha_i) \tag{15.16}$$

and

$$u_i\,v_i = -d'(\alpha_i) \quad . \tag{15.17}$$

It now follows that

$$\sum_{i=1}^{n} - \frac{\mathrm{adj}\,(A - \alpha_i\,1)}{d'(\alpha_i)} = 1 \tag{15.18}$$

since setting $x = \Sigma\,\xi_j v_j$ for an arbitrary column vector gives

$$\left[\sum - \frac{v_i\,u_i}{d'(\alpha_i)} \right] x = - \Sigma\,\xi_i \frac{v_i(u_i\,v_i)}{d'(\alpha_i)} = \Sigma\,\xi_i\,v_i = x \quad . \tag{15.19}$$

If we now set

$$B(\lambda) = [\text{adj } (A - \lambda 1)] \; Y \; [\text{adj } (A + \lambda 1)]^T \quad , \tag{15.20}$$

we find that

$$A \, B(\lambda) = \lambda \, B(\lambda) + d(\lambda) \; Y[\text{adj } (A + \lambda 1)]^T \tag{15.21}$$

and

$$B(\lambda) \, A^T = - \lambda \, B(\lambda) + d(-\lambda) \; [\text{adj } (A - \lambda 1)] \; Y \quad . \tag{15.22}$$

Adding the last equations,

$$\begin{aligned} A \, B(\lambda) + B(\lambda) \, A^T &= d(\lambda) \; Y[\text{adj } (A + \lambda 1)]^T \\ &\quad + d(-\lambda) \; [\text{adj } (A - \lambda 1)] \; Y \; . \end{aligned} \tag{15.23}$$

Remembering that we must have $\alpha_i + \alpha_j \neq 0$ if we are to have a solution for all Y, we impose this assumption in the form

$$d(- \alpha_i) \neq 0 \quad . \tag{15.24}$$

Hence, substituting $\lambda = \alpha_i$ in (15.23) and dividing by $d'(\alpha_i) \; d(- \alpha_i)$, we get on summing over $i=1,\ldots,n$,

$$AX + XA^T = \sum_{i=1}^{n} \left[\frac{\text{adj } (A - \alpha_i \; 1)}{- \; d'(\alpha_i)} \right] Y = Y \quad , \tag{15.25}$$

by (15.18), where

$$X = - \sum_{i=1}^{n} \frac{[\text{adj } (A - \alpha_i \; 1)] \; Y[\text{adj } (A + \alpha_i \; 1)]^T}{d'(\alpha_i) \; d(- \alpha_i)} \tag{15.26}$$

is therefore the explicit solution of $AX + XA^T = Y$.

BIBLIOGRAPHY

1. Afriat, S. N., Composite Matrices, Quart. J. Math (2) 5 (1954), 81-98.

2. Aitken, A. C., The Normal Form of Compound and Induced Matrices, Proc. London Math. Soc. II, 38 (1935), 354-376.

3. Arrow, K. J. and M. McManus, A Note on Dynamic Stability, Econometrica 26 (1958) 448-454.

4. Bass, Robert W., Zubov's Stability Criterion, Bol. Soc. Mat. Mexicana (2) 4 (1958), 26-29 (cf. M.R. 21 (1960) #7498).

5. Bedel'baev, A. K., On a Construction of the Function of Lyapunov as a Quadratic Form, Izv. Akad. Nauk Kazah. SSSR. Ser. Mat. Meh. 1956, No. 4(8), 24-37. (Russian) (cf. Math. Rev. 18 (1957), 308.)

6. Bellman, Richard, Introduction to Matrix Analysis, McGraw-Hill, N.Y. (1960).

7. Bickley, W. G., and J. McNamee, Matrix and Other Direct Methods for the Solution of Systems of Linear Difference Equations, Philos. Trans. Roy. Soc. London, Ser. A 252 (1960), 69-131. (cf. Math. Rev. 22 (1961) #4897.)

8. Bourbaki, N., Éléments de Mathématique, Livre II. Algèbre, ch. III. Algèbre Multilineaire, Actualites Sci. et Industr. #1044, Hermann, Paris (1948).

9. Bourbaki, N., Éléments de Mathématique, Algèbre, ch. 9, Formes Sesquilinéaires et Formes Quadratiques, Actualités Sci. et Industr. #1272, Hermann, Paris (1959).

10. Charaschal, V., On the Stability of Linear Systems of Linear Differential Equations of Second Order, Akad. Nauk. Kazah, SSSR. Trudy Sektor: Math. Meh. 1 (1958), 46-49 (Russian).

11. Chetayev, N. G., The Stability of Motion, Pergamon Press, London (1961).

12. Diliberto, Stephen P., A Note on Linear Ordinary Differential Equations, Proc. Amer. Math. Soc. 8 (1957), 462-464. (cf. Math. Rev. 19 (1958), 275.)

13. Egervary, E., On Hypermatrices Whose Blocks are Commutable in Pairs and Their Applications in Lattice Dynamics, Acta Scient. Math. Szeged 15 (1953), 211-222.

14. Enthoven, A. C., and K. J. Arrow, A Theorem on Expectations and the Stability of Equilibrium, Econometrica 24 (1956), 288-293.

15. Friedman, B., On n-Commutative Matrices, Math. Ann. 136 (1958), 343-7.

16. Friedman, B., Eigenvalues of Compound Matrices, N.Y.U. Math. Research Group Report, No. TW-16.

17. Friedman, B., Eigenvalues of Composite Matrices, Proc. Cambridge Philos. Soc. 57 (1961), 37-49.

18. Gantmacher, F. R., The Theory of Matrices, Vols. I, II, translated by K. A. Kirsch. Chelsea, New York (1959).

19. Gantmacher, Felix, Canonical Representation of Automorphisms of a Complex Semi-Simple Lie Group, Rec. Math. (Moscow) 5 (47), 101-146 (1939). (English, Russian summary) (cf. Math. Rev. 1 (1940), 163-164.)

20. Gantmacher, Felix, On the Classification of Real Simple Lie Groups, Rec. Math. [Mat. Sbornik] N.S. 5 (47), 217-250 (1939). (English, Russian summary) (cf. Math. Rev. 2 (1941), 5.)

21. Gautschi, W., Bounds of Matrices with Regard to a Hermitian Metric, Compositio Math. 12 (1954), 1-16.

22. Givens, Wallace, Fields of Values of a Matrix, Proc. Amer. Math. Soc. 3 (1952), 206-209.

23. Golubchikov, A. F., On the Structure of the Automorphisms of the Complex Simple Lie Groups, Dokl. Akad. Nauk, SSSR, 77 (1951), 7-9. (cf. Math. Rev. 12 (1951), 673.)

24. Hahn, Wolfgang, Theorie und Anwendung der direkten Methode von Lyapunov, Ergebnisse der Mathematik und ihrer Grenzgebiete. N.F., Heft 22. Springer-Verlag, Berlin-Göttingen-Heidelberg, 1959.

25. Hahn, Wolfgang, Ein Bemerkung zur zweiten Methode von Ljapunov. Math. Nachrichten 14 (1955), 349-354.

26. Hahn, Wolfgang, Bemerkungen zur einer Arbeit von Herrn Vejvoda, Math. Nachr. 20 (1959), 21-24. (cf. Math. Rev. 22 (1961) #128.)

27. Haynsworth, Emilie V., Special Types of Partitioned Matrices, J. Res. Nat. Bureau of Standards, Math. and Math. Physics 65B (1961), 7-12.

28. Hausner, M., Eigenvalues of Certain Operators on Matrices, Communications on Pure and Appl. Math. 14 (1961), 155-156.

29. Hestenes, M. R., and E. J. McShane, A Theorem on Quadratic Forms and Its Application in the Calculus of Variations, Trans. Amer. Math. Soc. 47 (1940), 501-512.

30. Ingraham, M. H. and M. C. Wolf, Relative Linear Sets and Similarity of Matrices Whose Elements Belong to a Division Algebra, Trans. Am. Math. Soc. 36 (1934), 16-31.

31. Ingraham, M. H. and H. C. Trimble, On the Matric Equation TA = BT + C, Am. J. Math. 63 (1941), 9-28.

32. Jacobson, N., Lectures in Abstract Algebra, Vol. II. Linear Algebra, van Nostrand (1953), New York.

33. Jacobson, N., Pseudo-linear Transformations, Annals of Math. 38 (1937), 484-507.

34. Khan, N. A., The Characteristic Roots of a Product of Matrices, Quart. J. Math. 7 (1956), 138-143.

35. Lewis, Daniel C., Differential Equations Referred to a Variable Metric, Am. J. Math. 73 (1951), 48-58.

36. Lewis, Daniel C., and Olga Taussky, Some Remarks Concerning the Real and Imaginary Parts of the Characteristic Roots of a Finite Matrix, Jour. of Mathematical Physics 1 (1960), 234-236.

37. Littlewood, D. E., On Induced and Compound Matrices, Proc. London Math. Soc. (2) 40 (1936), 370-381.

38. Lyapunov, A., Problème Général de la Stabilité du Mouvement, Ann. Math. Studies, No. 17, Princeton (1947). (Reprint of the translation of his 1893 work.)

39. MacDuffee, C. C., The Theory of Matrices, Ergebnisse der Math. u.i. Grenzgebiete Bd. 2, Berlin (1933).

40. Malkin, I. G., On the Construction of Lyapunov Functions for Systems of Linear Equations, Akad. Nauk SSSR. Prikl. Mat. Meh. 16 (1952), 239-242. (Russian) (cf. Math. Rev. 13 (1952), 944.)

41. Malkin, Joel G., Theorie der Stabilität einer Bewegung. (German translation of the (1952) Russian edition) R. Oldenbourg, Munich (1959). (cf. Math. Rev. 21 (1960) #2791 and 15 (1954), 873.)

42. Mirsky, L., The Spread of a Matrix, Mathematika 3 (1956), 127-130.

43. Muir, Thomas, The Theory of Determinants, Vols. I, II, III, and IV, Dover (reprint), N.Y. (1960).

44. Muir, Thomas, and Wm. H. Metzler, A Treatise on the Theory of Determinants, Dover (reprint) N.Y. (1960).

45. Ostrowski, A. M., Über Eigenwerte von Produkten Hermitenscher Matrizen, Hamburger Math. Abh. 23 (1959), 60-68.

46. Ostrowski, A. M., A Quantitative Formulation of Sylvester's Law of Inertia, Proc. Nat. Acad. Sci. U.S.A., 45 (1959), 740-744 (cf. Math. Rev. 22 #1588).

47. Ostrowski, A. M., and Hans Schneider, Paper on Stable Matrices presented at the Gatlinburg Conference in April, 1960, and to appear in Jour. Math. Analysis and Appl.

48. Reid, William T., Remarks on a Matrix Transformation for Linear Differential Equations, Proc. Amer. Math. Soc. 8 (1957), 708-712.

49. Rosenblum, Marvin, On the Operator Equation BX - XA = Q, Duke Math. J. 23 (1956), 263-269 (cf. Math. Rev. 18 (1957), 54).

50. Roth, W. E., On the Equation P(A,X) = 0 in Matrices, Trans. Am. Math. Soc. 35 (1933), 689-708.

51. Roth, W. E., On Direct Product Matrices, Bull. Amer. Math. Soc. 40 (1934), 461-468.

52. Rutherford, D. E., On the Condition that Two Zehfuss Matrices be Equal, Bull. Amer. Math. Soc. 39 (1933), 801-808.

53. Rutherford, D. E., On the Solution of the Matrix Equation AX + XB = C, Akad. van Wetenschappen, Amsterdam, Proceedings, 35 (1932), 54-59.

54. On the Rational Commutant of a Square Matrix, Akad. Wet., Amsterdam, Proc. 35 (1932), 870-875.

55. Rutherford, D. E., On the Rational Solution of the Matrix Equation SX = ST, Akad. Wet., Amsterdam, Proc. 36 (1933), 432-442.

56. Rutherford, D. E., Some Continuant Determinants Arising in Physics and Chemistry, Proc. Royal Soc. Edinburgh 62 (1947), 229-236; and, 63 (1952), 232-241.

57. Stephanos, Cyparissos, Sur une extension du calcul des substitutions lineaires, J. de Math. (5th ser.) 6 (1900), 73-128.

58. Taussky, Olga, Stable Matrices, Abstract 574-23 in Notices of AMS, vol. 7, #6, Issue #49, Nov. 1960, p. 743.

59. Taussky, Olga, Stable Matrices. II., Abstract 61T-158 in Notices of AMS, 8 (1960), 743.

60. Taussky, Olga, A Remark on a Theorem of Lyapunov, Jour. of Math. Analysis and Appl. 2 (1961), 105-107.

61. Taussky, Olga, A Generalization of a Theorem of Lyapunov (unpublished [preprint read]; part of a lecture given at the Matrix Conference in Gatlinburg, Tennessee, April 1961).

62. Turnbull, H. W., On the Reduction of Singular Matrix Pencils, Proc. Edinburgh Math. Soc. 4 (1935), 67-76.

63. Vejvoda, O., The Stability of a System of Differential Equations in the Complex Domain, Časopis Pěst. Mat. (82) (1957), 137-159. (Czech with Russian and English summaries.)

64. Walker, Gordon L., Direct Product and Lorentz Matrices, Cornell Univ. Dissertation (1941).

65. Wedderburn, J. H. M., Lectures on Matrices, Amer. Math. Soc. Colloq. Publ., Vol. XVII (1934).

66. Weitzenböck, R., Über die Matrixgleichung AX + XB = C, Proc. Amsterdam 35 (1932), 60-61.

67. Williamson, J., The Latent Roots of a Matrix of Special Type, Bull. Amer. Math. Soc. 37 (1931), 585-590.

68. Williamson, J., Quasi-Unitary Matrices, Duke Math. Jour. 3 (1937), 715-725.

69. Williamson, J., On the Normal Forms of Linear Canonical Transformations in Dynamics, Amer. Jour. of Math. 59 (1937), 599-617.

70. Zeipel, V. v., Om Determinanter, hvars elementer äro Binomialkoefficienter, Lunds Univ. Arsskrift, ii. (1865), 1-68.

71. Zubov, V. I., The Methods of A. M. Lyapunov and Their Application, Izdat. Leningrad Univ., Moscow (1957), 241 pp. (cf. Math. Rev. 19 (1958), 275.)

NOTE ADDED 10/26/61. After this paper was submitted for typing (on 9/15/61), we learned of a written account of the Gatlinburg lecture by Ostrowski to which reference is made above in § 12 and in [47]. This paper (first seen on 10/13/61) establishes the equality of the inertias of A and H when $\mathcal{R}$(AH) is positive definite, draws a number of interesting conclusions from this main theorem, studies the case in which A has pure imaginery eigenvalues and $\mathcal{R}$(A) is semidefinite, finds conditions for H-semistability (cf. [3] and [14]) and concludes with a result, stated without proof, connecting the inertias of A, H and $\mathcal{R}$(AH). The reference is:

72. Ostrowski, A. and H. Schneider, Some Theorems on the Inertia of General Matrices, Math. Research Center Tech. Summary Report #227 (April,1961), Univ. of Wisconsin, Madison, Wisc.

Item [4] in the bibliography was included on the basis of a reference in Math. Revs. and personal knowledge of relevant work by its author, but on receipt of a reprint, it became evident that reference to unpublished work of Bass would have been more pertinent.

The long paper [7] was also not seen until very recently. Whether or not the proposed "small" matrix techniques can replace the current "large" matrix methods used in the digital computer solution of partial differential equations requires further study. This seems most worthwhile in view of the large economies in computer time which could conceivably result.

X