2004

Positivity of eigenvalues of words in two positive definite letters

Ruchi Pratap
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Pratap, Ruchi, "Positivity of eigenvalues of words in two positive definite letters" (2004). *Master's Theses*. 2605.

DOI: <https://doi.org/10.31979/etd.4r5j-3utp>

https://scholarworks.sjsu.edu/etd_theses/2605

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

POSITIVITY OF EIGENVALUES OF WORDS
IN TWO POSITIVE DEFINITE LETTERS

A Thesis

Presented to

The Faculty of the Department of Mathematics

San Jose State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Ruchi Pratap

August 2004

UMI Number: 1424476

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI Microform 1424476

Copyright 2005 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© 2004

Ruchi Pratap
ALL RIGHTS RESERVED

APPROVED FOR THE DEPARTMENT OF MATHEMATICS

1-29-04

Dr. Jane Day, Thesis Advisor

Dr. Wasin So

Dr. Bradley Jackson

APPROVED FOR THE UNIVERSITY

Abstract

POSITIVITY OF EIGENVALUES OF WORDS IN TWO POSITIVE DEFINITE LETTERS

by Ruchi Pratap

A generalized word in two letters A and B can be written as

$$W(A, B) = A^{\alpha_1} B^{\beta_1} \dots A^{\alpha_N} B^{\beta_N},$$

where α_i and β_i are nonzero reals. The symbols A and B represent positive definite real matrices. The questions of interest are:

1. Is the trace of a given word positive for all such A and B?
2. Which words in positive definite letters have all eigenvalues positive, or at least one positive for such A and B?

This thesis is based on papers by Christopher J. Hillar, Charles R. Johnson, and Ilya M. Spitkovsky, which study generalized words in positive definite letters. Theorems in these papers present conditions on the exponents for which the answers to the above questions are affirmative. Large classes of words that do have one or the other of these properties are identified, and specific examples make it clear that other types of words do not.

Acknowledgements

The author gratefully acknowledges the enthusiastic support, encouragement, and supervision of her thesis advisor, Dr. Jane Day, throughout the thesis-writing period.

The author also wishes to thank her committee members, Dr. Wasin So and Dr. Bradley Jackson, for their valuable time and helpful comments and suggestions.

Table of Contents

Chapter 1	Overview	
1.1	Introduction	1
1.2	Definitions	2
1.3	Transformations that yield equivalent words	3
Chapter 2	Words with Positive Eigenvalues	
2.1	Introduction	7
2.2	Words with special exponent patterns	8
2.3	Nearly symmetric words	23
Chapter 3	Words That Do Not Always have Positive Eigenvalues	
3.1	Words with nonpositive eigenvalues	29
References		40
Appendices		
A	Positive definite and nonnegative matrices	41
B	Rank of a matrix	44
C	Theorems on congruence	46
D	Orthoprojections	49

CHAPTER 1 OVERVIEW

1.1 Introduction

The Bessis, Moussa and Villani (BMV) Conjecture, which is about quantum mechanical systems, was formulated in 1975 [BMV]. It has been proved in special cases but remains open in general. The conjecture says that, when A and B are Hermitian and B is positive definite, $\text{Trace}(\exp((A - \lambda B)))$ is the Laplace transform of a positive measure supported in $[0, \infty)$. This has been shown to be equivalent to the statement that when A and B are positive definite and m is any positive integer, the polynomial $f(\lambda) = \text{Trace}[(A + \lambda B)^m]$ has all coefficients nonnegative [LS].

The BMV conjecture has been proved for special types of matrices and when A and B are 2×2 , but it is not known whether or not it is true for any larger size matrices.

The papers [JH] and [JHS], which are the subject of this thesis, study the positivity of the coefficients in $f(\lambda)$. Each such coefficient is the trace of a sum of matrices of the form $A^{\alpha_1} B^{\beta_1} \dots A^{\alpha_N} B^{\beta_N}$, where the exponents are positive integers. One of the results in [JH], presented as (3.1.2) below, shows a way to construct 3×3 matrices A and B so that $\text{Trace}(A^2 B^2 A B)$ is negative. However, this is only one summand in a coefficient of $\text{Trace}[(A + \lambda B)^6]$, hence that coefficient could still be nonnegative.

In [JH] and [JHS], an expression of the form $A^{\alpha_1} B^{\beta_1} \dots A^{\alpha_N} B^{\beta_N}$, where the exponents are nonzero real numbers, is called a word in A, B . Some of the results proved in these papers include: any word in two 2×2 positive definite matrices with all exponents real

numbers of the same sign has all eigenvalues positive (Theorem 2.1.3), hence has positive trace; any symmetric word in positive definite A, B has positive eigenvalues, hence positive trace; and for any positive integers $p, q \geq 2$, there exist 3×3 positive matrices A and B such that the word ABA^pB^q has a negative trace (Theorem 3.1.3).

The appendices containing the matrix theory background that we need here are included at the end.

1.2 Definitions

1.2.1 **Word** [JHS]. A *generalized word*, or just *word*, in letters $A_1, A_2, \dots, A_n$ is an expression of the form $W(A_1, A_2, \dots, A_n) = A_1^{\alpha_1} \dots A_n^{\alpha_n}$ where the exponents are nonzero real numbers. A word in two letters A and B of the form

$$W(A, B) = A^{\alpha_1} B^{\beta_1} \dots A^{\alpha_N} B^{\beta_N}, \quad \alpha_i, \beta_i \in \mathbb{R} - \{0\} \quad (1)$$

is said to be of *class* N .

1.2.2 **A Symmetric Word** [JH]. A word is *symmetric* if it reads the same left to right as right to left. For example, $ABBABBA$ and $A^{1/2} B^3 A^{1/2}$ are symmetric but $ABABBA$ is not.

1.2.3 **Length of a Word** [JH]. When all exponents in a word are positive integers, the *length* of the word is the total number of letters present. For example, $AABABBAAB$ has length 10 and $AAAA$ has length 4.

We will write R^n and C^n to denote the spaces of real n -tuples and complex n -tuples, respectively.

1.3 Transformations that Yield Equivalent Words

1.3.1 Equivalent Words. Throughout this paper we assume that A and B are $n \times n$ real positive definite matrices, unless explicitly stated otherwise. The definition of positive definite and its basic properties are in Appendix A. We are interested in finding words $W(A,B)$ which have positive eigenvalues no matter what positive definite A and B are used. Thus we will call two words *equivalent* if one can be changed into the other by the following three types of transformations [JH, p 917].

(i) *Reversal*: this means writing the letters of a word in reverse order. The eigenvalues of $W(A,B)$ are not changed by reversal because the reverse of a word $W(A,B)$ is its transpose, since A and B are real and symmetric, and any matrix and its transpose have the same eigenvalues.

(ii) *Cyclic permutation*: this means moving the first letter of a word to the end of the word. This is a similarity, hence the eigenvalues do not change: for if $W = AW_1$ then $W_1A = A^{-1}WA$.

(iii) *Interchange of A and B*: this means creating $W(B,A)$ instead of $W(A,B)$. There is no need for us to study both $W(A,B)$ and $W(B,A)$ since we are interested here in what can be said about the positivity of eigenvalues of a word no matter what positive definite A and B are substituted for the letters. For example, the words $A^{\alpha_1}B^{\beta_1}$ and $B^{\alpha_1}A^{\beta_1}$ are equivalent for us.

We will write $W_1 \sim W_2$ to denote that words W_1 and W_2 are equivalent via one or more of the above transformations.

1.3.2 Example [JH]. There are only three non-equivalent words in two positive definite letters A and B , of length 6 with positive integer exponents and exactly 3 A 's.

Proof: The number of such words is ${}^6C_3 = 20$ (out of 6 places, choose 3 for A 's). We need only to calculate all the words of length 6 starting with A , since interchange of A and B would yield equivalent words:

Now we have ten distinct words:

- (i) A^3B^3
- (ii) ABA^2B^2
- (iii) AB^3A^2
- (iv) A^2BAB^2
- (v) $ABABAB$
- (vi) $ABAB^2A$
- (vii) AB^2A^2B
- (viii) AB^2ABA
- (ix) A^2B^2AB
- (x) A^2B^3A

We now verify by cases that only three of the ten words above are non-equivalent, those represented by (i), (ii) and (v).

(iii) $\sim$ (i): $AB^3A^2 \sim A^3B^3$ (by 4 cyclic permutations)

(iv) $\sim$ (ii): $A^2BAB^2 \sim BAB^2A^2$ (by 2 cyclic permutations) $\sim ABA^2B^2$ (by interchange of A and B)

(vi) $\sim$ (ii): $ABAB^2A \sim BAB^2A^2$ (by a cyclic permutation) $\sim ABA^2B^2$ (by interchange of A and B)

(vii) $\sim$ (ii): $AB^2A^2B \sim BAB^2A^2$ (by 5 cyclic permutation) $\sim ABA^2B^2$ (by interchange of A and B)

(viii) $\sim$ (ii): $AB^2ABA \sim ABA^2B^2$ (by 3 cyclic permutations)

(ix) $\sim$ (ii): $A^2B^2AB \sim ABA^2B^2$ (by 4 cyclic permutations)

(x) $\sim$ (i): $A^2B^3A \sim A^3B^3$ (by 5 cyclic permutations).

Finally, the words (i) A^3B^3 , (ii) ABA^2B^2 , and (v) $ABABAB$ are clearly non-equivalent.

CHAPTER 2 WORDS WITH POSITIVE EIGENVALUES

2.1 Introduction

We are interested in the following three questions [JH, Introduction].

Question 1. Does the polynomial $p(t)$, defined by $p(t) = \text{Tr}[(A+Bt)^m]$ have all coefficients positive for positive definite A and B ?

Question 2. Is the trace of a given word positive for positive definite A and B ?

Question 3. Are all the eigenvalues of a given word positive for positive definite A and B ?

In Sections 2.2 and 2.3 we will discuss a variety of words which the authors show do give affirmative answers to questions 2 and 3, and a few types of matrices A and B for which the answer to Question 1 is affirmative.

In Section 3 below we will explain some examples and theorems that the authors found, which show that the answers to Questions 2 and 3 are negative in general. However, Question 1 remains unanswered: no positive definite A and B and positive integer m are known for which any coefficient of $\text{Trace}[(A+Bt)^m]$ is negative, but there is no proof they must be nonnegative except when the matrices are 2×2 .

2.2 Words with Special Exponent Patterns

The definition of matrices being congruent, and the basic properties of congruence, are given in Appendix C.

2.2.1 Theorem [HJ, p. 917]. A symmetric word W in any finite number of real, symmetric, invertible letters is congruent to either the center letter (if its length is odd) or to the identity matrix I (if its length is even).

Proof: We know W and all of its factors are invertible. If the length of W is odd, then without loss of generality $W = W_1 A W_1^T$ and thus W is congruent to A . If the length of W is even, then W equals $W_1 W_1^T$ for some word W_1 , which equals $W_1 I W_1^T$, hence W is congruent to I . Thus by Sylvester's Theorem (4.2.4) $W(A, B)$ has all eigenvalues positive.

2.2.2 Theorem [JH, Lemma 1]. A symmetric word W in positive definite real letters is positive definite and thus has positive eigenvalues. That is, any symmetric word gives an affirmative answer to Question 3 and hence to Question 2 also.

Proof: Our letters are real and positive definite, hence symmetric. By (2.2.1), W is congruent to either its center letter or to I . In either case, W is congruent to a positive definite matrix, so its eigenvalues are positive by Sylvester's Theorem (4.2.4) in Appendix C. Since W is also symmetric and real, W is positive definite by (4.0.3).

2.2.3 Theorem [JH, Fact 6]. Let A and B be 2×2 positive definite real or complex matrices. Then any word $W(A, B)$ with all exponents positive or all negative has positive eigenvalues.

Proof: Suppose first that all exponents are positive. Let $W(A, B) = A^{d_1} B^{d_2} \dots A^{d_{k-1}} B^{d_k}$ where $d_1, d_2, \dots, d_k$ are positive real numbers. Since A is positive definite there exists a unitary matrix U and a diagonal matrix A_1 such that $A_1 = U^* A U$. Let $B_1 = U^* B U$. Note that B_1 is unitarily similar to B , which is positive definite, hence B_1 is also positive definite. Also the original word W is similar to $U^* A^{d_1} B^{d_2} \dots A^{d_{k-1}} B^{d_k} U$ which equals $(U^* A^{d_1} U) (U^* B^{d_2} U) (U^* \dots U) (U^* A^{d_{k-1}} U) (U^* B^{d_k} U)$

$$= A_1^{d_1} B_1^{d_2} \dots A_1^{d_{k-1}} B_1^{d_k}. \quad (I)$$

We will show that there exists a 2×2 unitary diagonal matrix D such that each entry of

$D^* B_1 D$ is real and positive (see Appendix A). We know $B_1 = \begin{bmatrix} a & z \\ \bar{z} & b \end{bmatrix}$ where $a, b \in \mathbb{R}$,

$a, b > 0$ and $z \bar{z} \geq 0$. If $z = 0$ we are done since W is then similar to a product of diagonal matrices with positive diagonal entries. Thus W has positive eigenvalues.

Suppose now that $z \neq 0$, and define $D = \begin{bmatrix} 1 & 0 \\ 0 & \bar{z}/|z| \end{bmatrix}$ so D is a 2×2 , diagonal, unitary

matrix. Then

$$B_2 = D^* B_1 D = \begin{bmatrix} 1 & 0 \\ 0 & \bar{z}/|z| \end{bmatrix} \begin{bmatrix} a & z \\ \bar{z} & b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \bar{z}/|z| \end{bmatrix} = \begin{bmatrix} a & |z| \\ |z| & b \end{bmatrix} > 0.$$

Therefore (I) is similar to $D^* A_1^{d_1} B_1^{d_2} \dots A_1^{d_{k-1}} B_1^{d_k} D$ which equals

$$(D^* A_1^{d_1} D) (D^* B_1^{d_2} D) (D^* \dots D) (D^* B_1^{d_k} D). \quad (II)$$

Note that $D^* A_1 D = A_1$ because D and A_1 are diagonal and $D^* D = I$. So (II) becomes $A_1^{d_1} B_1^{d_2} \dots A_1^{d_{k-1}} B_1^{d_k}$, which is a positive, invertible matrix. Then by the Perron-Frobenius Theorem (4.0.8) in Appendix A it has a positive eigenvalue.

Now $\det(W) = \det(A_1)^{d_1} \dots (\det B_1)^{d_k} > 0$ and this is the product of the two eigenvalues of W ; we know one of the eigenvalues is positive, so the other one has to be positive also. Thus both eigenvalues of the word $W(A_1, B_1)$ are positive, and $W(A, B)$ is similar to this.

If instead all exponents in $W(A, B)$ are negative, note $W(A, B)^{-1}$ is a word in positive definite A^{-1} and B^{-1} which has all exponents positive, hence its eigenvalues are positive, and thus those of $W(A, B)$ are.

2.2.4 Corollary [JH, p.919]. The polynomial $p(t) = \text{Tr}[(A + Bt)^m]$ has all positive coefficients for 2×2 positive definite A and B , for all positive integers m . Thus the answer to Questions 1, 2 and 3 are all affirmative in this special case.

Proof: On expanding $(A + Bt)^m$, each coefficient of a power of t is an expression which is the sum of different words in 2×2 positive definite A and B . From (2.2.3), each such word has positive eigenvalues and hence positive trace. Since the trace of the sum of matrices is equal to sum of their individual traces, the coefficient of each power of t in the trace of $(A + Bt)^m$ is positive. That is, the polynomial $p(t)$ has positive coefficients.

One can see that the technique of the proof of (2.2.3) does not work for the 3x3 case.

For instance let $B_1 = \begin{bmatrix} 3 & x & y \\ \overline{x} & 2 & z \\ \overline{y} & \overline{z} & 2 \end{bmatrix}$ with x, y and z different nonreal numbers of modulus

one. The leading principal minors of B_1 are all positive, so it is positive definite [HJ, p.

397]; but no matter what diagonal unitary $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & u & 0 \\ 0 & 0 & v \end{bmatrix}$ is chosen with u and v nonreal

and of modulus one, $D^*B_1D = \begin{bmatrix} 3 & ux & vy \\ \overline{ux} & 2 & \overline{vzu} \\ \overline{vy} & u(\overline{vz}) & 2 \end{bmatrix}$ has nonreal entries, hence cannot be

positive. Indeed, the following example shows that Theorem (2.2.3) is not true for 3x3 matrices.

2.2.5 Example [JH, p. 919]. Here is an example to show that Theorem (2.2.3) is not true for every word in two 3x3 positive definite matrices. Thus this example shows that the answers to Questions 2 and 3 at the beginning of this section are in general “no.”

$$\text{Let } A = \begin{bmatrix} 1 & 20 & 210 \\ 20 & 402 & 4240 \\ 210 & 4240 & 44903 \end{bmatrix} \text{ and } B = \begin{bmatrix} 36501 & -3820 & 190 \\ -3820 & 401 & -20 \\ 190 & -20 & 1 \end{bmatrix}.$$

Using Matlab, we get these approximate spectra for A and B :

$\sigma(A) = \{8.086e-005, 1.638, 45304\}$ and $\sigma(B) = \{2.227e-005, 1.2168, 36902\}$, so A and B are positive definite (refer (4.0.4) of Appendix A).

Let $W(A, B) = BABA^2B$. Let $C = BA$ so $W = C^2C^T$. Then

$$C = \begin{bmatrix} 1 & -20 & -20 \\ 0 & 2 & -20 \\ 0 & 0 & 3 \end{bmatrix}, W = \begin{bmatrix} -5199 & -6520 & 960 \\ 1920 & 2008 & -300 \\ -180 & -180 & 27 \end{bmatrix} \text{ and } \text{Tr } W = -3164. \text{ Since } W \text{ has a}$$

negative trace, at least one eigenvalue of W is negative.

2.2.6 Lemma [JHS, Lemma 2]. Let A and B be $n \times n$ real and positive definite. Suppose one of them has an eigenvalue of multiplicity at least $n - 1$, and in the word

$W(A, B) = A^{\alpha_1} B^{\beta_1} \dots A^{\alpha_N} B^{\beta_N}$ ($\alpha_j, \beta_j \in \mathbb{R} - \{0\}$), suppose all the exponents on the other matrix are of the same sign. Then $\rho(W(A, B))$ is an eigenvalue of $W(A, B)$, so $W(A, B)$ has at least one positive eigenvalue.

Proof: Let A be the matrix with eigenvalue λ_1 of multiplicity $n - 1$ and let λ_2 be its other eigenvalue. Then the word $W(A, B) = A^{\alpha_1} B^{\beta_1} \dots A^{\alpha_N} B^{\beta_N}$ has all β_i 's of the same sign.

We may suppose the β_i 's are positive. (If all $\beta_i < 0$, $C = B^{-1}$ is positive definite by (4.0.2)

so we can study $W(A, C) = A^{\alpha_1} C^{\gamma_1} \dots A^{\alpha_N} C^{\gamma_N}$ where all $\gamma_i > 0$). Consider the

following cases:

Case 1: Suppose $\beta_1, \dots, \beta_N$ are positive integers. Let U be a real orthogonal matrix such

that $A_0 = U^T A U = \begin{bmatrix} \lambda_1 I_{n-1} & 0 \\ 0 & \lambda_2 \end{bmatrix}$ (where λ_1 could equal λ_2). Let $B_1 = U^T B U =$

$\begin{bmatrix} B_{11} & b \\ b^T & \mu_n \end{bmatrix}$. Then B_1 is positive definite since B is, so B_{11} is positive definite as it is a

principal submatrix of positive definite B_1 . Let V be real orthogonal so that $V^T B_{11} V = D$

is diagonal. Let

$$\begin{aligned}
B_0 &= \begin{bmatrix} V & 0 \\ 0 & 1 \end{bmatrix}^T B_1 \begin{bmatrix} V & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} V^T & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} B_{11} & b \\ b^T & \mu_n \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} V^T B_{11} V & V^T b \\ b^T V & \mu_n \end{bmatrix} = \begin{bmatrix} D & V^T b \\ b^T V & \mu_n \end{bmatrix} \\
&= \begin{bmatrix} \mu_1 & 0 & \dots & 0 & \gamma_1 \\ 0 & \mu_2 & \dots & 0 & \gamma_2 \\ \vdots & & \dots & & \vdots \\ 0 & \dots & 0 & \mu_{n-1} & \gamma_{n-1} \\ \gamma_1 & \gamma_2 & & \gamma_{n-1} & \mu_n \end{bmatrix}.
\end{aligned}$$

If $V^T b$ has any negative entries, it is possible to rechoose V so that all entries of $V^T b$ will be nonnegative. To do this, define a real diagonal S by putting -1 as its j^{th} diagonal entry if the j^{th} entry of $V^T b$ is negative, and $+1$ as its j^{th} diagonal entry if the j^{th} entry of $V^T b$ is nonnegative. Note VS is orthogonal since S and V are, so we can use VS instead of V in the construction of B_0 above, and $(VS)^T b$ will have nonnegative entries.

Observe that $\mu_1, \dots, \mu_{n-1}$ are positive since they are eigenvalues of B_{11} . Also μ_n is positive since it is a diagonal entry of the positive definite B_1 . Thus B_0 is entrywise nonnegative and so is A_0 , hence $W(A_0, B_0)$ is nonnegative entrywise and invertible. Therefore $W(A_0, B_0)$ is permutation similar to a block diagonal matrix W' with nonnegative, irreducible diagonal blocks, hence $\rho(W') > 0$ and $\rho(W')$ is an eigenvalue of W' by the Perron-Frobenius Theorem (4.0.8). Since $W(A, B)$ is orthogonally similar to $W(A_0, B_0)$, hence to W' , Case 1 is done.

Case 2: Suppose $\beta_1 \dots \beta_N$ are rationals. Then $\beta_i = p_i/q_i$ where p_i and q_i are positive integers, so $B^{p_i/q_i} = (B^{1/q_i})^{p_i}$. Let Q be the least common denominator of the q_i 's and let

$B_1 = B^{1/Q}$; then B_1 is positive definite by (4.0.6) and $W(A, B) = W(A, B_1^Q)$ so we have reduced Case 2 to Case 1.

Case 3: Suppose $\beta_1 \dots \beta_N$ are arbitrary positive real numbers. Then for each $j = 1, \dots, N$ we introduce a sequence $\beta_j^{(k)}$ of positive rational numbers such that $\lim_{k \rightarrow \infty} \beta_j^{(k)} = \beta_j$. Let $W_k(A, B) = A^{\alpha_1} B^{\beta_1^{(k)}} \dots A^{\alpha_N} B^{\beta_N^{(k)}}$. Then the spectral radius of each of the matrices $W_k(A, B)$ is an eigenvalue of $W_k(A, B)$ by Case 2, so the limit matrix $W(A, B)$ has its spectral radius as an eigenvalue (see [HJ], Appendix D). Since $W(A, B)$ is invertible, it cannot have zero as an eigenvalue, hence $\rho(W(A, B))$ is a positive eigenvalue of $W(A, B)$.

2.2.7 Theorem [JHS, Theorem 1]. Suppose A and B are positive definite matrices, real or complex, such that A has exactly two distinct eigenvalues, $\lambda_1 > \lambda_2 > 0$. Then the spectrum of $A^m B A^{-m} B^{-1}$ is positive for all positive integers m if and only if A and B commute.

Proof: If $AB = BA$ then $W(A, B) = A^m B A^{-m} B^{-1} = B B^{-1} = I$, so W has all positive eigenvalues.

Conversely, assume the spectrum of $W(A, B) = A^m B A^{-m} B^{-1}$ is positive for every positive integer m . We want to show that A and B commute.

By (4.0.5) there exists a unitary U such that $A_1 = U^*AU = \begin{bmatrix} \lambda_1 I_{n_1} & 0 \\ 0 & \lambda_2 I_{n_2} \end{bmatrix}$, and we

have $n_1, n_2 > 0$. Let $B_1 = U^*BU$. Then $W(A_1, B_1) = U^*W(A, B)U$, and it suffices to show A_1 and B_1 commute.

Partition B_1 in the same way as A_1 : $B_1 = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$. To show A_1 and B_1 commute,

we need to prove that B_{12} and B_{21} are zero. Since B is positive definite so is B_1 and hence $B_{12} = B_{21}^*$, and the principal submatrices B_{11} and B_{22} are positive definite. Let $\gamma = \lambda_2 / \lambda_1$, so

$$A_1^m = \begin{bmatrix} \lambda_1^m I_{n_1} & 0 \\ 0 & \lambda_2^m I_{n_2} \end{bmatrix} = \lambda_1^m \begin{bmatrix} I_{n_1} & 0 \\ 0 & \gamma^m I_{n_2} \end{bmatrix}.$$

$$\text{Then } A_1^m B_1 A_1^{-m} = \lambda_1^m \begin{bmatrix} I_{n_1} & 0 \\ 0 & \gamma^m I_{n_2} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \lambda_1^{-m} \begin{bmatrix} I_{n_1} & 0 \\ 0 & \gamma^{-m} I_{n_2} \end{bmatrix}$$

$$= \begin{bmatrix} B_{11} & \gamma^{-m} B_{12} \\ \gamma^m B_{21} & B_{22} \end{bmatrix}.$$

If we let $C = B_{22} - B_{21}B_{11}^{-1}B_{12}$, then

$$B_1^{-1} = \begin{bmatrix} (B_{11} - B_{12}B_{22}^{-1}B_{21})^{-1} & B_{11}^{-1}B_{12}(-C)^{-1} \\ -C^{-1}B_{21}B_{11}^{-1} & C^{-1} \end{bmatrix} [\text{HJ, p. 18}].$$

Here C is positive definite because B_1 is positive definite [HJ, p. 475].

$$\text{Now } \gamma^m A_1^m B_1 A_1^{-m} B_1^{-1} = \begin{bmatrix} \gamma^m B_{11} & B_{12} \\ \gamma^{2m} B_{21} & \gamma^m B_{22} \end{bmatrix} B_1^{-1}, \text{ and } \gamma^m \rightarrow 0 \text{ as } m \rightarrow \infty \text{ since}$$

$0 < \gamma < 1$, so as $m \rightarrow \infty$ we have

$$\begin{aligned} \gamma^m A_1^m B_1 A_1^{-m} B_1^{-1} &\rightarrow \begin{bmatrix} 0 & B_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} (B_{11} - B_{12} B_{22}^{-1} B_{21})^{-1} & B_{11}^{-1} B_{12} (-C)^{-1} \\ -C^{-1} B_{21} B_{11}^{-1} & C^{-1} \end{bmatrix} \\ &= \begin{bmatrix} -B_{12} C^{-1} B_{21} B_{11}^{-1} & B_{12} C^{-1} \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

By hypothesis, $W(A_1, B_1)$ has positive spectrum so the eigenvalues of the limit matrix are nonnegative. Thus $(-B_{12} C^{-1} B_{21} B_{11}^{-1})$ has a nonnegative spectrum, which means the spectrum of $B_{12} C^{-1} B_{21} B_{11}^{-1}$ is nonpositive. However $B_{12} C^{-1} B_{21}$ is positive semidefinite since $B_{21} = B_{12}^*$: for any x , $x^* (B_{12} C^{-1} B_{12}^*) = y^* C^{-1} y \geq 0$ since C^{-1} is positive definite by (4.0.2). Therefore $B_{12} C^{-1} B_{21} B_{11}^{-1}$ is the product of a positive semidefinite and a positive definite matrix, so its eigenvalues are nonnegative by (4.2.5). Therefore here they must all be zero.

Again by Appendix C (4.2.5), because the positive semidefinite matrix $B_{12} C^{-1} B_{21}$ has all eigenvalues zero, it must be 0. Since C^{-1} is positive definite it has a positive definite square root [HJ, p. 405], so we have $0 = B_{12} C^{-1} B_{21} = (B_{12} C^{-1/2})(C^{-1/2} B_{12}^*) = (B_{12} C^{-1/2})(B_{12} C^{-1/2})^*$. For any matrix X , if $XX^* = 0$, then X must be 0, so we have $B_{12} C^{-1/2} = 0$, and $C^{-1/2}$ is invertible, so we can conclude that $B_{12} = 0$. This proves that B_1 commutes with A_1 .

2.2.8 Theorem [JHS, Theorem 3]. Suppose A and B are 2×2 real positive definite matrices. In the word $W(A, B) = A^{\alpha_1} B^{\beta_1} \dots A^{\alpha_N} B^{\beta_N}$, assume all α_i are of the same sign, or all β_i are of the same sign. Then all the eigenvalues of $W(A, B)$ are positive.

Proof: When $n = 2$, both A and B have eigenvalues of multiplicity $n - 1$. Thus the conditions of (2.2.6) are satisfied and therefore at least one eigenvalue of $W(A, B)$ is positive. Since $\det(W)$ is positive and is the product of its two eigenvalues, the second eigenvalue is also positive.

2.2.9 Theorem [JHS, Theorem 4]. Let A and B be 3×3 real positive definite matrices and suppose at least one of A or B has a multiple eigenvalue. Also suppose that all the powers of the other matrix in a word $W(A, B)$ are of the same sign. Then all the eigenvalues of $W(A, B)$ are positive.

Proof: Here $n - 1 = 2$ and we may suppose that A is the matrix with a multiple eigenvalue. Then one eigenvalue of A is of multiplicity at least $n - 1$. The conditions of (2.2.6) are met, so $\rho(W(A, B))$ is a positive eigenvalue of $W(A, B)$. Let $\eta_1 = \rho(W)$ and let η_2 and η_3 be the other two eigenvalues of W . Then $|\eta_3| \leq |\eta_2| \leq \eta_1$. Now

$$\begin{aligned} W(A, B)^{-1} &= (A^{\alpha_1} B^{\beta_1} \dots A^{\alpha_N} B^{\beta_N})^{-1} = B^{-\beta_N} A^{-\alpha_N} \dots B^{-\beta_1} A^{-\alpha_1} \\ &= (B^{-1})^{\beta_N} \dots (A^{-1})^{\alpha_1}. \end{aligned}$$

Thus $W(A, B)^{-1}$ is a word in A^{-1}, B^{-1} with positive exponents, A^{-1} and B^{-1} are both positive definite by Appendix A (4.0.2), and they also satisfy the conditions of (2.2.6).

Therefore $\rho(W^{-1}) > 0$ and $\rho(W^{-1})$ is an eigenvalue of W^{-1} . We also know $\rho(W^{-1}) = |1/\eta_3|$ and the eigenvalues of W^{-1} are $1/\eta_1, 1/\eta_2, 1/\eta_3$ hence $|1/\eta_3| = 1/\eta_3$ so $\eta_3 > 0$. Now we have η_1, η_3 real and positive, and $\det(W(A, B)) = \eta_1 \eta_2 \eta_3 > 0$ so η_2 is positive as well.

2.2.10 Theorem [JHS, Theorem 5]. Suppose that each of the $n \times n$ real positive definite matrices A and B has at most two distinct eigenvalues, and in $W(A, B)$, all powers of either A or B are of the same sign. Then for an arbitrary n , all the eigenvalues of $W(A, B)$ are positive.

Proof: Let λ_1, λ_2 be the eigenvalues of A , and μ_1, μ_2 be the eigenvalues of B . Let M be the eigenspace of A corresponding to λ_1 . Because A is symmetric, its eigenspaces are mutually orthogonal, hence $M^\perp$ is the eigenspace of A corresponding to λ_2 . Similarly, let N be the eigenspace of B corresponding to μ_1 , so $N^\perp$ is the eigenspace corresponding to μ_2 .

Case 1. If $\lambda_1 = \lambda_2$ or $\mu_1 = \mu_2$ then $A = \lambda_1 I$ or $B = \mu_1 I$ and in either case, the word $W(A, B)$ is a positive scalar times a power of A or B , hence is positive definite.

Case 2. Suppose $\lambda_1 \neq \lambda_2$ and $\mu_1 \neq \mu_2$. Define $P = \frac{1}{\lambda_1 - \lambda_2} (A - \lambda_2 I)$ and

$Q = \frac{1}{\mu_1 - \mu_2} (B - \mu_2 I)$, and observe that

$$P(x) = \begin{cases} x & \text{if } x \in M \\ 0 & \text{if } x \in M^\perp \end{cases} \quad \text{and} \quad Q(y) = \begin{cases} y & \text{if } y \in N \\ 0 & \text{if } y \in N^\perp. \end{cases}$$

Thus P and Q are orthoprojections and we can apply Theorem (4.3.6) of Appendix D.

Thus there exists a real orthogonal matrix U such that

$$U^T P U = \begin{bmatrix} P_1 & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & \dots & P_N \end{bmatrix} \quad \text{and} \quad U^T Q U = \begin{bmatrix} Q_1 & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & \dots & Q_N \end{bmatrix}$$

where for each j , P_j and Q_j are of same shape, either 1×1 or 2×2 .

Now $A = (\lambda_1 - \lambda_2)P + \lambda_2 I$ and $B = (\mu_1 - \mu_2)Q + \mu_2 I$ so it follows that

$$U^T A U = \begin{bmatrix} (\lambda_1 - \lambda_2)P_1 + \lambda_2 I & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & \dots & (\lambda_1 - \lambda_2)P_N + \lambda_2 I \end{bmatrix} \quad \text{and}$$

$$U^T B U = \begin{bmatrix} (\mu_1 - \mu_2)Q_1 + \mu_2 I & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & \dots & (\mu_1 - \mu_2)Q_N + \mu_2 I \end{bmatrix}$$

where each I that appears in these last two expressions is the same shape as the P_j to

which it is added. Therefore for each j , $A_j = (\lambda_1 - \lambda_2) P_j + \lambda_2 I$ and $B_j = (\mu_1 - \mu_2) Q_j + \mu_2 I$

are either 1×1 or 2×2 real positive definite matrices, and

$$U^T W(A, B) U = (U^T A^{\alpha_1} U U^T B^{\beta_1} U) \dots (U^T A^{\alpha_N} U U^T B^{\beta_N} U)$$

$$= \begin{bmatrix} W(A_1, B_1) & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & \dots & W(A_N, B_N) \end{bmatrix}.$$

From (2.2.8), the eigenvalues of each $W(A_j, B_j)$ are positive, hence this is also true for $U^T W(A, B) U$, and thus true for $W(A, B)$.

2.2.11 Definition [JHS, p.25]. Let $\alpha_1, \beta_1, \dots, \alpha_N, \beta_N$ be a sequence of nonzero real numbers and let k be a positive integer. We say $\alpha_1, \beta_1, \dots, \alpha_N, \beta_N$ is *k-good* if the word $W(A, B) = A^{\alpha_1} B^{\beta_1} \dots A^{\alpha_N} B^{\beta_N}$ has positive eigenvalues for all positive definite $k \times k$ matrices A and B .

Observe that any sequence $\alpha_1, \beta_1, \dots, \alpha_N, \beta_N$ where all α_i, β_i are positive or all are negative is 2-good because $A^{\alpha_1} B^{\beta_1} \dots A^{\alpha_N} B^{\beta_N}$ has both eigenvalues positive for all 2×2 positive definite A and B , by Theorem 2.2.3.

2.2.12 Theorem [JHS, p.25]. If j, k are both positive integers and $j < k$ then every k -good sequence is also j -good.

Proof: Let $\alpha_1, \beta_1, \dots, \alpha_N, \beta_N$ be a k -good sequence. Then $W(A, B) = A^{\alpha_1} B^{\beta_1} \dots A^{\alpha_N} B^{\beta_N}$ has positive eigenvalues when A and B are positive definite $k \times k$ matrices. Let A_1 and B_1

be $j \times j$ positive definite and define $A = \begin{bmatrix} A_1 & 0 \\ 0 & I_{k-j} \end{bmatrix}$ and $B = \begin{bmatrix} B_1 & 0 \\ 0 & I_{k-j} \end{bmatrix}$. Then $W(A, B) =$

$\begin{bmatrix} W(A_1, B_1) & 0 \\ 0 & I_{k-j} \end{bmatrix}$ is $k \times k$ and has all eigenvalues positive by hypothesis, so $W(A_1, B_1)$

must have positive eigenvalues.

2.2.13 **Definition** [JHS, p.25]. The *cancellation rule* for sequences $\alpha_1, \beta_1, \dots, \alpha_m, \beta_m$, $m \in \mathbb{N}$ is defined as follows: if $j \in \{1, \dots, m\}$ and if $\alpha_j \alpha_{j+1} > 0$ (where $\alpha_{m+1} = \alpha_m$), then α_j, β_j are discarded from the sequence. Similarly if $\beta_j \beta_{j+1} > 0$ then $\alpha_{j+1}, \beta_{j+1}$ are discarded.

2.2.14 **Definition** [JHS]. A sequence $\alpha_1, \beta_1, \dots, \alpha_m, \beta_m$ is said to be *irreducible of reduced class m* if no cancellations are possible. This means that consecutive α 's have alternate signs and consecutive β 's have alternate signs.

2.2.15 **Theorem** [JHS, Theorem 6]. Any sequence of real numbers $\alpha_1, \beta_1, \dots, \alpha_N, \beta_N$ which is of the reduced class $m \equiv 2 \pmod{4}$ or $m \equiv 3 \pmod{4}$ is not 2-good. That is, if $W(A, B)$ is the word with $\alpha_1, \beta_1, \dots, \alpha_N, \beta_N$ as its exponents, then there exist 2×2 positive definite A and B for which $W(A, B)$ has a negative eigenvalue.

Proof: Let $A = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \end{bmatrix}$ and $B = \begin{bmatrix} 1/2 + \varepsilon & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$ for some $\varepsilon > 0$, so

$$A^{-1} = \frac{1}{\varepsilon} \begin{bmatrix} \varepsilon & 0 \\ 0 & 1 \end{bmatrix} \text{ and } B^{-1} = \frac{2}{\varepsilon} \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 + \varepsilon \end{bmatrix}.$$

$$\text{Let } P = \lim_{\varepsilon \rightarrow 0} A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \text{ Then for } \alpha > 0, \lim_{\varepsilon \rightarrow 0} A^\alpha = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}^\alpha = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = P.$$

$$\text{Also } \lim_{\varepsilon \rightarrow 0} \varepsilon A^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = I - P. \text{ Hence for } \alpha > 0, \lim_{\varepsilon \rightarrow 0} (\varepsilon A^{-1})^\alpha = (I - P)^\alpha = I - P.$$

$$\text{Let } Q = \lim_{\varepsilon \rightarrow 0} B = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}. \text{ Then for } \beta > 0, \lim_{\varepsilon \rightarrow 0} B^\beta = Q^\beta \text{ and one can check that}$$

$$Q^\beta = Q.$$

Similarly, $\lim_{\varepsilon \rightarrow 0} ((\varepsilon/2)B^{-1}) = I - Q$ so for $\beta > 0$, $\lim_{\varepsilon \rightarrow 0} ((\varepsilon/2)B^{-1})^\beta = (I - Q)^\beta = I - Q$.

Let $W(A,B) = A^{\alpha_1}B^{\beta_1} \dots A^{\alpha_N}B^{\beta_N}$. Then whatever sign pattern $\alpha_1, \beta_1, \dots, \alpha_N, \beta_N$ has, there will exist positive real numbers c and d such that $\lim_{\varepsilon \rightarrow 0} \varepsilon^c(\varepsilon/2)^d A^{\alpha_1}B^{\beta_1} \dots A^{\alpha_N}B^{\beta_N} = P_1Q_1 \dots P_NQ_N$ where $P_i = P$ if $\alpha_i > 0$, $P_i = I - P$ if $\alpha_i < 0$, $Q_i = Q$ if $\beta_i > 0$, and $Q_i = I - Q$ if $\beta_i < 0$. In fact, c equals the negative of the sum of the negative α_i , while d is the negative of the sum of the negative β_i .

If $\alpha_i\alpha_{i+1} > 0$ then α_i, α_{i+1} are either both positive or both negative. A factor $P_iQ_iP_{i+1}$ is either PQP or $P(I - Q)P$, which both reduce to $(1/2)P$, or $(I - P)Q(I - P)$ or $(I - P)(I - Q)(I - P)$, which reduce to $(1/2)(I - P)$. Similarly if $\beta_i\beta_{i+1} > 0$ then the factors $Q_iP_{i+1}Q_{i+1}$ are QPQ or $Q(I - P)Q$ which reduce to $(1/2)Q$, or $(I - Q)P(I - Q)$ or $(I - Q)(I - P)(I - Q)$, which reduce to $(1/2)(I - Q)$. After doing cancellation to get a reduced word with exponents $a_1, b_1, \dots, a_m, b_m$, the a_i 's will alternate in sign and so will the b_i 's. Further we can use A or A^{-1} and B or B^{-1} originally, to ensure that $a_1, b_1 > 0$. Thus the reduced limit word has to be of the form

$(PQ(I - P)(I - Q)) \dots (PQ(I - P)(I - Q)) = (PQ(I - P)(I - Q))^{m/2}$ where m is even, or $(PQ(I - P)(I - Q)) \dots (PQ(I - P)(I - Q)) PQ = (PQ(I - P)(I - Q))^{(m-1)/2} PQ$ where m is odd.

By induction we can see that for any positive integer k ,

$$(PQ(I - P)(I - Q))^k = \frac{1}{4^k} \begin{bmatrix} (-1)^k & (-1)^{k-1} \\ 0 & 0 \end{bmatrix}.$$

Thus $PQ(I - P)(I - Q))^{m/2}$ is equal to $\frac{1}{2^m} \begin{bmatrix} (-1)^{m/2} & (-1)^{(m/2)-1} \\ 0 & 0 \end{bmatrix}$ if m is even,

or $\frac{1}{2^{m-1}} \begin{bmatrix} (-1)^{(m-1)/2} & (-1)^{(m-1)/2-1} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 0 & 0 \end{bmatrix}$ if m is odd. Therefore the trace of the

reduced limit word is negative if and only if either m is even and $m/2$ is odd, which implies m is of the form $4n + 2$, or m is odd and $(m-1)/2$ is odd, which implies m is of the form $4n + 3$.

2.3 Nearly Symmetric Words

2.3.1 **Definition** [JH]. We call a generalized word *nearly symmetric* if it is either symmetric or the product of two symmetric words.

2.3.2 **Proposition** [JH]. If A and B are symmetric and invertible and $W(A, B)$ is nearly symmetric then the three transformations: (i) reversal, (ii) cyclic permutation and (iii) interchange of A and B , yield nearly symmetric words.

Proof: (i) If a word is symmetric then its reversal is the same word again. If a word is the product of two symmetric words, say $W_1 W_2$, then $\text{Reversal}(W_1 W_2) = W_2 W_1$ which is again a nearly symmetric word as it is the product of two symmetric matrices.

(ii) Let $W_1 W_2$ be a nearly symmetric word in the letters A, B where W_1 and W_2 are symmetric. We may suppose $W_1 W_2 = A \hat{W} W_2$ where $W_1 = A \hat{W}$ and $\hat{W}$ could be I. Then cyclic permutation of $W_1 W_2$ yields $\hat{W} W_2 A = \hat{W} A^{-1} A W_2 A = (\hat{W} A^{-1})(A W_2 A)$. Since $A W_2 A$ is symmetric we only have to prove that $\hat{W} A^{-1}$ is also symmetric. We have $W_1 = A \hat{W}$ so $\hat{W} A^{-1} = A^{-1} W_1 A^{-1}$, which is symmetric.

(iii) Let $W = W_1 W_2$ be a nearly symmetric word in the letters A, B, where W_1 and W_2 are symmetric. Then $W_1(B, A)$ and $W_2(B, A)$ are symmetric, hence $W(B, A)$ is nearly symmetric. Therefore a nearly symmetric word in A and B is again nearly symmetric after the interchange of A and B.

The following concept and lemma will be used in Theorem (2.3.5).

2.3.3 Definition [JH]. A diagonalizable matrix with positive eigenvalues is called *quasi-positive*.

2.3.4 Lemma [JH, Lemma 2]. The $n \times n$ matrix Q is quasi-positive if and only if $Q = AB$, for some positive definite A and B . Moreover if $Q = SDS^{-1}$, with D a diagonal matrix with positive diagonal entries, and if $Q = AB$ where A and B are positive definite, then there exists a positive definite matrix E that commutes with D so that $A = SES^*$ and $B = (S^{-1})^* E^{-1} D S^{-1}$.

Proof: Let $Q = AB$, where A and B are positive definite matrices. Then Q is similar to $(A^{-1/2}) AB (A^{1/2})$, which equals $A^{1/2} B A^{1/2}$. Then by Sylvester's Law of Inertia (4.2.4), Q and B have same inertia, and B is positive definite so Q has positive eigenvalues. Since Q is similar to the diagonalizable matrix B , Q is also diagonalizable hence Q is quasi positive.

Conversely, suppose Q is quasi-positive. Then there exist invertible S and diagonal D with positive diagonal entries so that $Q = SDS^{-1}$. Let E be any positive definite diagonal matrix and note E commutes with D . Define $A = SES^*$ and $B = (S^{-1})^* E^{-1} D S^{-1}$. Then $Q = AB$ because $Q = SDS^{-1} = SES^* (S^{-1})^* E^{-1} D S^{-1} = (SES^*) ((S^{-1})^* E^{-1} D S^{-1}) = AB$.

We will now show that A and B are positive definite. Since A and E are congruent, A has positive eigenvalues. Also, $A^* = (SES^*)^* = SES^* = A$, so A is positive definite. Similarly, $E^{-1}D$ is a positive definite diagonal matrix congruent to B , and $B = B^*$ because E and D commute, hence B is also positive definite.

For the rest, suppose $Q = AB$ is a factorization of Q into positive definite A and B , and $Q = SDS^{-1}$, where D is diagonal with positive diagonal entries. We need to construct positive definite E which commutes with D so that $A = SES^*$ and $B = (S^{-1})^* E^{-1} D S^{-1}$ will be true. Since $Q = AB$, $B = A^{-1}Q$; B is positive definite so $B = B^*$, hence $A^{-1}Q = Q^* A^{-1}$, which implies $AQ^* = QA$. Substituting $Q = SDS^{-1}$ gives:

$$A (S^{-1})^* D S^{-1} = SDS^{-1} A$$

This gives $S^{-1} A (S^{-1})^* D S^{-1} = D S^{-1} A$

$$(S^{-1} A (S^{-1})^*) D = D (S^{-1} A (S^{-1})^*)$$

This means $S^{-1}A(S^{-1})^*$ commutes with D . If we let $E = S^{-1}A(S^{-1})^*$, we have shown this E commutes with D , and E is positive definite because it is symmetric and congruent to A so its eigenvalues are all positive.

$$\begin{aligned}\text{Finally, } B &= A^{-1}Q = (SES^*)^{-1}(SDS^{-1}). \\ &= (S^*)^{-1}E^{-1}S^{-1}SDS^{-1} \\ &= (S^{-1})^*E^{-1}DS^{-1}.\end{aligned}$$

2.3.5 Theorem [JH, Theorem 3]. Every nearly symmetric word $W(A,B)$ in two real positive definite letters is quasi-positive, hence has all eigenvalues positive.

Proof: If $W(A,B)$ is a symmetric word then W is positive definite by (2.2.2) so it has positive eigenvalues. Otherwise $W(A,B) = W_1W_2$ where both W_1 and W_2 are symmetric, hence positive definite by (2.2.2). Therefore W is quasi-positive by (2.3.4), hence has positive eigenvalues.

2.3.6 Theorem [JH, Theorem 4]. If A and B are symmetric, a word $W(A,B)$ in which one of the letters appears at most twice is nearly symmetric.

Proof: Let B be the letter that appears at most twice.

Case 1. If the word contains no B then it is a power of A and hence symmetric.

Case 2. If the letter B appears only once, then the word will be of the form A^pBA^q where $p, q \geq 0$. If $p \geq q$ then we have $W(A,B) = A^{p-q}(A^qBA^q)$ which is a product of two symmetric words, hence nearly symmetric. If $p \leq q$ then we have $W(A,B) =$

$(A^p B A^p) A^{q-p}$ which is again nearly symmetric.

Case 3. If the word contains exactly two B's then it will be of the form

$A^p B A^q B A^t$, where $p, q, t \geq 0$. Then $W(A, B) = (A^p B A^q B A^p) A^{t-p}$ if

$t \geq p$ or $W(A, B) = A^{p-t} (A^t B A^q B A^t)$ if $p \geq t$.

We now know that a word which is not nearly symmetric must have at least 3 each of two different letters A and B, hence has length at least 6. By Example (1.3.2), there are only three equivalence classes for words of length 6 which have 3 A's and 3 B's. By inspection we see that only ABA^2B^2 is not nearly symmetric.

2.3.7 Corollary [JH, Corollary 5]. Every nearly symmetric word in real positive definite A and B, and thus every such word of length less than six, has only positive eigenvalues.

Proof: By (2.3.6) we know every nearly symmetric word in such A and B has only positive eigenvalues. If a word is of length less than 6 then without loss of generality we may consider three cases: zero B's, one B and two B's. (Three A's and three B's will not occur since the length would no longer be less than 6.) By (2.3.6), such a word is nearly symmetric, hence has only positive eigenvalues by (2.3.5).

Thus nearly symmetric words are very nice, but unfortunately the following result shows that they occur less and less frequently as the length of words increases.

2.3.8 **Theorem** [JH, p.918]. The fraction of nearly symmetric words of length L to the total number of words of length L approaches zero as L goes to infinity.

Proof: The number of nearly symmetric words of length L is $O(L)(2^{3L/4})$ [K]. The number of total words of length L is 2^L , since each place in such a word can be filled with A or B.

Then $\lim_{L \rightarrow \infty} (L)(2^{(3/4)L}) / 2^L = \lim_{L \rightarrow \infty} L / 2^{L/4} = 0$.

CHAPTER 3 WORDS THAT DO NOT ALWAYS HAVE POSITIVE EIGENVALUES

3.1 Words with Nonpositive Eigenvalues

3.1.1 Theorem [JH, Theorem 8]. Let $W(A,B)$ be any word for which there exist positive definite A and B such that W has an eigenvalue that is not positive. Then, for any positive integer k , there exist positive definite letters A and B such that W^k has a nonpositive eigenvalue.

Proof: Let A and B be positive definite matrices that give W a nonpositive eigenvalue λ . If k is a positive integer then λ^k is an eigenvalue of $W(A, B)^k$. If λ^k is nonpositive, we are done. If λ^k is positive we will show a way to create a new pair of positive definite matrices A' and B' for which $W(A', B')^k$ has a nonpositive eigenvalue.

An outline of the approach used is as follows: First, we parameterize a pair of positive definite matrices in terms of a real variable t , $0 \leq t \leq 1$, and then examine the eigenvalues of the word W^k evaluated at those matrices. Second, using the continuity of eigenvalues on matrix entries, we then show that for some t between 0 and 1, $W(A(t), B(t))^k$ has a nonpositive eigenvalue.

Let λ_A be the largest eigenvalue of A , and let λ_B be the largest eigenvalue of B .

Define the following parameterization:

$$A(t) = t(\lambda_A I - A) + A \quad \text{and} \quad B(t) = t(\lambda_B I - B) + B, \quad \text{for } 0 \leq t \leq 1.$$

Note that $A(t)$ and $B(t)$ are positive definite for all $0 \leq t \leq 1$ since $t(\lambda_A I - A)$ is positive semidefinite: for if the eigenvalues of A are $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n = \lambda_A$ then the eigenvalues of

$\lambda_A I - A$ are $\lambda_A - \lambda_1 \geq \lambda_A - \lambda_2 \geq \dots \geq \lambda_A - \lambda_n = 0$. Similarly, the eigenvalues of $B(t)$ are all nonnegative for all such t .

Also notice that $A(1) = \lambda_A I$ and $B(1) = \lambda_B I$, giving the word $W(A(1), B(1))$ positive eigenvalues. Since $A(0) = A$ and $B(0) = B$, $W(A(0), B(0))$ has an eigenvalue that is not positive, by assumption.

Let $W_t = W(A(t), B(t))$ and notice $\det(W_t) > 0$ for each $t \in [0, 1]$. Define $\Gamma = \{ t \in [0, 1] : \text{each eigenvalue of } W_t \text{ is positive} \}$. Then $0 \notin \Gamma$ by hypothesis, and $1 \in \Gamma$ since W_1 has positive eigenvalues. Thus Γ is not empty and its greatest lower bound, say t_M , exists and satisfies $0 < t_M \in \Gamma$.

Since the eigenvalues of a matrix depend continuously on its entries [HJ, Appendix D], t_M must be in Γ : for there exist W_t arbitrary close to W_{t_M} which have real positive eigenvalues, so those of W_{t_M} must be real and nonnegative; and $\det(W_{t_M}) \neq 0$ so none of its eigenvalues are zero. Therefore they are all positive, hence $t_M \in \Gamma$.

Now, let k be an integer, $k \geq 2$. Using continuity of eigenvalues again and the fact that all eigenvalues of W_{t_M} are real and positive, there is an $\varepsilon > 0$ such that for any $t \in (t_{M-\varepsilon}, t_{M+\varepsilon})$, each eigenvalue $\lambda = re^{i\theta}$ of W_t has $-\pi/k < \theta < \pi/k$. By definition of Γ and the greatest lower bound, there must be some $t \in (t_{M-\varepsilon}, t_M)$ for which W_t has an eigenvalue λ that is not real. So for this λ , $-\pi/k < \theta < \pi/k$ and $\theta \neq 0$, hence $k\theta$ is not an integer multiple of π , so λ^k is not real. Therefore λ^k , which is an eigenvalue of W_t^k , is not real and hence nonpositive. $\square$

3.1.2 **Theorem** [JH, Theorem 9]. If j and k are different positive integers, then there exist a 3×3 , real, quasi-positive matrix S such that $S^k(S^T)^j$ has a negative trace.

Proof. We can assume $k > j$ since $S^j(S^T)^k$ has the same trace as its transpose $S^k(S^T)^j$.

Let $S = \begin{bmatrix} 1 & x & z \\ 0 & a & y \\ 0 & 0 & b \end{bmatrix}$, where $1 < a < b$ and $y, z \neq 0$. We will show that if x is sufficiently

large, then $S^k(S^T)^j$ has negative trace.

$$\begin{aligned} \text{Observe that } S^2 &= \begin{bmatrix} 1 & x+ax & z+xy+zb \\ 0 & a^2 & ay+by \\ 0 & 0 & b^2 \end{bmatrix} = \begin{bmatrix} 1 & X_2 & Z_2 \\ 0 & a^2 & Y_2 \\ 0 & 0 & b^2 \end{bmatrix}, \\ S^3 &= \begin{bmatrix} 1 & x+a(x+ax) & z+y(x+ax)+b(z+xy+zb) \\ 0 & a^3 & a^2y+b(ay+by) \\ 0 & 0 & b^3 \end{bmatrix} = \begin{bmatrix} 1 & x+aX_2 & z+yX_2+bZ_2 \\ 0 & a^3 & a^2y+bY_2 \\ 0 & 0 & b^3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & X_3 & Z_3 \\ 0 & a^3 & Y_3 \\ 0 & 0 & b^3 \end{bmatrix} \text{ and in general } S^k = \begin{bmatrix} 1 & x+aX_{k-1} & z+yX_{k-1}+bZ_{k-1} \\ 0 & a^k & a^{k-1}y+bY_{k-1} \\ 0 & 0 & b^k \end{bmatrix} \\ &= \begin{bmatrix} 1 & X_k & Z_k \\ 0 & a^k & Y_k \\ 0 & 0 & b^k \end{bmatrix}. \end{aligned}$$

Also observe that $X_k = x + a X_{k-1}$

$$\begin{aligned} &= x + a(x + a X_{k-2}) = x + ax + a^2 X_{k-2} \\ &= x + ax + a^2(x + a X_{k-3}) = x + ax + a^2x + a^3 X_{k-3} \\ &= x(1 + a + a^2 + \dots + a^{k-1}) \end{aligned}$$

$$= x (a^k - 1) / (a - 1) \text{ because } a \neq 1. \quad (3.1.3)$$

$$\begin{aligned} \text{Similarly, } Y_k &= y a^{k-1} + b Y_{k-1} \\ &= y a^{k-1} + b(y a^{k-2} + b Y_{k-2}) \\ &= y(a^{k-1} + b a^{k-2}) + b^2(y a^{k-3} + b Y_{k-3}) \\ &= y(a^{k-1} + b a^{k-2} + b^2 a^{k-3} + \dots + b^{k-1}) \\ &= y(a^k - b^k) / (a - b) \text{ because } a \neq b. \end{aligned} \quad (3.1.4)$$

$$\begin{aligned} \text{And } Z_k &= z + y X_{k-1} + b Z_{k-1} \\ &= z + y X_{k-1} + b(z + y X_{k-2} + b Z_{k-2}) = z(1 + b) + y(X_{k-1} + b X_{k-2}) + b^2 Z_{k-2} \\ &= z(1 + b + b^2 + \dots + b^{k-2}) + y(X_{k-1} + b X_{k-2} + b^2 X_{k-3} + \dots + b^{k-2} X_1) + b^{k-1} z \end{aligned}$$

$$\begin{aligned} \text{By (3.1.3), } Z_k &= z(1 + b + b^2 + \dots + b^{k-2} + b^{k-1}) + \\ &\quad xy \left(\frac{a^{k-1} - 1}{a - 1} + b \frac{a^{k-2} - 1}{a - 1} + b^2 \frac{a^{k-3} - 1}{a - 1} + \dots + b^{k-2} \right) \end{aligned} \quad (3.1.5)$$

$$\begin{aligned} \text{And by (3.1.4) this equals } & xy \frac{1}{a - 1} \left(\frac{a^k - b^k}{a - b} - \frac{b^k - 1}{b - 1} \right) + z \frac{b^k - 1}{b - 1} \\ &= xy C_k + z D_k \end{aligned} \quad (3.1.6)$$

$$\text{where } C_k = \frac{1}{a - 1} \left(\frac{a^k - b^k}{a - b} - \frac{b^k - 1}{b - 1} \right) \text{ and } D_k = \frac{b^k - 1}{b - 1} \text{ depend only on } a, b, \text{ and } k.$$

Thus the trace of $S^k(S^T)^j$ can be computed explicitly in terms of x, y, z, a, b, k, j :

$$\begin{aligned}
\text{Tr}[S^k(S^T)^j] &= \text{Tr} \left[\begin{pmatrix} 1 & X_k & Z_k \\ 0 & a^k & Y_k \\ 0 & 0 & b^k \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ X_j & a^j & 0 \\ Z_j & Y_j & b^j \end{pmatrix} \right] \\
&= (1 + X_k X_j + Z_k Z_j) + (a^{k+j} + Y_k Y_j) + b^{k+j} \\
&= 1 + a^{k+j} + b^{k+j} + x^2 \frac{a^k - 1}{a - 1} \cdot \frac{a^j - 1}{a - 1} + y^2 \frac{a^k - b^k}{a - b} \cdot \frac{a^j - b^j}{a - b} + x^2 y^2 C_k C_j \\
&\quad + xy (C_k D_j + C_j D_k) z + D_k D_j z^2 \\
&= P + Qz + Rz^2.
\end{aligned}$$

Recall that $1 < a < b$, let $y = x$, and view $\text{Tr}[S^k(S^T)^j]$ as a quadratic polynomial in z . Since $R > 0$, this polynomial will take on negative values if and only if its discriminant $Q^2 - 4PR$ is positive. This discriminant is a quadratic in x^2 ; therefore, if we can show that its leading coefficient is always positive, this will demonstrate that for sufficiently large values of x , the discriminant will also be positive. The coefficient of x^4 in the discriminant is:

$$(C_k D_j + C_j D_k)^2 - 4D_k D_j (C_k C_j) = (C_k D_j - C_j D_k)^2.$$

This is nonnegative, hence we have to prove that whenever $k \neq j$, $C_k D_j \neq C_j D_k$, so we calculate:

$$\begin{aligned}
C_k D_j - C_j D_k &= \frac{1}{a-1} \left(\frac{a^k - b^k}{a-b} - \frac{b^k - 1}{b-1} \right) \left(\frac{b^j - 1}{b-1} \right) - \frac{1}{a-1} \left(\frac{a^j - b^j}{a-b} - \frac{b^j - 1}{b-1} \right) \left(\frac{b^k - 1}{b-1} \right) \\
&= \frac{1}{(a-1)(b-1)} \left[\left(\frac{a^k - b^k}{a-b} - \frac{b^k - 1}{b-1} \right) (b^j - 1) - \left(\frac{a^j - b^j}{a-b} - \frac{b^j - 1}{b-1} \right) (b^k - 1) \right].
\end{aligned}$$

Thus if we prove that the expression $a^j b^{k+1} - a^k + b^k + a^k b - b^{k+1} + a^k b^j - a^k b^{j+1} - a^j b^k - b^j + a^j - a^j b + b^{j+1}$ is never zero unless $k = j$, we will be done. On factorization, we get

$a^j b^k (b-1) + a^k (b-1) - b^k (b-1) - a^k b^j (b-1) + b^j (b-1) - a^j (b-1)$. Since $(b-1) \neq 0$,

we need to prove only that $f(a, b) = a^j b^k + a^k - b^k - a^k b^j + b^j - a^j$ is never zero when $k \neq j$.

To do that, examine the following polynomial in x :

$$g(x) = f(a, x) = x^k (a^j - 1) + x^j (1 - a^k) + a^k - a^j.$$

We note that $g(1) = 0$ and $g(a) = 0$. Since $a > 1$ and $k > j$, there are two sign changes in

the consecutive coefficients of $g(x)$. The Descartes rule of signs states that for a

polynomial $P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ with real coefficients, the number of

positive real roots of the equation $P_n(x) = 0$ is either equal to the number of sign changes

between successive terms of $P_n(x)$ or a reduction of this number by a multiple of 2 [H,

p.223]. Since a and 1 are two different roots, g has no more positive roots. Hence,

$g(b) \neq 0$ because $b \neq 1$ and $b \neq a$.

This completes the proof that $S^k (S^T)^j$ will have negative trace when x is sufficiently large.

Observe that the same conclusion can be drawn if $S = \begin{bmatrix} a_1 & x_1 & z_1 \\ 0 & b_1 & y_1 \\ 0 & 0 & c_1 \end{bmatrix}$ where

$0 < a_1 < b_1 < c_1$ and $y_1, z_1 \neq 0$; just apply the above proof to $(1/a_1)S$. $\square$

Example (2.1.5) illustrates the proof above, and it seems likely that the result above was motivated by that example. There a specific matrix C was constructed which is like the matrix S used above, and it was shown that $C^2 C^T$ has negative trace.

3.1.3 Theorem [JH, Theorem 10]. For integers $p, q \geq 2$ and the word $W = ABA^p B^q$, there exist positive definite matrices A and B such that $W(A, B)$ has a negative trace.

Proof: By transposing and using cyclic permutations, we see that $ABA^p B^q$ has the same trace as $BAB^q A^p$. Let $F(p, q) = \text{Tr}[ABA^p B^q] = \text{Tr}[BAB^q A^p]$.

Suppose A and B are fixed positive definite real matrices. Then $A = U^T D U$ and $B = V^T E V$ where U and V are real orthogonal matrices, and $D = \text{diag}(a, b, c)$ and $E = \text{diag}(r, s, t)$ are diagonal matrices with $a, b, c, r, s, t > 0$.

Since $\text{Tr}[BAB^q A^p] = \text{Tr}[UBAB^q A^p U^T] = \text{Tr}[UBAB^q U^T D^p]$ and $\text{Tr}[ABA^p B^q] = \text{Tr}[VABA^p B^q V^T] = \text{Tr}[VABA^p V^T E^q]$, we have

$$F(p, q) = \text{Tr} \left[UBAB^q U^T \begin{pmatrix} a^p & 0 & 0 \\ 0 & b^p & 0 \\ 0 & 0 & c^p \end{pmatrix} \right] = g_1(q)a^p + g_2(q)b^p + g_3(q)c^p \quad (1)$$

$$\text{and } F(p, q) = \text{Tr} \left[VABA^p V^T \begin{pmatrix} r^q & 0 & 0 \\ 0 & s^q & 0 \\ 0 & 0 & t^q \end{pmatrix} \right] = h_1(p)r^q + h_2(p)s^q + h_3(p)t^q \quad (2)$$

Observe that $g_i(q)$ is a linear function in r^q, s^q, t^q and $h_i(p)$ is a linear function in a^p, b^p, c^p .

If $F(p, q)$ is known for $2 \leq p, q \leq 4$, we claim the following recurrence relations are valid:

$$F(p, q) = (a + b + c) F(p - 1, q) - (ab + bc + ca) F(p - 2, q) + (abc) F(p - 3, q)$$

for any $p \geq 5$ and $q \geq 2$. (3)

$$F(p, q) = (r + s + t) F(p, q - 1) - (rs + st + tr) F(p, q - 2) + (rst) F(p, q - 3)$$

for any $p \geq 2$ and $q \geq 5$. (4)

Proof of (3): The expression on the right equals

$$\begin{aligned} & (a + b + c) (g_1(q)a^{p-1} + g_2(q)b^{p-1} + g_3(q)c^{p-1}) \\ & - (ab + bc + ca)(g_1(q)a^{p-2} + g_2(q)b^{p-2} + g_3(q)c^{p-2}) \\ & + abc(g_1(q)a^{p-3} + g_2(q)b^{p-3} + g_3(q)c^{p-3}) \\ & = (g_1(q)a^p + g_1(q)a^{p-1}b + g_1(q)a^{p-1}c) + (g_2(q)ab^{p-1} + g_2(q)b^p + g_2(q)b^{p-1}c) + (g_3(q)ac^{p-1} \\ & + g_3(q)bc^{p-1} + g_3(q)c^p) - (g_1(q)a^{p-1}b + g_1(q)a^{p-2}bc + g_1(q)a^{p-1}c) - (g_2(q)ab^{p-1} + g_2(q)b^{p-1}c \\ & + g_2(q)ab^{p-2}c) - (g_3(q)abc^{p-2} + g_3(q)bc^{p-1} + g_3(q)ac^{p-1}) + (g_1(q)a^{p-2}bc + g_2(q)ab^{p-2}c \\ & + g_3(q)abc^{p-2}) \\ & = g_1(q)a^p + g_2(q)b^p + g_3(q)c^p = F(p, q). \end{aligned}$$

The proof of (4) is similar.

$$\text{Now let } A = \begin{bmatrix} 1 & 20 & 210 \\ 20 & 402 & 4240 \\ 210 & 4240 & 44903 \end{bmatrix} \text{ and } B = \begin{bmatrix} 36501 & -3820 & 190 \\ -3820 & 401 & -20 \\ 190 & -20 & 1 \end{bmatrix}. \text{ We will show}$$

that the trace of $W(A, B)$ is negative for these matrices.

The characteristic polynomials of A and B are as follows:

$P_A(t) = t^3 - 45306t^2 + 74211t - 6$; $P_B(t) = t^3 - 36903t^2 + 44903t - 1$. The coefficients of these can be expressed in terms of their roots, a, b, c and r, s, t respectively:

$$(a+b+c), (ab+bc+ca), abc, (r+s+t), (rs+st+tr), rst,$$

and these expressions appear in (3) and (4). In this example we have

$$F(p,q) = 45306 F(p-1, q) - 74211 F(p-2, q) + 6 F(p-3, q) \quad (5)$$

$$F(p,q) = 36903 F(p, q-1) - 44903 F(p, q-2) + F(p, q-3) \quad (6)$$

We will now show that $F(p, q)$ is negative for all $p, q \geq 2$. For the smaller values $2 \leq p, q \leq 4$, the following table gives the values of $F(p, q)$.

	$q = 2$	$q = 3$	$q = 4$
$p = 2$	-3164	-171233664	-6318893781764
$p = 3$	-219049002	-10537988104302	-388873536893369802
$p = 4$	-9923997300324	-477421308542380824	-17617832833924812095724

Now we will prove that for all $p, q \geq 2$, the following inequalities hold:

$$F(p, q) < 0, \quad (7)$$

$$F(p, q) < 10 F(p-1, q) \text{ for } p > 2 \quad (8)$$

$$F(p, q) < 10 F(p, q-1) \text{ for } q > 2. \quad (9)$$

By the table, (7) – (9) are valid for $2 \leq p, q \leq 4$. Assume this to be true for all

$2 \leq p, q < N$, where $N \geq 5$. We want to show that (7) – (9) are true for $2 \leq p, q \leq N$. First

verify (7) – (9) for $p = N$ and $q < N$, and for $p < N$ and $q = N$:

$$F(N, q) = 45306 F(N-1, q) - 74211 F(N-2, q) + 6 F(N-3, q) \quad (\text{by (5)})$$

$$< 45306 F(N-1, q) - 74211 F(N-2, q) \quad (\text{by (7)})$$

$$< 45306 F(N-1, q) - 7421.1 F(N-1, q) \quad (\text{by (8)})$$

$$= 37884.9 F(N-1, q)$$

$$< 10 F(N-1, q) \quad (10)$$

$$< 0$$

In the same way, using (6), (7) and (9), we get:

$$\begin{aligned}
F(p, N) &= 36903 F(p, N-1) - 44903 F(p, N-2) + F(p, N-3) && \text{(by (6))} \\
&< 36903 F(p, N-1) - 44903 F(p, N-2) && \text{(by (7))} \\
&< 36903 F(p, N-1) - 4490.3 F(p, N-1) && \text{(by (9))} \\
&= 32412.7 F(p, N-1) \\
&< 10 F(p, N-1) && (11) \\
&< 0
\end{aligned}$$

To show $F(N, N) < 10 F(N, N-1)$, substitute $F(N-1, N)$, $F(N-2, N)$ and $F(N-3, N)$ from (6) into the right-hand side of (5).

$$\begin{aligned}
F(N, N) &= 1671927318 F(N-1, N-1) - 2034375318 F(N-1, N-2) + 45306 F(N-1, N-3) \\
&\quad - 2738608533 F(N-2, N-1) + 3332296533 F(N-2, N-2) - 74211 F(N-2, N-3) \\
&\quad + 221418 F(N-3, N-1) - 269418 F(N-3, N-2) + 6 F(N-3, N-3) \\
&< 1671927318 F(N-1, N-1) - 203437531.8 F(N-1, N-1) \\
&\quad - 273860853.3 F(N-1, N-1) - 74.211 F(N-1, N-1) - 269.418 F(N-1, N-1) \\
&= 1194628589.271 F(N-1, N-1). && (12)
\end{aligned}$$

From (11), substituting $p = N-1$, we get

$$36903 F(N-1, N-1) = F(N-1, N) + 44903 F(N-1, N-2) - F(N-1, N-3)$$

$$< F(N-1, N) - (1/100) F(N-1, N-1)$$

Thus, $36903.01 F(N-1, N-1) < F(N-1, N)$ and substituting this result in (12) gives us
 $F(N, N) < 1194628589.271 F(N-1, N-1) < 10 F(N-1, N)$.

In order to prove the inequality $F(N, N) < 10 F(N-1, N)$, we do the same steps, using (10) instead of (11):

$$\begin{aligned} 45306 F(N-1, N) &= F(N, N-1) + 74211 F(N-2, N) - 6 F(N-3, N) \\ &< F(N, N-1) - (1/100) F(N-1, N) \end{aligned}$$

Therefore, $45306.01 F(N-1, N) < F(N, N-1)$ and again from (12) we obtain the inequality
 $F(N, N) < 10 F(N, N-1)$. This completes the induction showing that for all $p, q \geq 2$,
 $F(p, q) < 0$. $\square$

References:

- [BMV] D. Bessis, P. Moussa, M. Villani. “ Monotonic converging variational approximations to the Functional Integrals in quantum Statistical Mechanics,” *Journal of Mathematical Physics*, 16 (1975), 2318-2325.
- [Davis] C. Davis. “Seperation of Two Linear Subspaces,” *Acta Scientiarum Mathematicarum*, 19 (1958), 172-187.
- [FIS] S. Friedberg, A.J. Insel, L. Spence. *Linear Algebra*. Third Edition, Prentice Hall, N.J., 1999.
- [Halmos] P. L. Halmos. “Two Subspaces,” *Trans. Amer. Math. Society*, 144 (1969), 381-389.
- [H] L. I. Holder. *College Algebra*. Fifth Edition. Wadsworth Publishing Company, Belmont, CA, 1991.
- [HJ] R. Horn, C.R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1999.
- [HJS] C.J. Hillar, C.R. Johnson, and I.M. Spitkovsky. “Positive Eigenvalues and Two-letter Generalized Words,” *The Electronic Journal of Linear Algebra*, 9 (2002), 21-26.
- [JH] C.R. Johnson, C.J. Hillar. “Eigenvalues of words in two positive definite letters,” *SIAM J. Matrix Analysis Applications*, 23 (2002), 916-928.
- [K] R. Kemp. On the number of words in the language $\{w \in \Sigma^* \mid w = w^R\}^2$, *Discrete Mathematics*, 40 (1982), 225-234.
- [Lay] D. Lay. *Linear Algebra and its Applications*. Third Edition. Addison-Wesley, Boston, 2003.
- [LS] E. Lieb, R. Seiringer. “Equivalent Form of the Bessis-Moussa-Villani Conjecture,” Manuscript Submitted for Publication.

APPENDIX A

Positive Definite and Nonnegative Matrices

4.0.1 Definition [HJ, p. 396]. An $n \times n$ Hermitian matrix A is said to be *positive definite* (positive semidefinite) if $x^*Ax > 0$ for all $0 \neq x \in \mathbb{C}^n$ ($x^*Ax \geq 0$ for all $x \in \mathbb{C}^n$).

4.0.2 Theorem [HJ, p. 398]. Each eigenvalue of a positive (semi) definite matrix is a positive (nonnegative) real number. If A is positive definite, then A^{-1} is also Hermitian and has positive eigenvalues, hence is positive definite.

4.0.3 Theorem [Lay, p. 461]. A real matrix is positive definite (semi) definite if and only if $A = A^T$ and $x^T Ax > 0$ for all $0 \neq x \in \mathbb{R}^n$ ($x^T Ax \geq 0$ for all $0 \neq x \in \mathbb{R}^n$).

4.0.4 Theorem [Lay, p. 461]. A real matrix A is positive (semi) definite if and only if $A = A^T$ and all the eigenvalues of A are positive (nonnegative).

4.0.5 Theorem. If A is a positive (semi) definite matrix then there exist a unitary matrix U such that U^*AU is a real diagonal matrix. If A is a real matrix, we can choose U to be a real orthogonal matrix.

Proof: By Schur's theorem [HJ, p.79] for a given $n \times n$ matrix A with eigenvalues $\lambda_1, \dots, \lambda_n$, there is an $n \times n$ unitary matrix U such that $U^*AU = T$ is upper triangular, with diagonal entries $t_{ii} = \lambda_i, i = 1, \dots, n$.

Since $A = A^*$ here, $T = T^*$. But T^* is lower triangular and is equal to the upper triangular T only if it is diagonal. Therefore $T = T^* = D$ where D is a real diagonal matrix with λ_i , $i = 1, \dots, n$ as its diagonal entries. $\square$

4.0.6 Theorem. If A is positive definite then A^α can be defined for each real number α so that A^α is positive definite and $(A^\alpha)^{1/\alpha} = A$ is true when $\alpha \neq 0$. If A is positive semidefinite then A^α can be defined for each real $\alpha > 0$ so that A^α is positive semidefinite and $(A^\alpha)^{1/\alpha} = A$. In either case, if $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A , then $\lambda_1^\alpha, \dots, \lambda_n^\alpha$ are the eigenvalues of A^α .

Proof: Since A is positive definite its eigenvalues are positive, and $A = UDU^*$ where U is unitary and $D = \text{diag}(\lambda_1, \dots, \lambda_n)$. Since α is real and each $\lambda_i > 0$, $D^\alpha = \text{diag}(\lambda_1^\alpha, \dots, \lambda_n^\alpha)$ is uniquely defined. Then A^α is defined to be $UD^\alpha U^*$. Each λ_i^α is real and positive and A^α is Hermitian, hence it is positive definite, and clearly $(A^\alpha)^{1/\alpha} = A$ when $\alpha \neq 0$.

If $\alpha > 0$ and A is positive semidefinite, some eigenvalues may be zero, but A^α can be defined the same way because $\alpha > 0$, and A^α will be positive semidefinite such that $(A^\alpha)^{1/\alpha} = A$.

Clearly the eigenvalues of A^α are of the form λ_i^α as claimed. $\square$

4.0.7 Definitions. An $n \times n$ matrix A is *irreducible* if $n \geq 2$ and there is no permutation

matrix P such that $P^T A P = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$, where B and D are square. A matrix is *positive*,

denoted $A > 0$ (*nonnegative*, denoted $A \geq 0$), if each $a_{ij} > 0$ (each $a_{ij} \geq 0$).

4.0.8 Perron-Frobenius Theorem [HJ, p. 500]. Let A be $n \times n$. If $A > 0$, or more generally if $A \geq 0$ and A is irreducible, then $\rho(A) > 0$ and $\rho(A)$ is an eigenvalue of A .

APPENDIX B

Rank of a Matrix

4.1.1 Definition. The *rank* of a matrix A is the dimension of the column space of A , that is, $\text{rank}(A) = \dim(\text{Col}(A))$.

4.1.2 Theorem. If A is an invertible $n \times n$ matrix and B is an $n \times n$ matrix then $\text{rank}(AB) = \text{rank}(B)$.

Proof: Let $b_1, \dots, b_n$ be the column vectors of B , so $B = [b_1 \dots b_n]$. By renumbering we may suppose $\{b_1, \dots, b_p\}$ is a basis for $\text{Col}(B)$ where $p \leq n$, so $\text{rank}(B) = p$.

Assume $\sum_{i=1}^p \alpha_i (Ab_i) = 0$ which gives $A \sum_{i=1}^p \alpha_i b_i = 0$. Since A is invertible thus $\sum_{i=1}^p \alpha_i b_i = 0$

and because $\{b_1, \dots, b_p\}$ is linearly independent each α_i is zero. Therefore

$\{Ab_1, \dots, Ab_p\}$ is linearly independent.

Also note for any $j > p$, $b_j = \sum_{i=1}^p \beta_i b_i$ for some scalars β_i . Therefore $Ab_j = A \sum_{i=1}^p \beta_i b_i =$

$\sum_{i=1}^p \beta_i Ab_i$ for $j > p$. Thus $\{Ab_1, \dots, Ab_p\}$ is a basis for $\text{Col}(AB)$, hence $\text{rank}(AB) = p$. $\square$

4.1.3 Theorem. For any matrix A , $\text{rank}(A) = \text{rank}(A^T)$.

Proof: If two matrices A and B are row equivalent, then their row spaces are same [Lay, p. 236]. If B is the row reduced echelon form of A then the nonzero rows of B are linearly independent because no nonzero row is a linear combination of the nonzero rows below it. Thus the nonzero rows of B form a basis for the row space of B and obviously their number is equal to the number of pivot columns in A , which form a basis for $\text{Col}A$. So we have

$$\dim(\text{Row}(A)) = \dim(\text{Row}(B)) = \dim(\text{Col}(A)) = \text{rank}(A).$$

Observe that $\text{Row}(A)$ equals $\text{Col}(A^T)$ and the dimension of $\text{Col}(A^T)$ is the rank of A^T .

Therefore rank of A equals rank of A^T $\square$.

4.1.4 Corollary. If B is an $n \times k$ matrix and A is an invertible $k \times k$ matrix then $\text{rank}(BA) = \text{rank}(B)$.

Proof: By Theorem (4.1.3) we know $\text{rank}(BA)$ equals $\text{rank}(BA)^T$ which equals $\text{rank}(A^T B^T)$. But by Theorem (4.1.2) $\text{rank}(A^T B^T) = \text{rank } B^T$. Again using (4.1.3), $\text{rank}(B)^T = \text{rank}(B)$, thus proving the result.

APPENDIX C

Theorems on Congruence

4.2.1 Definition. If there exists a nonsingular matrix S such that $B = SAS^*$, then we will say B is *congruent* to A . In [HJ] this concept is called star congruence.

4.2.2 Definition [HJ, p. 221]. Let A be a real, symmetric matrix. The *inertia* of A is defined to be the ordered triple $i(A) = (i_+(A), i_-(A), i_0(A))$, where $i_+(A)$ equals the number of positive eigenvalues of A , $i_-(A)$ equals the number of negative eigenvalues of A , and $i_0(A)$ equals the number of zero eigenvalues of A . The *inertia matrix* of A is the diagonal matrix formed by inserting $i_+(A)$ 1's, followed by $i_-(A)$ (-1) 's and then 0's in the diagonal positions.

4.2.3 Theorem [HJ, p.222]. Every real, symmetric matrix A is congruent by a real orthogonal matrix to its *inertia matrix*, which is known once the inertia of A is known.

Proof: Because A is real and symmetric, $A = U\Lambda U^T$ where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ and U is a real, orthogonal matrix [Appendix]. We may assume that the diagonal entries in Λ are arranged as positive, negative and zero; $\lambda_1, \lambda_2, \dots, \lambda_p > 0$, $\lambda_{p+1}, \dots, \lambda_q < 0$, and $\lambda_{q+1} = \dots = \lambda_n = 0$. Then $p = i_+(A)$, $q-p = i_-(A)$ and $n-q = i_0(A)$. If we set

$$D = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_p}, \sqrt{-\lambda_{p+1}}, \dots, \sqrt{-\lambda_q}, 1, \dots, 1)$$

then clearly D is a nonsingular, real, diagonal matrix and Λ can be written as:

$$\Lambda = D \begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_{q-p} & 0 \\ 0 & 0 & 0 \end{bmatrix} D = D I_A D \text{ where } I_A \text{ is the inertia matrix of } A. \text{ Thus}$$

$$A = U \Lambda U^T = U D I_A D U^T = S I_A S^T \text{ where } I_A \text{ is the inertia matrix of } A \text{ and } S = U D.$$

Therefore A is congruent to I_A . $\square$

4.2.4 Theorem [HJ, p.223]. *Sylvester's Law of Inertia*: Let A and B be $n \times n$, real, symmetric matrices. Then A and B are congruent if and only if A and B have the same inertia.

Proof: Let A and B have the same inertia. Then by the previous theorem there exist invertible S_A and S_B so that $A = S_A I_K S_A^T$ and $B = S_B I_K S_B^T$ where $I_K = I_A = I_B$. Then $I_K = S_B^{-1} B (S_B^T)^{-1}$, hence $A = T B T^T$ where $T = S_A S_B^{-1}$ is a nonsingular matrix.

Therefore A and B are congruent to each other. Conversely, suppose A and B are congruent. Then $A = S B S^T$ where S is an $n \times n$ invertible matrix. Congruent matrices have same rank by (4.1.2) and (4.1.4). We know from Theorem 4.2.3 that there exist invertible P and Q and inertia matrices I_A and I_B such that

$$P A P^T = I_A = \begin{bmatrix} I_t & & \\ & -I_{r-t} & \\ & & 0 \end{bmatrix} \text{ and } Q B Q^T = I_B = \begin{bmatrix} I_s & & \\ & -I_{r-s} & \\ & & 0 \end{bmatrix}. \text{ We need to prove that}$$

$s = t$. Let us suppose $s < t$. Note $I_A = P S B S^T P^T = P S Q^{-1} I_B (Q^T)^{-1} S^T P^T = R I_B R^T$ where $R = P S Q^{-1}$.

Let $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{0} \end{bmatrix}$ where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{0} \neq \mathbf{x}_1 \in \mathbb{R}^t$. Then $\mathbf{x}^T \mathbf{I}_A \mathbf{x} = \sum_{i=1}^t |x_i|^2 > 0$. We partition

$\mathbf{R}^T$ as $\mathbf{R}^T = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 \\ \mathbf{R}_3 & \mathbf{R}_4 \end{bmatrix}$, so that $\mathbf{R}_1$ is sxt. Since $s < t$, there exists a free variable in

$\mathbf{R}_1 \mathbf{x}_1 = \mathbf{0}$, so we may choose $\mathbf{x}_1 \neq \mathbf{0}$ in $\mathbb{R}^t$ such that $\mathbf{R}_1 \mathbf{x}_1 = \mathbf{0}$.

Let $\mathbf{y} = \mathbf{R}_3 \mathbf{x}_1 = \begin{bmatrix} y_1 \\ \vdots \\ y_{n-s} \end{bmatrix}$. Then $\mathbf{R}^T \mathbf{x} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 \\ \mathbf{R}_3 & \mathbf{R}_4 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{R}_3 \mathbf{x}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix}$, where $\mathbf{0} \in \mathbb{R}^s$

and $\mathbf{y} \in \mathbb{R}^{n-s}$. Note $\mathbf{y} \neq \mathbf{0}$ since $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{R}^T$ is invertible. So

$$\mathbf{x}^T \mathbf{I}_A \mathbf{x} = \mathbf{x}^T \mathbf{R} \mathbf{I}_B \mathbf{R}^T \mathbf{x} = \begin{bmatrix} \mathbf{0} & \mathbf{y} \end{bmatrix}^T \begin{bmatrix} \mathbf{I}_s & & \\ & -\mathbf{I}_{r-s} & \\ & & 0 \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix} = -\sum_{i=s+1}^r |y_i|^2 < 0. \text{ But this is impossible,}$$

so $s = t$. $\square$

4.2.5 Theorem. Let A and B be $n \times n$ matrices where A is positive definite and B is positive semidefinite. Then BA has nonnegative eigenvalues. If all eigenvalues of BA are zero, then $B = 0$.

Proof: Observe that BA is congruent to $A^{-1/2}BA(A^{-1/2})^*$ and this equals $A^{-1/2}BA^{1/2}$. As this is similar to B , the eigenvalues of BA are nonnegative.

If the eigenvalues of BA are zero then by the similarity, the eigenvalues of B are zero; as B is positive semidefinite, this implies $B = 0$. $\square$

APPENDIX D

Orthoprojections

The purpose of this section is to prove Theorem 4.3.6 below, which is needed in proving the Theorem above. This theorem says in essence that if A and B are orthoprojections on a finite dimensional inner product space V , there is a decomposition of V into mutually orthogonal subspaces which are 1×1 or 2×2 , so that each of these is invariant under both A and B . In [HJS], two references are given for this, [Davis] and [Halmos]. The specific fact we need does follow from those. However, the focus of those papers is to describe the way a pair of subspaces of an infinite dimensional space intersect, so as to know when two different pairs of subspaces are unitarily equivalent. Thus the presentations in [Davis] and [Halmos] are more abstract than we need, and the exact fact we need does not appear in either. So we give a basic constructive proof here.

4.3.1 Definitions and notation. Let V be a finite dimensional inner product space [FIS. Sec. 6.3]. Let $A: V \rightarrow V$ be a linear operator. In [FIS, Theorem 6.9] it is shown that there exists a unique linear operator $A^*: V \rightarrow V$ such that $\langle A(x), y \rangle = \langle x, A^*(y) \rangle$ for all x, y in V . The map A^* is called the *adjoint* of A . An operator A is said to be *self-adjoint* if $A = A^*$.

An *orthoprojection* is a linear operator $A: V \rightarrow V$ such that for some subspace M of V , $A(x) = x$ for each x in M and $A(x) = 0$ for each x in $M^\perp$. This is also called the *orthoprojection*, or *orthogonal projection*, of V on M .

If S is an A -invariant subspace of V , let $A|_S: S \rightarrow S$ denote the restriction of A to S .

If γ is a basis for V , let $[A]_\gamma$ denote the matrix representation of A with respect to γ [FIS, Sec. 2.2].

4.3.2 Theorem. Let V be a finite dimensional complex inner product space and let $A: V \rightarrow V$ be a linear operator. Then the following results are true:

- (a) If γ is an orthonormal basis for V then $[A^*]_\gamma = [A]_\gamma^*$ [FIS, Theorem 6.10].
- (b) The operator A is self-adjoint if and only if there exists an orthonormal basis for V consisting of eigenvectors of A [FIS, Theorem 6.17].
- (c) A is self-adjoint if and only if there exists orthonormal basis γ for V such that $[A]_\gamma$ is Hermitian.
- (d) If S is an A -invariant subspace of V and A is self-adjoint, then $A|_S: S \rightarrow S$ is self-adjoint.

Proof: Proofs of (a) and (b) can be found in the given references.

Proof of (c). If A is self adjoint, by (b) there exists an orthonormal basis for V consisting of eigenvectors of A . Then $[A]_\gamma$ is diagonal and from (a) we have $[A]_\gamma^* = [A^*]_\gamma$ which equals $[A]_\gamma$ since A is self adjoint. So $[A]_\gamma$ is Hermitian. Conversely, suppose γ is an orthonormal basis for V such that $[A]_\gamma$ is Hermitian. Then from (a) we have $[A^*]_\gamma = [A]_\gamma^*$ which equals $[A]_\gamma$ since $[A]_\gamma$ is Hermitian. Thus $[A]_\gamma = [A^*]_\gamma$, so

$A = A^*$, hence A is self-adjoint.

Proof of (d). Suppose $A = A^*$ and S is an A -invariant subspace of V . If x, y are in S , then $\langle A|_S(x), y \rangle = \langle A(x), y \rangle$, which by existence of the adjoint operator equals $\langle x, A^*(y) \rangle$; this equals $\langle x, A(y) \rangle$ because A is self-adjoint, which equals $\langle x, A|_S(y) \rangle$. But $\langle A|_S(x), y \rangle = \langle x, (A|_S)^*(y) \rangle$, and thus $A|_S = (A|_S)^*$ because the adjoint is unique.

4.3.3 Lemma. Let A be an $n \times n$ Hermitian matrix, and let W be a k -dimensional subspace of C^n which is A -invariant. Then $W^\perp$ is A -invariant and there exists an orthonormal basis $\{x_1, \dots, x_n\}$ for C^n of eigenvectors of A such that $\{x_1, \dots, x_k\}$ is a basis for W and $\{x_{k+1}, \dots, x_n\}$ is a basis for $W^\perp$.

Proof: Let $\{y_1, \dots, y_k\}$ be an orthonormal basis for W and $\{y_{k+1}, \dots, y_n\}$ be an orthonormal basis for $W^\perp$ [FIS, Theorem 6.7]. Let $U = [y_1 \ y_2 \ \dots \ y_n]$ so U is orthogonal since it is square with orthonormal columns. Then $U^*AU = U^* [Ay_1 \ \dots \ Ay_k \ \dots \ Ay_n]$. Since W is A -invariant, $Ay_1, \dots, Ay_k$ are in W , hence will be linear combinations of $y_1, \dots, y_k$ only. Thus

$$U^*AU = \begin{bmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_n^* \end{bmatrix} \begin{bmatrix} a_{11}y_1 + \dots + a_{k1}y_k & \dots & a_{1k}y_1 + \dots + a_{kk}y_k & \dots & a_{1n}y_1 + \dots + a_{nn}y_n \end{bmatrix}$$

$$\left[\begin{array}{ccc|ccc} a_{11} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & \dots & a_{2k} & \dots & a_{2n} \\ \vdots & \dots & \vdots & \dots & \vdots \\ a_{k1} & \dots & a_{kk} & \dots & a_{kn} \\ \hline 0 & \dots & 0 & \dots & a_{k+1n} \\ \vdots & \dots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & \dots & a_{nn} \end{array} \right] = \begin{bmatrix} E & F \\ 0 & G \end{bmatrix} \text{ where } E \text{ is } k \times k \text{ and } 0 \text{ is a zero block. Since } A$$

is Hermitian, U^*AU is also. Therefore $F = 0$ and that implies that $Ay_{k+1}, \dots, Ay_n$ are linear combinations of $\{y_{k+1}, \dots, y_n\}$ only, hence $W^\perp$ is also A -invariant.

Since A is Hermitian, then by (4.3.2 (c)), $x \rightarrow Ax$ is a self-adjoint linear operator on C^n which we will also call A . Thus $A|_W: W \rightarrow W$ and $A|_{W^\perp}: W^\perp \rightarrow W^\perp$ are Hermitian operators by (4.3.2 (d)), so by (4.3.2 (b)) there are orthonormal bases $\{x_1, \dots, x_k\}$ for W and $\{x_{k+1}, \dots, x_n\}$ for $W^\perp$, of eigenvectors of $A|_W$ and $A|_{W^\perp}$ respectively. Then $\{x_1, \dots, x_n\}$ is an orthonormal basis for C^n , of eigenvectors of A .

4.3.4 Theorem [Halmos], [Davis]. Let V be a finite dimensional inner product space. Suppose there exist subspaces M and N such that $M \cap N$, $M \cap N^\perp$, $M^\perp \cap N$ and $M^\perp \cap N^\perp$ are all trivial. Such subspaces are said to be *in general position*. Let α, β be nonzero real scalars. We define linear transformations $A, B: V \rightarrow V$ by

$$A(x) = \begin{cases} \alpha x & \text{if } x \in M \\ 0 & \text{if } x \in M^\perp \end{cases}, \quad B(y) = \begin{cases} \beta y & \text{if } y \in N \\ 0 & \text{if } y \in N^\perp \end{cases}$$

Thus if $\alpha = \beta = 1$, A and B are orthoprojections. Also, the following results are true:

(a) $M, N, M^\perp$ and $N^\perp$ have the same dimension.

(b) $V = M \oplus N$, the direct sum of M and N .

(c) A and B are self-adjoint and there exists an orthonormal basis

$\gamma = \{u_1, \dots, u_n\}$ for V such that

$$[A]_\gamma = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \vdots \\ 0 & 0 & \dots & A_p \end{bmatrix} \text{ and } [B]_\gamma = \begin{bmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \dots & \dots & \dots & \vdots \\ 0 & 0 & \dots & B_p \end{bmatrix} \text{ where for each } i,$$

A_i and B_i are both 1×1 or both 2×2 .

Proof: We will give a more constructive proof than that in [Halmos]. First we will show that A is self-adjoint. By [FIS, theorem 6.7], there exists an orthonormal basis

$\gamma = \{u_1, \dots, u_n\}$ for V such that $\{u_1, \dots, u_k\}$ is a basis for M and $\{u_{k+1}, \dots, u_n\}$ is a basis

for $M^\perp$. Then $[A]_\gamma = \begin{bmatrix} \alpha I_k & 0 \\ 0 & 0 \end{bmatrix}$, a Hermitian matrix, hence A is self-adjoint by Theorem

(4.3.2 (c)).

Next observe that $\text{Nul}(A) = M^\perp$ and $\text{Range } A = M$. For if x is in $M^\perp$ then $A(x) = 0$ which implies x is in $\text{Nul}(A)$, so $M^\perp$ is contained in $\text{Nul}(A)$. Conversely, if x is in $\text{Nul}(A)$ then $Ax = 0$. Also, $x = m + m'$ where m is in M and m' is in $M^\perp$ [FIS, Theorem 6.7]. Since A is linear, this gives $0 = Ax = Am + Am' = \alpha m + 0$. As $\alpha \neq 0$, thus $m = 0$ and therefore $x = m'$. This means $\text{Nul}(A)$ is contained in $M^\perp$. Therefore $\text{Nul}(A) = M^\perp$.

To prove that $\text{Range}(A)$ equals M , let x be an element of V ; then $x = m + m'$ as before. Then $Ax = Am + Am' = \alpha m$, which is in M . Thus $\text{Range}(A) \subseteq M$. Conversely,

let m be in M . Then since $\alpha \neq 0$, the element $(1/\alpha)m$ in M maps to m , so m is in $\text{Range}(A)$. So we have $M \subseteq \text{Range}(A)$, hence $\text{Range}(A) = M$.

(a) The map $A|_N$ maps N one-one into M : for $\text{Nul}(A|_N) = N \cap \text{Nul}(A) = N \cap M^\perp = \{0\}$. Thus A is one-one from N to M by [FIS, Theorem 2.4]. By a similar argument we can prove $B|_M$ maps M one-one to N . Then the composition $(A|_N)(B|_M)$ maps M one-one into M , hence maps M onto M because M is finite dimensional. This implies $A|_N$ maps N onto M , hence $A|_N: N \rightarrow M$ is an isomorphism. So M and N have same dimension.

The hypotheses are the same on each pair M and $N^\perp$, $M^\perp$ and N , and $M^\perp$ and $N^\perp$, so all four subspaces have the same dimension.

(b) We have proved $\dim N = \dim M^\perp$, and $M \cap N$ is trivial, so $\dim (M + N) = \dim(M) + \dim(N)$ [FIS, p.54]. Also $\dim(V) = \dim(M) + \dim(M^\perp)$ by [FIS, Theorem.6.7] so $\dim(V) = \dim (M+N)$. Thus $M+N$ is a subspace of V whose dimension is equal to that of V , so $M+N = V$ by [FIS, Theorem 1.11]. Also $M \cap N = \{0\}$, hence $V = M \oplus N$.

(c) Let z be an eigenvector of $A+B$. Then $z \neq 0$ and $(A + B)z = cz$ for some scalar c . We know that V is the direct sum of M and N , hence there exist unique vectors $x \in M$ and $y \in N$ such that $z = x + y$. Then

$$Az = \alpha x + Ay \text{ and } Bz = Bx + \beta y. \quad (\text{i})$$

Therefore $(A + B)z = cz = c(x + y) = \alpha x + Ay + Bx + \beta y$ which implies

$cx - \alpha x - Ay = \beta y - cy + Bx$. Observe that the left side of this equality is in M while the right is in N . Since $M \cap N$ is trivial, both sides equal zero, which gives

$$(c - \alpha)x = Ay \text{ and } (c - \beta)y = Bx. \quad (\text{ii})$$

Define $S = \text{Span } \{x, y\}$. We claim $\dim(S) = 2$. To see this, first suppose $x = 0$. Then from (ii) we get $Ay = 0$, which means y is in $M^\perp$. Thus y is in N and $M^\perp$ and $M^\perp \cap N = \{0\}$ hence $y = 0$. But then $z = 0$ which is not true, so we see that x cannot be zero. In the same way, we can show $y \neq 0$. Then $\{x, y\}$ is linearly independent since $ax + by = 0$ implies $ax = -by$, ax is in M and $-by$ is in N , hence $ax = -by = 0$; since $x, y \neq 0$ we have $a, b = 0$. Therefore $\dim(S) = 2$ is true.

We also claim S is invariant under A and B and $A + B$. For $Ax = \alpha x$ and $Ay = (c - \alpha)x$. By (ii), S is invariant under A . In the same way, S is invariant under B , hence under $A + B$.

Since A and B are self-adjoint so is $A + B$, hence $A|_S, B|_S$ and $A|_S + B|_S$ are self-adjoint by (4.3.2(d)). Also by (4.3.3), there exist orthonormal eigenvectors u_1, u_2 for $A + B$ which form a basis for S , and orthonormal eigenvectors $u_3, \dots, u_n$ which form a basis for $S^\perp$. Then $\{u_1, \dots, u_n\}$ is an orthonormal basis for V [FIS, Theorem 6.7]. If V is two dimensional we are done as $S = V$. If not, suppose that two dimensional subspaces $S_1, \dots, S_k$ have been found which are mutually orthogonal and invariant under A and B . Then $S_1 + \dots + S_k$ is invariant under A and B , and if $S_1 + \dots + S_k = V$ we are done. If not, $(S_1 + \dots + S_k)^\perp$ is nontrivial and invariant under A and B by (4.3.3), hence under $A + B$. Thus there exists an eigenvector z of $A + B$ in $(S_1 + \dots + S_k)^\perp$ and by the argument above, z can be used to construct a new two dimensional subspace S_{k+1} which is invariant under A and B . This subspace will be in $(S_1 + \dots + S_k)^\perp$ so it is orthogonal to $S_1, \dots, S_k$. Continue in this way until we have two dimensional subspaces $S_1, \dots, S_m$ which are

mutually orthogonal, invariant under A and B, and which satisfy $S_1 + \dots + S_m = V$. Let $\{u_0, v_0\}, \{u_1, v_1\}, \dots, \{u_m, v_m\}$ be orthonormal bases for $S, S_1, \dots, S_m$ respectively. Then $\gamma = \{u_0, v_0, u_1, v_1, \dots, u_m, v_m\}$ is an orthonormal basis for V and

$$[A]_\gamma = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \vdots \\ 0 & 0 & \dots & A_M \end{bmatrix} \text{ and } [B]_\gamma = \begin{bmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \dots & \dots & \dots & \vdots \\ 0 & 0 & \dots & B_M \end{bmatrix}$$

where for each i , A_i and B_i are 2×2 .

4.3.5 Lemma. Let U and W be subspaces of an inner product space V . Then

$$U^\perp \cap W^\perp = (U + W)^\perp.$$

Proof: Since $U \subseteq (U + W)$ so $U^\perp \supseteq (U + W)^\perp$. Similarly $W^\perp \supseteq (U + W)^\perp$, hence

$(U + W)^\perp$ is contained in $U^\perp \cap W^\perp$. Conversely, let $x \in U^\perp \cap W^\perp$. Then for any $u \in U$ and $w \in W$, we have $\langle u + w, x \rangle = \langle u, x \rangle + \langle w, x \rangle = 0 + 0 = 0$. Since $u + w \in U + W$, therefore $x \in (U + W)^\perp$. Thus $U^\perp \cap W^\perp \subseteq (U + W)^\perp$. Therefore $U^\perp \cap W^\perp = (U + W)^\perp \square$.

Two references are given in [HJS] for the following theorem in [HJS, Theorem 5], but we have written our own constructive proof.

4.3.6 Theorem [Halmos], [Davis]. Let M and N be any two subspaces of an inner product space V , not necessarily in general position. Let A and B be the orthoprojections of V on M and N respectively. That is,

$$A(x) = \begin{cases} x & \text{if } x \in M \\ 0 & \text{if } x \in M^\perp \end{cases}, B(y) = \begin{cases} y & \text{if } y \in N \\ 0 & \text{if } y \in N^\perp \end{cases}.$$

Then there exists an orthonormal basis γ for V such that

$$[A]_\gamma = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \vdots \\ 0 & 0 & \dots & A_p \end{bmatrix}, \quad [B]_\gamma = \begin{bmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \dots & \dots & \dots & \vdots \\ 0 & 0 & \dots & B_p \end{bmatrix}$$

and for each i , A_i and B_i are both 1×1 or both 2×2 .

Proof: Since $M + N$ is a subspace of V , we have $V = (M + N) \oplus (M + N)^\perp$. As shown in the proof of (4.3.4), taking $\alpha, \beta = 1$, A and B are self-adjoint.

Observe that $(M + N)^\perp = M^\perp \cap N^\perp$ by (4.3.5), and every vector in $M^\perp \cap N^\perp$ is mapped to zero by both A and B , so we may choose any orthonormal basis η for $(M + N)^\perp$ and each vector in η will be an eigenvector of both A and B .

Now consider $M + N$. Henceforth the “ $\perp$ ” notation, as $M^\perp, N^\perp$ etc., will mean orthogonal complement in $M + N$. First observe that since $(M + N)^\perp$ is invariant under both of the self adjoint operators A and B , by Lemma (4.3.3), $((M + N)^\perp)^\perp = M + N$ is also invariant under both.

We proceed in steps.

- (a) Since $M \cap N$ is contained in both M and N , it is orthogonal to $M^\perp \cap N$ (with elements of $M^\perp$) and $M \cap N^\perp$ (with elements of $N^\perp$). Similarly, $M^\perp \cap N$ and

$M \cap N^\perp$ are orthogonal to each other. Thus $M \cap N$, $M^\perp \cap N$ and $M \cap N^\perp$ are mutually orthogonal subspaces of $M + N$ and it is easy to see that each of them is invariant under both A and B .

- (b) Let β_1 , β_2 and β_3 be orthonormal bases for $M \cap N$, $M^\perp \cap N$ and $M \cap N^\perp$ respectively. Then from (4.3.3) each vector in $\beta = \beta_1 \cup \beta_2 \cup \beta_3$ is an eigenvector of both A and B .
- (c) Let $S = (M \cap N) + (M^\perp \cap N) + (M \cap N^\perp)$, so β is an orthonormal basis for S . Thus S is invariant under A and B . Since A and B are self-adjoint maps, $S^\perp$ is also A , B invariant by Lemma (4.3.3).
- (d) Extend $\beta_1 \cup \beta_3$ to an orthonormal basis for M , $\beta_1 \cup \beta_3 \cup \beta_M$. Since β_M consists of elements of M and β_2 is in $M^\perp$, β_M is also orthogonal to β_2 . Thus β_M is orthogonal to S . Similarly, extend $\beta_1 \cup \beta_2$ to an orthonormal basis for N , $\beta_1 \cup \beta_2 \cup \beta_N$ and notice β_N is also orthogonal to β_3 , hence to S .
- (e) We prove β_1 , β_2 , β_3 , β_M and β_N are mutually disjoint sets. The sets β_M and β_N need not be orthogonal to each other. In order to prove they are mutually disjoint, suppose $x \in \beta_M \cap \beta_N$. Then $x \in M \cap N \subseteq S$ and we know β_M is orthogonal to S , so $x \in S \cap S^\perp = \{0\}$; but x is in the linearly independent set β_M , so cannot be zero, hence x cannot exist and $\beta_M \cap \beta_N$ is empty. All the other pairs β_i , β_j are independent sets which are orthogonal to each other hence have no elements in common.

(f) Let $\gamma = \beta \cup \beta_M \cup \beta_N$. We claim γ is a basis for $M + N$. Let $|\alpha|$ denote the cardinality of a set α . First, note $\dim(M + N) = \dim(M) + \dim(N) - \dim(M \cap N)$ [FIS, p.54], and the latter equals

$$(|\beta_1| + |\beta_3| + |\beta_M|) + (|\beta_1| + |\beta_2| + |\beta_N|) - |\beta_1| = |\beta_1| + |\beta_2| + |\beta_3| + |\beta_M| + |\beta_N| = |\gamma|.$$

Thus the number of elements in γ equals the dimension of $M + N$, and γ spans $M + N$ since its span contains M and N , hence γ must be linearly independent by [FIS, Theorem 1.11].

(g) Let $M_1 = \text{Span}(\beta_M)$ and $N_1 = \text{Span}(\beta_N)$. We claim that $(M_1 + N_1)^\perp = S$. First notice $\beta_M \cup \beta_N$ is independent by (f) and disjoint by (e), so this is a basis for $M_1 + N_1$. Also, $S \subseteq M_1^\perp$ because β is orthogonal to β_M . Similarly $S \subseteq N_1^\perp$. Thus $S \subseteq M_1^\perp \cap N_1^\perp = (M_1 + N_1)^\perp$ by (4.3.5).

Conversely, since $M_1 + N_1$ is a subspace of $M + N$, we have

$$\begin{aligned} \dim(M_1 + N_1)^\perp &= \dim(M + N) - \dim(M_1 + N_1) \\ &= (|\beta| + |\beta_M| + |\beta_N|) - (|\beta_M| + |\beta_N|) = |\beta|. \end{aligned}$$

So S is a subspace of $(M_1 + N_1)^\perp$ with the same dimension, hence S must equal $(M_1 + N_1)^\perp$ by [FIS, Theorem 1.11].

(h) By (4.3.3) $M_1 + N_1$ is invariant under A and B because its orthogonal complement S is.

(i) We now claim that in the subspace $M_1 + N_1$, M_1 and N_1 are in general position.

Let M_1' and N_1' denote the orthogonal complements of M_1 and N_1 in $M_1 + N_1$.

We must verify each intersection:

1. $M_1 \cap N_1 = \{0\}$ because β_M and β_N are disjoint and their union is independent by (e) and (f).
2. $M_1 \cap N_1' = \{0\}$. For suppose $x \in M_1 \cap N_1'$. Then $x \in N_1'$ implies x is orthogonal to β_N ; and $x \in M_1 = \text{Span}(\beta_M)$ implies x is orthogonal to $\beta_1 \cup \beta_2$, by (d). Therefore x is orthogonal to N . So $x \in M_1 \cap N^\perp = M \cap N^\perp = S$; therefore $x \in M_1 \cap S = \{0\}$, hence $x = 0$.
3. $M_1' \cap N_1 = \{0\}$ is proved in the same way.
4. $M_1' \cap N_1' = \{0\}$. For $M_1' \cap N_1'$ is the set of vectors in $M_1 + N_1$ which are orthogonal to both M_1 and N_1 , and this is contained in $M_1^\perp \cap N_1^\perp$, the set of vectors in $M + N$ which are orthogonal to both M_1 and N_1 ; by (4.3.5) $M_1^\perp \cap N_1^\perp = (M_1 + N_1)^\perp$ which equals S by (g). Thus $M_1' \cap N_1' \subseteq (M_1 \cap N_1) \cap S = \{0\}$.

We have shown that M_1 and N_1 are in general position in $M_1 + N_1$, hence by (4.3.4) there is an orthonormal basis $\alpha = \{u_1, v_1, u_2, v_2, \dots, u_k, v_k\}$ for $M_1 + N_1$ such that for each i , $S_i = \text{Span}\{u_i, v_i\}$ is invariant under both A and B . We have also shown that $M_1 + N_1$ and S are mutually orthogonal subspaces and their sum is $M + N$. Thus $\alpha \cup \beta$ is an orthonormal basis for $M + N$, and then $\gamma = \alpha \cup \beta \cup \eta$ is an orthonormal basis for V such that $[A]_\gamma$ and $[B]_\gamma$ have the desired form. $\square$