Steen A. Andersson Hans K. Brøns
Søren Tolver Jensen

Distribution of Eigenvalues in Multivariate Statistical Analysis

Preprint
February
1982

2

Institute of Mathematical Statistics
University of Copenhagen

Steen A. Andersson*, Hans K. Brøns*
and Søren Tolver Jensen*

DISTRIBUTION OF EIGENVALUES IN MULTIVARIATE STATISTICAL ANSLYSIS

Preprint 1982 No. 2

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

February 1982

* The authors are Professors at the Institute of Mathematical Statistics of the University of Copenhagen. A portion of Andersson's work was done while visiting the Department of Statistics, University of Washington.

DISTRIBUTION OF EIGENVALUES
IN MULTIVARIATE STATISTICAL ANALYSIS

by

Steen A. Andersson†, Hans K. Brøns† and Søren Tolver Jensen†

ABSTRACT

Ten invariant multivariate testing problems involving the real, complex, or quaternion structure of covariance matrices are considered. In each problem the the maximal invariant statistic and its distribution are described, as well as the maximum likelihood estimators and likelihood ratio test statistics. These results are obtained by means of a new, unified method based on invariance arguments.

†The authors are Professors at the Institute of Mathematical Statistics of the University of Copenhagen. A portion of Andersson's work was done while visiting the Department of Statistics, University of Washington.

1. INTRODUCTION.

In this paper we consider ten testing problems in multivariate analysis. These are the problems of testing that (a) a covariance matrix has complex structure; (b) a covariance matrix with complex structure has real structure; (c) a covariance matrix with complex structure has quaternion structure; (d) a covariance matrix with quaternion structure has complex structure; (e) two sets of variates are independent, when their joint covariance matrix has real, complex, or quaternion structure; (f) two covariance matrices are identical, when both have real, complex, or quaternion structure. See Andersson [1] for a definition of these structures.

Some of these ten problems have been treated in the literature while the others are new. Together they occur in a fundamental way in a general algebraic theory of normal statistical models developed by Andersson, Brøns, and Jensen [5]. The class of normal models considered in [5] includes most of the structured families of covariance matrices which have appeared in the literature. It is shown in [5] that every testing problem where both the null hypothesis H_0 and the alternative hypothesis $H_1 (H_0 \subset H_1)$ belong to this class of models can be decomposed into simpler problems, each of which has the form of one of the ten problems described in (a) - (f) above.

In the present paper these ten testing problems are treated in detail in a unified manner suggested by the general theory. Each problem is invariant under a group of linear transformations, and our main aim is to obtain a concrete representation of the orbit projection, i.e., the maximal invariant statistic, and to find a representation of its distribution in terms of a density with respect to the Lebesgue measure. Furthermore, the maximum likelihood estimators and likelihood ratio test statistics are obtained and their distributions discussed.

For each problem the representation of the orbit projection is obtained using results on simultaneous reduction of certain types of forms on vector spaces, i.e., simultaneous diagonalization of matrices. The orbit projection is thus represented in terms of eigenvalues of matrices with certain structures. Many of these results on simultaneous reduction of forms are well-known and all but one occur in Bourbaki [9]. The exception is apparently new and is given here as Lemma 8.

Traditionally, distributions in multivariate analysis are derived by calculations involving Jacobians of high dimensions. To use this method would require laborious calculations for each of the ten testing problems and, to make the proofs rigorous, it would be necessary to use arguments from differential geometry. In this paper a new method for obtaining the distribution of eigenvalues is presented in Lemmas 4, 5 and 6. This method, based on actions of a set of certain polynomial transformations on a space of positive definite symmetric matrices and on the space of eigenvalues, has several advantages over the classical approach. The method once presented for one problem, is readily adapted to all other problems. It is elementary, requiring only the calculation of 2×2 Jacobians at most. Lastly, the method is rigorous, not avoiding difficulties involving null sets of multiple eigenvalues.

In the existing literature about the complex normal and complex Wishart distribution one usually treats complex variates. It is important to note that in this paper all matrices considered have real elements; by "matrices with complex or quarternion structure" we refer to real matrices with certain additional structures defined by the corresponding complex and quarternion matrices.

The ten testing problems are treated in Sections 2-7. In each of these sections the joint density of the eigenvalues is derived up to an unspecified norming constant. The exact values of all norming constants are derived simultaneously in Section 8 using a new method involving recursion formulae. The moments of each likelihood ratio test statistic are readily obtained from these norming constants.

For the real case, the distribution of eigenvalues arising in problems (e) and (f) was obtained in 1939 by Hsu [14], Fisher [12], and Roy [19], and are treated in detail in Anderson [8]. For the complex case, the parallel results in (e) and (f) can be obtained from work of Khatri [15]. For the quaternion case these two results appear in the thesis of Gabrielsen [13] written under the supervision of S. Tolver Jensen, obtained by methods similar to Khatri's.

Problem (b) was first treated by Khatri [16], while problem (a) first appeared in Andersson [3], which also contains a statistical interpretation of the eigenvalues and eigenvectors. Further results on problems (a), (b), (c), and (d), including a study of the noncentral distributions, are given in Andersson and Perlman [6], [7].

Related problems concerning distributions of eigenvalues have been studied by mathematicians and physicists in the context of statistical mechanics; c.f. Mehta [17] and Porter [18].

2. TESTING THE HYPOTHESIS THAT A COVARIANCE MATRIX HAS COMPLEX STRUCTURE.

Let E be a p-dimensional vector space over the field $\mathbb{C}$ of complex numbers. By restricting the scalar multiplication to the subfield $\mathbb{R}$ of real numbers E is also a 2p-dimensional vector space over $\mathbb{R}$, and if $e_1,\ldots,e_p$ is a basis for E as a vector space over $\mathbb{C}$ then

(1) $$e_1,\ldots,e_p,\ ie_1,\ldots,ie_p$$

becomes a basis for E as a vector space over $\mathbb{R}$. If f is a $\mathbb{C}$-linear map of E into E with matrix $A+iB$ w.r.t. $e_1,\ldots,e_p$ then f considered as a $\mathbb{R}$-linear map has matrix

(2) $$\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

w.r.t. (1). Since composition of linear maps corresponds to multiplication of matrices it is seen that the set $GL(p, \mathbb{C})$ of nonsingular $2p \times 2p$ matrices of the form (2) for $q = 2p$ is a subgroup of the group $GL(q, \mathbb{R})$ of all nonsingular $q \times q$ matrices.

Let $\phi\colon E \times E \to \mathbb{C}$ be a hermitian left sesquilinear or symmetric bilinear complex form on E (see [9] or [1]) and $C + iD = (\phi(e_\alpha, e_\beta))$ the matrix of ϕ w.r.t. $e_1,\ldots, e_p$. Then $\delta = \mathrm{Re} \circ \phi\colon E \times E \to \mathbb{R}$, where Re denotes the real part of a complex number, is a symmetric bilinear real form on E.

The form ϕ is hermitian left sesquilinear if and only if $C + iD$ is hermitian, i.e., C is symmetric and D is antisymmetric. In this case the matrix of δ w.r.t. (1) is

$$(3) \qquad \begin{pmatrix} C & -D \\ D & C \end{pmatrix} .$$

Moreover, $\phi \circ (f \times f)$ is also hermitian left sesquilinear and the matrix of $\mathrm{Re} \circ (\phi \circ (f \times f)) = (\mathrm{Re} \circ \phi) \circ (f \times f)$ w.r.t. (1) is

$$(4) \qquad \begin{pmatrix} A & -B \\ B & A \end{pmatrix}' \begin{pmatrix} C & -D \\ D & C \end{pmatrix} \begin{pmatrix} A & -B \\ B & A \end{pmatrix} .$$

Since ϕ is positive definite if and only if $\mathrm{Re} \circ \phi$ is positive definite it follows that the action

$$(5) \qquad \begin{aligned} GL(p, \mathbb{C}) \times H^+(p, \mathbb{C}) &\to H^+(p, \mathbb{C}) \\ (M, T) &\to MTM' , \end{aligned}$$

where $H^+(p, \mathbb{C})$ is the set of positive definite matrices of the form (3), is well defined. It follows from the first equation in (14) below that this action is transitive. Furthermore, for $q = 2p$ the action (5) is a restriction of the transitive action

$$(6) \qquad \begin{aligned} GL(q, \mathbb{R}) \times H^+(q, \mathbb{R}) &\to H^+(q, \mathbb{R}) \\ (M, S) &\to MSM' , \end{aligned}$$

where $H^+(q, \mathbb{R})$ is the set of all positive definite symmetric $q \times q$ matrices. Since the action (6) is proper ([10], p. III.27) it follows that the action (5) is also proper.

The form ϕ is symmetric bilinear if and only if $C + iD$ is symmetric, i.e., C and D are symmetric. In this case the matrix of δ w.r.t. the basis (1) is

$$(7) \qquad \begin{pmatrix} C & -D \\ -D & -C \end{pmatrix} .$$

Moreover, $\phi \circ (f \times f)$ is also symmetric bilinear and the matrix of $\mathrm{Re} \circ (\phi \circ (f \times f)) = (\mathrm{Re} \circ \phi) \circ (f \times f)$ w.r.t. (1) is

$$(8) \qquad \begin{pmatrix} A & -B \\ B & A \end{pmatrix}' \begin{pmatrix} C & -D \\ -D & -C \end{pmatrix} \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \in \mathcal{S}(p, \mathbb{C})$$

where $\mathcal{S}(p, \mathbb{C})$ is the set of matrices of the form (7).

We are now ready to consider the first statistical problem. Let $x_1, \ldots, x_N$, $N \geq 2p$, be independently distributed observations from a normal distribution on $\mathbb{R}^{2p}$ with mean vector 0 and unknown covariance matrix $\Sigma \in H^+(2p, \mathbb{R})$. A minimal sufficient statistic is the empirical covariance matrix $S = \frac{1}{N} \sum x_n x_n'$, the maximum likelihood estimator, which follows a Wishart distribution on $H^+(2p, \mathbb{R})$ with N degrees of freedom and parameter $\frac{1}{N} \Sigma$. This distribution has the density

$$(9) \qquad \left(\frac{\det S}{\det \Sigma} \right)^{N/2} \exp(-N/2(\mathrm{tr}\Sigma^{-1}S)), \; S \in H^+(2p, \mathbb{R})$$

w.r.t. a measure $\nu \equiv \nu_{\mathbb{R},2p,N}$ on $H^+(2p, \mathbb{R})$, which is invariant under the action (6). Since this action is transitive ν is uniquely determined by the condition that the integral of (9) is 1.

Let H_0 denote the hypothesis that $\Sigma \in H^+(p, \mathbb{C})$, i.e., that Σ has complex structure. The statistical problem of testing H_0 is invariant under the restriction of the action (6) to the subgroup $GL(p, \mathbb{C})$. Every invariant test statistic has a unique factorization through the orbit projection

$$\Pi: H^+(2p, \mathbb{R}) \to H^+(2p, \mathbb{R})/GL(p, \mathbb{C}), \tag{10}$$

where the right hand side denotes the set of orbits. The main problem is to find a representation of (10) as a function into $\mathbb{R}^\ell$ for some ℓ and, when H_0 is true, to represent the distribution of Π by a density w.r.t. a Lebesgue measure on $\mathbb{R}^\ell$.

Let

$$J \equiv J_p = \begin{pmatrix} 0 & -I_p \\ I_p & 0 \end{pmatrix}, \tag{11}$$

where I_p is the $p \times p$ identity matrix. It is seen that the linear map

$$t: H^+(2p, \mathbb{R}) \to H^+(p, \mathbb{C}) \tag{12}$$

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \longrightarrow 1/2 \begin{pmatrix} S_{11} + S_{22} & S_{12} - S_{21} \\ S_{21} - S_{12} & S_{11} + S_{22} \end{pmatrix}$$

is well defined, because $t(S) = \frac{1}{2}(S + JSJ') \in H^+(p, \mathbb{C})$. Since J is the matrix (w.r.t. the basis (1)) for scalar multiplication by i, it follows that J commutes with all matrices of the form (2), and thus also that t commutes with the actions of $GL(p, \mathbb{C})$. Moreover, the residual

$$(13) \qquad S - t(S) = 1/2 \begin{pmatrix} S_{11} - S_{22} & S_{12} + S_{21} \\ S_{21} + S_{12} & S_{22} - S_{11} \end{pmatrix}$$

has the form (7).

<u>Lemma 1</u>. Let ϕ be a positive definite hermitian left sesquilinear form and ψ a symmetric bilinear form on E. Then there exists a basis for E such that the matrices of ϕ and ψ are respectively the identity matrix I_p and a diagonal matrix

$$\Lambda = \begin{pmatrix} \lambda_1 & & o \\ & \ddots & \\ o & & \lambda_p \end{pmatrix} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p),$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$.

<u>Proof</u>. Bourbaki [9], p. 123. $\square$

An equivalent formulation of the lemma is that there exists a $\mathbb{C}$-linear map $f: E \to E$ such that $\phi \circ (f \times f)$ and $\psi \circ (f \times f)$ have matrices I_p and Λ resp. w.r.t. the original basis $e_1, \ldots, e_p$. Since the matrices for $\mathrm{Re} \circ \phi$ and $\mathrm{Re} \circ \psi$ transform according to (4) and (8) the lemma also has an equivalent formulation in terms of $2p \times 2p$ real matrices. Let $T \in H^+(p, \mathbb{C})$ and $R \in S(p, \mathbb{C})$. Then there exists an $M \in GL(p, \mathbb{C})$ such that

$$(14) \qquad MTM' = I_{2p} \quad \text{and} \quad MRM' = \begin{pmatrix} \Lambda & o \\ o & -\Lambda \end{pmatrix}.$$

It is seen from (14) that $\pm\lambda_1,\ldots,\pm\lambda_p$ are uniquely determined as the eigenvalues of R w.r.t. T, i.e., the solutions to the equation

$$\det(R - \lambda T) = 0. \tag{15}$$

From (14) and (15) with $T = t(S)$ and $R = S - t(S)$ and the fact that $S = t(S) + (S - t(S))$ is positive definite, it then follows that there exists an $M \varepsilon GL(p, \mathbb{C})$ such that

$$MSM' = \begin{pmatrix} I_p + \Lambda & 0 \\ 0 & I_p - \Lambda \end{pmatrix} \tag{16}$$

and that (10) can be represented by

$$\Pi : H^+(2p, \mathbb{R}) \to \Lambda_p, \tag{17}$$

where $\Lambda_p = \{(\lambda_1,\ldots,\lambda_p) \varepsilon \mathbb{R}^p | 1 > \lambda_1 \geq \ldots \geq \lambda_p \geq 0\}$ and $\Pi(S)$ is the ordered family of nonnegative eigenvalues of $S - t(S)$ w.r.t. $t(S)$.

That the function Π in (17) is continuous follows from

Lemma 2. The ordered family of eigenvalues of a symmetric $m \times m$ real matrix R w.r.t. a positive definite symmetric $m \times m$ real matrix T depends continuously on (T,R).

Proof. Let (T_n, R_n) be a sequence of pairs of a positive definite symmetric $m \times m$ real matrix and a symmetric $m \times m$ real matrix such that $(T_n, R_n) \to (T,R)$. One has to show that the ordered family $(\lambda_{1n},\ldots,\lambda_{mn})$ of solutions to the equation

$$\det(R_n - \lambda T_n) = a_n \prod_{i=1}^{m} (\lambda - \lambda_{in}) = 0$$

converges to the ordered family $(\lambda_1,\ldots,\lambda_m)$ of solutions to the equation

$$\det(R - \lambda T) = a \prod_{i=1}^{m} (\lambda - \lambda_i) = 0.$$

Since the sequence (T_n, R_n) is bounded, the sequence $(\lambda_{1n}, \ldots \lambda_{mn})$ is also bounded. It is therefore enough to show that every convergent subsequence $(\lambda_{1n'}, \ldots, \lambda_{mn'})$ converges to $(\lambda_1, \ldots, \lambda_m)$. But if $(\lambda_{1n'}, \ldots, \lambda_{mn'}) \to (\mu_1, \ldots, \mu_m)$ it follows that $a_n \Pi(\lambda - \lambda_{in'})$ converges to both $a\Pi(\lambda - \lambda_i)$ and $a\Pi(\lambda - \mu_i)$, and the uniqueness of the roots of a polynomial gives that $(\mu_1, \ldots, \mu_m) = (\lambda_1, \ldots, \lambda_m)$. $\square$

Remark. Since the action of $GL(p, \mathbb{C})$ on $H^+(2p, \mathbb{R})$ is proper it is known that the final topology induced by Π on $H^+(2p, \mathbb{R})/GL(p, \mathbb{C})$ is locally compact ([11], p. 39). Since the right hand side of (16) depends continuously on $(\lambda_1, \ldots, \lambda_p) \in \Lambda_p$, the lemma above shows in fact that the representation of (10) by (17) is also topological, i.e., the identification of $H^+(2p, \mathbb{R})/GL(p, \mathbb{C})$ with Λ_p is a homeomorphism.

Theorem 1. The maximum likelihood estimator of Σ under H_0 is $t(S)$ and the likelihood ratio statistic for testing H_0 is

$$\prod_{\gamma=1}^{p} (1 - \lambda_\gamma^2)^{N/2}, \tag{18}$$

where $(\lambda_1, \ldots, \lambda_p) = \Pi(S)$. Under the hypothesis H_0 the statistics $t(S)$ and $\Pi(S)$ are independently distributed. The distribution of $t(S)$ has density

$$\left(\frac{\det T}{\det \Sigma}\right)^{N/2} \exp\left(-N/2(\operatorname{tr}\Sigma^{-1} T)\right), \; T \in H^+(p, \mathbb{C}) \tag{19}$$

w.r.t. a unique measure $\nu_{\mathbb{C},p,N}$ which is invariant under the action (5). The distribution of $\Pi(S)$ has density (18) w.r.t. a measure κ on Λ_p which is uniquely defined by

(20) $$(t,\Pi)(\nu_{\mathbb{R},2p,N}) = \nu_{\mathbb{C},p,N} \otimes \kappa .$$

Furthermore, κ has the density

(21) $$\prod_{1 \le \alpha < \beta \le p}(\lambda_\alpha^2 - \lambda_\beta^2) \prod_{\gamma=1}^{p} \lambda_\gamma (1-\lambda_\gamma^2)^{-\frac{2p+1}{2}}$$

w.r.t. a Lebesgue measure on Λ_p.

Remark. The distribution given by the density (19) is the complex Wishart distribution with N degrees of freedom and parameter $\frac{1}{N}\Sigma \in H^+(p, \mathbb{C})$.

Proof. The density of S is given by (9). Since $GL(p, \mathbb{C})$ is a group it follows that $\Sigma^{-1} \in H^+(p, \mathbb{C})$ when $\Sigma \in H^+(p, \mathbb{C})$, and a direct calculation shows that $\mathrm{tr}\Sigma^{-1}S = \mathrm{tr}\Sigma^{-1}t(S)$. Since t commutes with the action of $GL(p, \mathbb{C})$, the function $\det t(S)/\det S$, $S \in H^+(2p, \mathbb{R})$, is invariant. Using the representation (16) the density (9) can be rewritten as

(22) $$\prod_{\gamma=1}^{p}(1-\lambda_\gamma)^{N/2} \left(\frac{\det T}{\det \Sigma}\right)^{N/2} \exp(-N/2 \ \mathrm{tr}(\Sigma^{-1}T)),$$

where $T = t(S)$ and $(\lambda_1,\ldots,\lambda_p) = \Pi(S)$. The first sentence of the theorem follows from (22).

Next, the distribution of $(t,\Pi)(S) \equiv (t(S),\Pi(S))$ has the density (22), where $T \in H^+(p, \mathbb{C})$ and $(\lambda_1,\ldots,\lambda_p) \in \Lambda_p$, w.r.t. $(t,\Pi)(\nu_{\mathbb{R},2p,N})$. Since the action (5) is transitive, $\nu_{\mathbb{C},p,N}$ is uniquely determined by the condition that the integral of (19) must be 1. The theorem now follows from Lemmas 3 and 5 below.

□

Lemma 3. Let G be a locally compact group, which acts properly on a locally compact space X and properly and transitively on a locally compact space Y. Let furthermore $t:X \to Y$ be a continuous map which commutes with the actions of G, and let $\Pi:X \to X/G$ denote the orbit projection. Then the map (t,Π) is proper. If ν is an invariant measure on X and ν_0 is an invariant measure on Y then there exists a unique measure κ on the locally compact space X/G such that $(t,\Pi)(\nu) = \nu_0 \otimes \kappa$.

Proof. (Compare Bourbaki [11], p. 39) Since the action of G on X is proper the final topology on X/G is locally compact and every compact subset of X/G has the form $\Pi(K)$, where $K \subseteq X$ is compact. Thus every compact subset of $Y \times X/G$ is contained in a compact subset of the form $L \times \Pi(K)$, where $L \subseteq Y$ is compact. One has to show that $(t,\Pi)^{-1}(L \times \Pi(K)) \subseteq X$ is compact. Since the action of G on Y is proper, the set $\{g \in G \mid \exists z \in K: gt(z) \in L\} = \{g \in G \mid gt(K) \cap L \neq \emptyset\} \equiv P(t(K),L)$ is compact, (Bourbaki [10], p. III.33). Thus one has $(t,\Pi)^{-1}(L \times \Pi(K)) =$ $\{x \in X \mid t(x) \in L,\ \Pi(x) \in \Pi(K)\} = \{x \in X \mid t(x) \in L,\ \exists z \in K\ \exists g \in G:\ x = gz\} =$ $\{gz \in X \mid z \in K,\ g \in G:\ gt(z) \in L\} \subseteq P(t(K), L)K$, which is compact.

To prove the second assertion, let h and f be non-negative continuous functions with compact supports on Y and X/G respectively. Since t commutes with the actions it is seen that the positive linear map $h \to \int h(t(x))\ f(\Pi(x))\ d\nu(x)$ defines an invariant measure on Y. Because of the uniqueness of an invariant measure on Y there exists a non-negative constant $\kappa(f)$ such that $\nu_0(h)\kappa(f) = \int h(t(x))f(\Pi(x))d\nu(x)$. It is easy to see that κ is a positive linear map, so that κ is a measure on X/G. $\square$

All assertions in Theorem 1 have now been proved except for (21). To prove this we have developed a new method, based on invariance under polynomial transformations, which also can be applied to the eigenvalue problems in the following sections.

Since $H^+(p, \mathbb{C}) \cap S(p, \mathbb{C}) = \emptyset$ and $t(T) = T$ for $T \varepsilon H^+(p, \mathbb{C})$, it follows that the linear map

$$(23) \qquad H^+(2p, \mathbb{R}) \to E$$
$$S \to (t(S), S - t(S)),$$

where $E = \{(T,R) \varepsilon H^+(p, \mathbb{C}) \times S(p, \mathbb{C}) \mid T + R \varepsilon H^+(2p, \mathbb{R})\}$, is well-defined, one-to-one, and onto. The fact that t commutes with the actions of $GL(p, \mathbb{C})$ gives that (23) also commutes with the actions of $GL(p, \mathbb{C})$ on $H^+(2p, \mathbb{R})$ and E, the latter given by

$$(24) \qquad GL(p, \mathbb{C}) \times E \to E$$
$$(M, (T,R)) \to (MTM', MRM'),$$

which is well-defined because of (8). Let $r(x) = \sum_{j=o}^{m} a_j x^{2j+1}$ be an odd real polynomial such that $r(0) = 0$, $r(1) = 1$, and $Dr(x) > 0$, $0 \leq x \leq 1$. Let

$$(25) \qquad r(T,R) = \sum_{j=o}^{m} a_j (RT^{-1})^{2j} R, \ (T,R) \varepsilon E,$$

and

$$(26) \qquad \tilde{r}(S) = t(S) + r(t(S), S - t(S)), \ S \varepsilon H^+(2p, \mathbb{R}).$$

<u>Lemma 4</u>. The map $\tilde{r}$ defined by (26) is a diffeomorphism of $H^+(2p, \mathbb{R})$ onto $H^+(2p, \mathbb{R})$ which commutes with the action of $GL(p, \mathbb{C})$. Moreover,

$$(27) \qquad t \circ \tilde{r} = t \quad \text{and} \quad \Pi \circ \tilde{r} = \bar{r} \circ \Pi,$$

where $\bar{r}$ is the homeomorphism of Λ_p onto Λ_p given by

$$(28) \qquad \bar{r}(\lambda_1,\ldots,\lambda_p) = (r(\lambda_1),\ldots,r(\lambda_p)),\ (\lambda_1,\ldots,\lambda_p) \in \Lambda_p.$$

The Jacobian of $\tilde{r}$ is an invariant function and is given by

$$(29) \qquad \det D\tilde{r}(S) = Dr(0)^2{}^{|C|} \prod_A \frac{r(\lambda_\alpha)^2 - r(\lambda_\beta)^2}{\lambda_\alpha{}^2 - \lambda_\beta{}^2} \prod_B \frac{r(\lambda_\alpha)}{\lambda_\alpha} Dr(\lambda_\alpha),$$

where $(\lambda_1,\ldots,\lambda_p) = \Pi(S)$, $A \equiv A(\lambda_1,\ldots,\lambda_p) = \{(\alpha,\beta)|\alpha < \beta,\ \lambda_\alpha > \lambda_\beta\}$, $B \equiv B(\lambda_1,\ldots,\lambda_p) = \{(\alpha,\beta)|\alpha \leq \beta,\ \lambda_\alpha = \lambda_\beta > 0\}$, $C \equiv C(\lambda_1,\ldots,\lambda_p) = \{(\alpha,\beta)|\alpha \leq \beta,\ \lambda_\alpha = \lambda_\beta = 0\}$, and $|C|$ denotes the number of elements in C.

Proof. It is seen from (25) that $r(MTM', MRM') = Mr(T,R)M'$ for $M \in GL(p, \mathbb{C})$ and $(T,R) \in E$. It follows from the representation (14) that M can be chosen such that $MTM' = I_{2p}$ and $MRM' = R_o = \mathrm{diag}(\lambda_1,\ldots,\lambda_p,-\lambda_1,\ldots,-\lambda_p)$. Then $Mr(T,R)M' = r(MTM',MRM') = r(I_{2p},R_o) = \mathrm{diag}(r(\lambda_1),\ldots,r(\lambda_p),-r(\lambda_1),\ldots,-r(\lambda_p))$, and it is seen that $(T, r(T,R)) \in E$. Using the isomorphism (23) it is then seen that $\tilde{r}(S) \in H^+(2p,\mathbb{R})$, $\tilde{r}$ corresponds to the mapping $(T,R) \to (T, r(T,R))$ of E into E, $\tilde{r}$ commutes with the action, and that (27) holds.

The next step is to show that $\tilde{r}$ is one-to-one. Since $\tilde{r}$ commutes with the action of $GL(p, \mathbb{C})$ it is enough to show that $(T, r(T,R)) = (I_{2p}, r(I_{2p}, R_o))$ implies that $T = I_{2p}$ and $R = R_o$. This means one has to show that $r(I_{2p},R) = r(I_{2p},R_o)$ implies that $R = R_o$. Since $\bar{r}$ is one-to-one it follows that R and R_o have the same eigenvalues w.r.t. I_{2p} or in other words there exists $M \varepsilon GL(p, \mathbb{C})$ such that $MM' = I_{2p}$ and $MR_oM' = R$. Then $Mr(I_{2p}, R_o)M' = r(MM', MR_oM') = r(I_{2p}, R) = r(I_{2p}, R_o) = \mathrm{diag}(r(\lambda_1),\ldots,r(\lambda_p),-r(\lambda_1),\ldots,-r(\lambda_p))$, and since M is orthogonal, one has

$$M\, r(I_{2p}, R_o) = r(I_{2p}, R_o)M. \tag{30}$$

For every odd polynomial q, (30) implies that M commutes with $\mathrm{diag}(q(r(\lambda_1)),\ldots,q(r(\lambda_p)),-q(r(\lambda_1)),\ldots,-q(r(\lambda_p)))$. Since q can be chosen such that $q(r(\lambda_\alpha)) = \lambda_\alpha$, $\alpha = 1,\ldots,p$, one obtains that M commutes with R_o and therefore that $R = MR_oM' = R_oMM' = R_o$. Therefore $\tilde{r}$ is one-to-one.

Since $\tilde{r}$ commutes with the action of $GL(p, \mathbb{C})$ and $\overline{r}$ is onto it follows from (27) that $\tilde{r}$ also is onto.

The fact that $\tilde{r}$ commutes with the action gives that the Jacobian is an invariant function. It is therefore enough to calculate $\det D\tilde{r}(S)$ when $S = I_{2p} + R_o$. Using the isomorphism (23) the Jacobian of $\tilde{r}$ is the same as the Jacobian of the mapping $(T,R) \to (T,r(T,R))$, which again is the same as the Jacobian of $R \to r(T,R)$. Thus one has to find the absolute value of the determinant of the mapping

$$dR \to \sum_{j=o}^{m} a_j \sum_{k=o}^{2j} R_o^k (dR) R_o^{2j-k} \quad . \tag{31}$$

Since dR has the form (7), where $C = (c_{\alpha\beta})$ and $D = (d_{\alpha\beta})$, it is seen that the mapping (31) multiplies $c_{\alpha\beta}$, $\alpha \leq \beta$, by

$$(32) \qquad \sum_{j=o}^{m} a_j \sum_{k=o}^{2j} \lambda_\alpha^k \lambda_\beta^{2j-k}$$

and $d_{\alpha\beta}$, $\alpha \leq \beta$, by

$$(33) \qquad \sum_{j=o}^{m} a_j \sum_{k=o}^{2j} \lambda_\alpha^k (-\lambda_\beta)^{2j-k}$$

The Jacobian is therefore a product of all these factors. If $(\alpha,\beta) \varepsilon B$ (32) is equal to $Dr(\lambda_\alpha)$ and (33) is equal to $r(\lambda_\alpha)/\lambda_\alpha$. If $(\alpha,\beta) \varepsilon C$ both (32) and (33) are equal to $Dr(0)$. If $(\alpha,\beta) \varepsilon A$ we have two geometric progressions, and it is seen that (32) is equal to $(r(\lambda_\alpha) - r(\lambda_\beta))/(\lambda_\alpha - \lambda_\beta)$ and that (33) is equal to $(r(\lambda_\alpha) + r(\lambda_\beta))/(\lambda_\alpha + \lambda_\beta)$. Since (29) is positive it follows that $\tilde{r}^{-1}$ is differentiable. $\square$

<u>Lemma 5</u>. The measure κ in Theorem 1 has the density (21) w.r.t. a Lebesgue measure on Λ_p.

<u>Proof</u>. The invariant measure $\nu_{\mathbb{R},2p,N}$ on $H^+(2p,\mathbb{R})$ has density $|\det S|^{-(2p+1)/2}$ w.r.t. a Lebesgue measure ([11], p. 93). Since $\tilde{r}$ is a diffeomorphism $\tilde{r}^{-1}(\nu_{\mathbb{R},2p,N})$ has the density

$$(34) \qquad (\det S/\det \tilde{r}(S))^{(2p+1)/2} |\det D\tilde{r}(S)|$$

w.r.t. $\nu_{\mathbb{R},2p,N}$. It follows from Lemma 4 and (16) that (34) is an invariant function $g(\pi(S))$, $S \varepsilon H^+(2p,\mathbb{R})$, where

(35) $$g(\lambda_1,\dots,\lambda_p) =$$

$$\prod_{\gamma=1}^{p}\left[\frac{1-\lambda_\gamma^2}{1-r(\lambda_\gamma)^2}\right]^{\frac{2p+1}{2}} Dr(0)^{2|C|}\prod_A \frac{r(\lambda_\alpha)^2 - r(\lambda_\beta)^2}{\lambda_\alpha^2 - \lambda_\beta^2}\prod_B \frac{r(\lambda_\alpha)}{\lambda_\alpha} Dr(\lambda_\alpha),$$

$$(\lambda_1,\dots,\lambda_p) \in \Lambda_p.$$

Using (27) we obtain that $\nu_{\mathbb{C},p,N} \otimes \overline{r}^{-1}(\kappa) = (1\times\overline{r}^{-1})(\nu_{\mathbb{C},p,N}\otimes\kappa) = (1\times\overline{r}^{-1})((t,\Pi)(\nu_{\mathbb{R},2p,N})) = (t,\Pi)(\tilde{r}^{-1}(\nu_{\mathbb{R},2p,N})) = (t,\Pi)((g\circ\Pi)\nu_{\mathbb{R},2p,N}) = \nu_{\mathbb{C},p,N}\otimes g\kappa$, so it is seen that

(36) $$\overline{r}^{-1}(\kappa) = g\kappa.$$

Let

$$k(\lambda_1,\cdots,\lambda_p) = \prod_{\gamma=1}^{p}(1-\lambda_\gamma^2)^{-(2p+1)/2}\prod_A(\lambda_\alpha^2-\lambda_\beta^2)\prod_B\lambda_\alpha,\ (\lambda_1,\dots\lambda_p)\in\Lambda_p$$

and let $\mu = (1/k)\kappa$. Then κ has density $k(>0)$ w.r.t. μ and it follows from (35) that $\overline{r}^{-1}(\mu)$ has density

$$Dr(0)^{2|C|}\prod_B Dr(\lambda_\alpha)$$

w.r.t. μ. By considering the restrictions of μ to the "faces"

$$\{(\lambda_1,\cdots,\lambda_p)\in\mathbb{R}^p \mid 1>\lambda_1=\cdots=\lambda_{m_1}>\lambda_{m_1+1}=\cdots=\lambda_{m_q}>\lambda_{m_{q+1}}=\cdots=\lambda_p=0\}$$

of Λ_p, where $1\le m_1<\dots<m_q\le p$, $q=1,2,\dots,p$, it follows from the

next lemma that μ is a Lebesgue measure on Λ_p. Hence μ is concentrated on the interior of Λ_p, and on this subset $k(\lambda_1,\cdots,\lambda_p)$ is equal to (21). $\square$

<u>Lemma 6</u>. Let μ be a measure on the interior $\{(\delta_1,\cdots,\delta_q)\in\mathbb{R}^q \mid 1>\delta_1>\cdots>\delta_q>0\}$ of Λ_q, $q=1,2,\cdots$, such that $\bar{r}^{-1}(\mu)$ has density

$$\prod_{\alpha=1}^{q} Dr(\delta_\alpha)^{n_\alpha},$$

$n_\alpha = 1,2,\cdots$, $\alpha=1,\cdots,q$, w.r.t. μ for any uneven polynomial r such that $r(0)=0$, $r(1)=1$, and $Dr(x)>0$, $0\le x\le 1$. If μ is not identically zero then $n_\alpha=1$, $\alpha=1,\cdots,q$, and μ is the restriction of a Lebesgue measure to the interior of Λ_q.

<u>Proof</u>. The condition on μ is that

$$\int f d\mu = \int f(r(\delta_1),\cdots,r(\delta_q)) \prod_{\alpha=1}^{q} Dr(\delta_\alpha)^{n_\alpha}\, d\mu(\delta_1,\cdots,\delta_q) \tag{37}$$

for any continuous function f with compact support. For a fixed f it follows from the Weierstrass approximation theorem that (37) even holds for any continuously differentiable function r with $r(0)=0$, $r(1)=1$, and $Dr(x)>0$, $0\le x\le 1$. [If the polynomials q_n tend uniformly to $\sqrt{Dr}$ on $[0,1]$ then the polynomials $r_n(x) = \int_0^x q_n(t)^2dt + x(1-\int_0^1 q_n(t)^2dt)$ are odd, $r_n(0)=0$, $r_n(1)=1$, and r_n and Dr_n tend uniformly to r and Dr, respectively, on $[0,1]$. Hence $Dr_n(x)>0$, $0\le x\le 1$, for n sufficiently large.] By monotone convergence (37) is extended to the case where f is an indicator function for a compact subset. Let $I = \underset{\alpha}{X}[a_\alpha, a_\alpha + c_\alpha]$ and $J = \underset{\alpha}{X}[b_\alpha, b_\alpha + c_\alpha]$ be two rectangles in the interior of Λ_q with sidelengths $c_\alpha > 0$, $\alpha = 1,\cdots,q$. Since $1 > a_1 + c_1 > a_1 > \cdots > a_p + c_p > a_p > 0$ and $1 > b_1 + c_1 > b_1 > \cdots > b_p + c_p > b_p > 0$ there exists a continuously differentiable function r with $r(0) = 0$, $r(1) = 1$, and $Dr(x) > 0$, $0 \leq x \leq 1$, such that $r(x) = x + (b_\alpha - a_\alpha)$, $x \in [a_\alpha, a_\alpha + c_\alpha]$, $\alpha = 1,\cdots,q$. Then $Dr(x) = 1$ for $x \in [a_\alpha, a_\alpha + c_\alpha]$, $\alpha = 1,\cdots,q$, and if f is the indicator function of J it follows from (37) that $\mu(I) = \mu(J)$. Hence μ is translation invariant. If μ is not identically zero it must be the restriction of a Lebesgue measure, and then it is clear that $n_\alpha = 1$, $\alpha = 1,\cdots,q$. $\square$

3. TESTING THE HYPOTHESIS THAT A COVARIANCE MATRIX WITH COMPLEX STRUCTURE HAS REAL STRUCTURE.

Let $x_1,\cdots,x_N$, $N \geq p$, be i.i.d. observations from a normal distribution on $\mathbb{R}^{2p}$ with mean vector 0 and unknown covariance matrix

$$\Sigma = \begin{pmatrix} \Gamma & -\Psi \\ \Psi & \Gamma \end{pmatrix} \in H^+(p, \mathbb{C})$$

It follows from section 2 that the empirical covariance matrix transformed by the mapping (12) is a minimal sufficient reduction which follows a complex Wishart distribution with N degrees of freedom and parameter $\frac{1}{N}\Sigma$. This distribution has the density (19) w.r.t. $\nu_{\mathbb{C},p,N}$.

For $p > 1$ we shall consider the hypothesis H_0 that $\psi = 0$, i.e., that Σ has a real structure. The statistical problem of testing H_0 is invariant under the restriction of the action (5) to the subgroup $GL(p, \mathbb{R}) \otimes I_2 = \{\mathrm{diag}(A,A) | A \in GL(p, \mathbb{R})\}$ of $GL(p, \mathbb{C})$. The problem is to find a representation of the orbit projection

$$\Pi : \mathcal{H}^+(p, \mathbb{C}) \to \mathcal{H}^+(p, \mathbb{C})/GL(p, \mathbb{R}) \otimes I_2 \tag{38}$$

and, when H_0 is true, the distribution of Π.

The group $GL(p, \mathbb{R}) \otimes I_2$ acts on $\mathcal{H}^+(p, \mathbb{R}) \otimes I_2 \equiv \{\mathrm{diag}(H,H) | H \in \mathcal{H}^+(p, \mathbb{R})\}$ by restriction of (5). The linear map

$$t : \mathcal{H}^+(p, \mathbb{C}) \to \mathcal{H}^+(p, \mathbb{R}) \otimes I_2$$

$$S = \begin{pmatrix} H & -F \\ F & H \end{pmatrix} \to \begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix}$$

commutes with the actions of $GL(p, \mathbb{R}) \otimes I_2$, and the residual

$$S - t(S) = \begin{pmatrix} 0 & -F \\ F & 0 \end{pmatrix} \in \mathcal{A}(p, \mathbb{R}) \otimes J_1,$$

where

$$\mathcal{A}(p, \mathbb{R}) \otimes J_1 = \left\{ \begin{pmatrix} 0 & -F \\ F & 0 \end{pmatrix} \middle| \; F \in \mathcal{A}(p, \mathbb{R}) \right\}$$

and $\mathcal{A}(p, \mathbb{R})$ is the set of all antisymmetric $p \times p$ real matrices.

Lemma 7. Let ϕ be a positive definite symmetric bilinear real form and ψ an antisymmetric bilinear real form on a p-dimensional real vector space. Then there exists a basis such that the matrices of ϕ and ψ are I_p and

$$(39)\qquad \Lambda = \begin{pmatrix} 0 & -\lambda_1 & 0 & 0 \ldots 0 \\ +\lambda_1 & 0 & 0 & 0 \ldots 0 \\ 0 & 0 & 0 & -\lambda_2 \ldots 0 \\ 0 & 0 & \lambda_2 & 0 \ldots 0 \\ . & . & . & . \ldots . \end{pmatrix},$$

respectively, where $\lambda_1 \geq \cdots \geq \lambda_{[p/2]} \geq 0$.

Proof. Bourbaki [9], p. 123. □

An equivalent formulation of the lemma is that for every $H \varepsilon \mathcal{H}^+(p, \mathbb{R})$ and $F \varepsilon \mathcal{A}(p, \mathbb{R})$ there exists $A \varepsilon GL(p, \mathbb{R})$ such that $AHA' = I_p$ and $AFA' = \Lambda$. This yields another equivalent formulation in terms of $2p \times 2p$ real matrices. For $T \equiv \begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix} \varepsilon \mathcal{H}^+(p, \mathbb{R}) \otimes I_2$ and $R \equiv \begin{pmatrix} 0 & -F \\ F & 0 \end{pmatrix} \varepsilon \mathcal{A}(p, \mathbb{R}) \otimes J_1$, there exists

$$M \equiv \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \varepsilon GL(p, \mathbb{R}) \otimes I_2$$

such that

$$(40)\qquad MTM' = I_{2p} \quad \text{and} \quad MRM' = \begin{pmatrix} 0 & -\Lambda \\ \Lambda & 0 \end{pmatrix}.$$

It is seen from (40) that $\pm\lambda_1,\cdots,\pm\lambda_{[p/2]}$ are uniquely determined as the eigenvalues of R w.r.t. T, each with multiplicity two. (When p is odd 0 is always an eigenvalue with multiplicity two.)

For $T = t(S)$ and $R = S - t(S)$ it now follows that there exists an $M \varepsilon GL(p, \mathbb{R}) \otimes I_2$ such that

$$(41) \qquad MSM' = \begin{pmatrix} I_p & -\Lambda \\ \Lambda & I_p \end{pmatrix},$$

and that (38) can be represented by

$$(42) \qquad \Pi: H^+(p, \mathbb{C}) \to \Lambda_{[p/2]},$$

where $\Pi(S)$ is the ordered family of the nonnegative eigenvalues of $S - t(S)$ w.r.t. $t(S)$. As in section 2 it is seen that this representation is also topological.

<u>Theorem 2</u>. The maximum likelihood estimator of Σ under H_o is $t(S)$ and the likelihood ratio statistic for testing H_o is

$$(43) \qquad \prod_{\gamma=1}^{[p/2]} (1 - \lambda_\gamma^2)^N,$$

where $(\lambda_1,\cdots,\lambda_{[p/2]}) = \Pi(S)$. Under the hypothesis H_o the statistics $t(S)$ and $\Pi(S)$ are independently distributed. The distribution of $t(S)$ has the density

$$(44) \qquad \left(\frac{\det T}{\det \Sigma}\right)^{N/2} \exp(-N/2 \ \mathrm{tr}(\Sigma^{-1}T)), T \varepsilon H^+(p, \mathbb{R}) \otimes I_2$$

w.r.t. a unique measure $\nu'_{\mathbb{R},p,N}$ which is invariant under the action of $GL(p,\mathbb{R}) \otimes I_2$ on $H^+(p,\mathbb{R}) \otimes I_2$. The distribution of $\Pi(S)$ has density (43) w.r.t. a measure κ on $\Lambda_{[p/2]}$ which is uniquely defined by

$$(t,\Pi)(\nu_{\mathbb{C},p,N}) = \nu'_{\mathbb{R},p,N} \otimes \kappa .$$

Furthermore, κ has the density

$$\prod_{1 \le \alpha < \beta \le [p/2]} (\lambda_\alpha^2 - \lambda_\beta^2)^2 \prod_{\gamma=1}^{[p/2]} \lambda_\gamma^{2\varepsilon} (1 - \lambda_\gamma^2)^{-p} , \tag{45}$$

where $\varepsilon = p - 2[p/2]$, w.r.t. a Lebesgue measure on $\Lambda_{[p/2]}$.

Remark. The maximum likelihood estimator of Γ under H_o is H. Since the mapping $T = \mathrm{diag}(H,H) \to H$ from $H^+(p,\mathbb{R}) \otimes I_2$ onto $H^+(p,\mathbb{R})$ transforms $\nu'_{\mathbb{R},p,N}$ into $\nu_{\mathbb{R},p,2N}$, it is seen from (44) that H is Wishart distributed with $2N$ degrees of freedom and parameter $\frac{1}{2N}\Gamma$.

Proof. If in section 2 one replaces $GL(p,\mathbb{C})$ by $GL(p,\mathbb{R}) \otimes I_2$, $H^+(2p,\mathbb{R})$ by $H^+(p,\mathbb{C})$, $H^+(p,\mathbb{C})$ by $H^+(p,\mathbb{R}) \otimes I_2$, $S(p,\mathbb{C})$ by $A(p,\mathbb{R}) \otimes J_1$, $\nu_{\mathbb{R},2p,N}$ by $\nu_{\mathbb{C},p,N}$; $\nu_{\mathbb{C},p,N}$ by $\nu'_{\mathbb{R},p,N}$ and $(\lambda_1,\cdots,\lambda_p)$ by $(\lambda_1,\cdots,\lambda_{[p/2]})$, then the proof is completely analogous to the proof of Theorem 1 except for the following changes: The determinant of the right hand side of (41) is $\prod_{\gamma=1}^{[p/2]} (1 - \lambda_\gamma^2)^2$ which gives (43). The invariant measure $\nu_{\mathbb{C},p,N}$ has density $|\det S|^{-p/2}$ w.r.t. a Lebesgue measure on $H^+(p,\mathbb{C})$ (Bourbaki [11], p. 93) which gives the factor $\prod_{\gamma=1}^{[p/2]} (1 - \lambda_\gamma^2)^{-p}$ in (45). Finally, we shall show below that the Jacobian of

$$dR \to \sum_{j=1}^{m} a_j \sum_{k=0}^{2j} R_0^k \,(dR) R_0^{2j-k} \ ,$$

where

$$dR = \begin{pmatrix} 0 & -F \\ F & 0 \end{pmatrix} \varepsilon\ A(p,\mathbb{R}) \otimes J_1 \quad \text{and} \quad R_0 = \begin{pmatrix} 0 & -\Lambda \\ \Lambda & 0 \end{pmatrix},$$

is

$$(46) \qquad \prod_{\gamma=1}^{[p/2]} Dr(\lambda_\gamma) \prod_{1 \le \alpha < \beta \le [p/2]} \left(\frac{r(\lambda_\alpha)^2 - r(\lambda_\beta)^2}{\lambda_\alpha^2 - \lambda_\beta^2}\right)^2 \prod_{\gamma=1}^{[p/2]} \left(\frac{r(\lambda_\gamma)}{\lambda_\gamma}\right)^{2\varepsilon}$$

when $1 > \lambda_1 > \cdots > \lambda_{[p/2]} > 0$, which gives the remaining factors in (45).

To prove (46), first note that $J_1^{2j+1} = (-1)^j J_1$ implies

$$\sum_{j=}^{m} a_j \sum_{k=0}^{2j} \begin{pmatrix} 0 & -\Lambda \\ \Lambda & 0 \end{pmatrix}^k \begin{pmatrix} 0 & -F \\ F & 0 \end{pmatrix} \begin{pmatrix} 0 & -\Lambda \\ \Lambda & 0 \end{pmatrix}^{2j-k} = \begin{pmatrix} 0 & -F_0 \\ F_0 & 0 \end{pmatrix},$$

where $F_0 = \sum_{j=1}^{m} a_j (-1)^j \sum_{k=0}^{2j} \Lambda^k F \Lambda^{2j-k}$. Now partition $F = (F_{\alpha,\beta})$, where $F_{\alpha,\beta}$ is 2×2, $1 \le \alpha,\beta \le [p/2]$, and where if p is odd, $F_{\gamma,[p/2]+1}$ is 2×1, $1 \le \gamma \le [p/2]$. Then $F_{\gamma,\gamma}$ for $1 \le \gamma \le [p/2]$ is of the form tJ_1 and is mapped into

$$\sum_{j=1}^{m} a_j (-1)^j \sum_{k=0}^{2j} \begin{pmatrix} 0 & -\lambda_\gamma \\ \lambda_\gamma & 0 \end{pmatrix}^k \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} \begin{pmatrix} 0 & -\lambda_\gamma \\ \lambda_\gamma & 0 \end{pmatrix}^{2j-k} =$$

$$\sum_{j=1}^{m} a_j \lambda_\gamma^{2j}(2j+1) \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} = Dr(\lambda_\gamma) \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix},$$

which gives the factor $Dr(\lambda_\gamma)$ in (46). For $1 \le \alpha < \beta \le [p/2]$, $F_{\alpha,\beta} \equiv \begin{pmatrix} v & s \\ w & u \end{pmatrix}$ is mapped into

$$\sum_{j=1}^{m} a_j(-1)^j \sum_{k=o}^{j} \begin{pmatrix} 0 & -\lambda_\alpha \\ \lambda_\alpha & 0 \end{pmatrix}^{2k} \begin{pmatrix} v & s \\ w & u \end{pmatrix} \begin{pmatrix} 0 & -\lambda_\beta \\ \lambda_\beta & 0 \end{pmatrix}^{2j-2k} +$$

$$\sum_{j=1}^{m} a_j(-1)^j \sum_{k=o}^{j-1} \begin{pmatrix} 0 & -\lambda_\alpha \\ \lambda_\alpha & 0 \end{pmatrix}^{2k+1} \begin{pmatrix} v & s \\ w & u \end{pmatrix} \begin{pmatrix} 0 & \lambda_\beta \\ -\lambda_\beta & 0 \end{pmatrix}^{2j-2k-1}$$

$$= \sum_{j=1}^{m} a_j \sum_{k=o}^{j} \lambda_\alpha^{2k} \lambda_\beta^{2j-2k} \begin{pmatrix} v & s \\ w & u \end{pmatrix}$$

$$- \sum_{j=1}^{m} a_j \sum_{k=o}^{j-1} \lambda_\alpha^{2k+1} \lambda_\beta^{2j-2k-1} \begin{pmatrix} -u & w \\ s & -v \end{pmatrix} =$$

$$\begin{pmatrix} av + bu & as - bw \\ aw - bs & au + bv \end{pmatrix} ,$$

where

$$a = (\lambda_\alpha r(\lambda_\alpha) - \lambda_\beta r(\lambda_\beta))/(\lambda_\alpha^2 - \lambda_\beta^2) \quad \text{and}$$

$$b = (\lambda_\beta r(\lambda_\alpha) - \lambda_\alpha r(\lambda_\beta))/(\lambda_\alpha^2 - \lambda_\beta^2).$$

The determinant of this mapping is $(a^2 - b^2)^2 = [(r(\lambda_\alpha)^2 - r(\lambda_\beta)^2)/(\lambda_\alpha^2 - \lambda_\beta^2)]^2$, which gives the (α,β)'th factor in (46). If p is odd then $F_{\gamma,[p/2]+1} \equiv \binom{v}{s}$ is mapped into

$$\sum_{j=1}^{m} a_j(-1)^j \begin{pmatrix} 0 & -\lambda_\gamma \\ \lambda_\gamma & 0 \end{pmatrix}^{2j} \binom{r}{s} = \frac{r(\lambda_\gamma)}{\lambda_\gamma} \binom{r}{s},$$

which gives the factor $(r(\lambda_\gamma)/\lambda_\gamma)^{2\varepsilon}$ in (46). $\square$

4. TESTING THE HYPOTHESIS THAT A COVARIANCE MATRIX WITH COMPLEX STRUCTURE HAS QUATERNION STRUCTURE.

Let $\mathbb{H}$ denote the division algebra over $\mathbb{R}$ of quaternions and let $1,i,j,k$ be a canonical basis; i.e., $i^2 = j^2 = k^2 = -1$, $ij = -ji = k$, $jk = -kj = i$ and $ki = -ik = j$. The conjugate of a quaternion $q = a + ib + jc + kd$ is denoted by $\overline{q} = a - ib - jc - kd$. Define the $\mathbb{R}$-linear map ρ of $\mathbb{H}$ into $\mathbb{H}$ by $\rho(a + ib + jc + kd) = a + ib - jc + kd$. It is seen that ρ is an involutive antiautomorphism of $\mathbb{H}$; i.e. $\rho(\rho(q)) = q$ and $\rho(q_1q_2) = \rho(q_2)\rho(q_1)$ for $q, q_1, q_2 \varepsilon \mathbb{H}$. Moreover, the set $\{q \varepsilon \mathbb{H} | \rho(q) = \overline{q}\} = \{a + jc | a,c \varepsilon \mathbb{R}\}$ is isomorphic to the field of complex numbers, so it is here denoted by $\mathbb{C}$. (More generally, it can be shown that there is a one to one correspondence between the embeddings of the field of complex numbers into $\mathbb{H}$ and the involutive antiautomorphisms of $\mathbb{H}$ other than the conjugation operator.)

Let E be a p-dimensional right vector space over $\mathbb{H}$. Since E is a right vector space the scalar multiplication of an $x \varepsilon E$ by $q \varepsilon \mathbb{H}$ is denoted xq. By restricting the scalar multiplication to the subalgebra of real numbers, E is also a 4p-dimensional vector space over $\mathbb{R}$, and if $e_1, \cdots, e_p$ is a basis for E as a vector space over $\mathbb{H}$ then

$$e_1,\ldots,e_p, e_1 i,\ldots,e_p i, e_1 j,\ldots,e_p j, e_1 k,\ldots,e_p k \tag{47}$$

becomes a basis for E as a vector space over $\mathbb{R}$.

If f is an $\mathbb{H}$-linear map of E into E with matrix $A + iB_1 + jB_2 + kB_3$ w.r.t. $e_1,\ldots,e_p$ then f considered as an $\mathbb{R}$-linear map has matrix

$$(48)\qquad M = \begin{pmatrix} A & -B_1 & -B_2 & -B_3 \\ B_1 & A & -B_3 & B_2 \\ B_2 & B_3 & A & -B_1 \\ B_3 & -B_2 & B_1 & A \end{pmatrix}$$

w.r.t. (47). Since composition of linear maps corresponds to multiplication of matrices it is seen that the set $GL(p,\mathbb{H})$ of all nonsingular $4p \times 4p$ matrices of the form (48) is a subgroup of $GL(2p,\mathbb{C})$

Let $\phi : E \times E \to \mathbb{H}$ be a hermitian left sesquilinear quaternion form on E (see [9] or [1]) with matrix $\Phi = C + iD_1 + jD_2 + kD_3 = (\phi(e_\alpha, e_\beta))$. Then C is symmetric and D_1, D_2 and D_3 are antisymmetric. Moreover, $\delta \equiv \mathrm{Re} \circ \phi$, where Re denotes the real part of a quarternion, is a symmetric bilinear real form on E and the matrix of δ w.r.t. (47) is

$$(49)\qquad \Delta = \begin{pmatrix} C & -D_1 & -D_2 & -D_3 \\ D_1 & C & -D_3 & D_2 \\ D_2 & D_3 & C & -D_1 \\ D_3 & -D_2 & D_1 & C \end{pmatrix} .$$

The form $\phi \circ (f \times f)$ is also hermitian left sesquilinear and the matrix of $\mathrm{Re} \circ (\phi \circ (f \times f)) = (\mathrm{Re} \circ \phi) \circ (f \times f)$ w.r.t. (47) is

$$(50)\qquad M' \Delta M.$$

Since ϕ is positive definite if and only if $\mathrm{Re} \circ \phi$ is positive definite it follows that the action

$$(51)\qquad \begin{aligned} GL(p,\mathbb{H}) \times H^+(p,\mathbb{H}) &\to H^+(p,\mathbb{H}) \\ (M,T) &\to MTM', \end{aligned}$$

where $H^{+}(p,\mathbb{H})$ is the set of positive definite matrices of the form (49), is well defined. It follows from the first equation in (59) below that this action is transitive. Moreover, since it is a restriction of the proper action (6) it is also proper. We remark that $H^{+}(p,\mathbb{H})$ is a subset of $H^{+}(2p, \mathbb{C})$.

Next let $\phi: E \times E \to \mathbb{H}$ be a hermitian left ρ-sesquilinear quaternion form on E, i.e., $\phi(xq_1, yq_2) = \rho(q_1)\,\phi(x,y)q_2$, $\phi(y,x) = \rho(\phi(x,y))$, and $\phi(x_1 + x_2, y) = \phi(x_1, y) + \phi(x_2, y)$, $q_1, q_2 \varepsilon \mathbb{H}$, $y, x, x_1, x_2 \varepsilon E$. If $\Phi = C + iD_1 + jD_2 + kD_3 = (\phi(e_\alpha, e_\beta))$ is the matrix of ϕ then C, D_1, and D_3 are symmetric and D_2 is antisymmetric. Moreover, $\delta \equiv \mathrm{Re} \circ \phi$ is a symmetric bilinear real form on E, and the matrix of δ w.r.t. (47) is

$$(52) \qquad \Xi = \begin{pmatrix} C & -D_1 & -D_2 & -D_3 \\ -D_1 & -C & D_3 & -D_2 \\ D_2 & D_3 & C & -D_1 \\ -D_3 & D_2 & -D_1 & -C \end{pmatrix} .$$

The form $\phi \circ (f \times f)$ is also hermitian left ρ-sesquilinear and the matrix of $\mathrm{Re} \circ (\phi \circ (f \times f)) = (\mathrm{Re} \circ \phi) \circ (f \times f)$ w.r.t. (47) is

$$(53) \qquad M' \Xi M \varepsilon S(p, \mathbb{H}),$$

where $S(p,\mathbb{H})$ is the set of matrices of the form (52).

Let $x_1, \ldots, x_N$, $N \geq 2p$, be i.i.d. observations from a normal distribution on $\mathbb{R}^{4p}$ with mean vector 0 and unknown covariance matrix $\Sigma \varepsilon H^{+}(2p, \mathbb{C})$. As in section 3 one obtains as a minimal sufficient statistic an observation S from a complex Wishart distribution on $H^+(2p, \mathbb{C})$ with N degrees of freedom and parameter $\frac{1}{N}\Sigma$. This distribution has the density (19) (with p replaced by 2p) w.r.t. $\nu_{\mathbb{C},2p,N}$.

Let H_0 denote the hypothesis that $\Sigma \in H^+(p, \mathbb{H})$, i.e., that Σ has quaternion structure. The statistical problem of testing H_0 is invariant under the restriction of the action (5) (with p replaced by 2p) to the subgroup $GL(p, \mathbb{H})$. The problem is to find a representation of the orbit projection

$$\Pi : H^+(2p, \mathbb{C}) \to H^+(2p, \mathbb{C})/GL(p, \mathbb{H}) \tag{54}$$

and, when H is true, the distribution of Π.

Let

$$J \equiv J_p = \begin{pmatrix} 0 & 0 & 0 & -I_p \\ 0 & 0 & I_p & 0 \\ 0 & -I_p & 0 & 0 \\ I_p & 0 & 0 & 0 \end{pmatrix}. \tag{55}$$

It is seen that the linear map

$$t : H^+(2p, \mathbb{C}) \to H^+(p, \mathbb{H})$$

$$S = \begin{pmatrix} H_{11} & H_{12} & -F_{11} & -F_{12} \\ H_{21} & H_{22} & -F_{21} & -F_{22} \\ F_{11} & F_{12} & H_{11} & H_{12} \\ F_{21} & F_{22} & H_{21} & H_{22} \end{pmatrix} \to \frac{1}{2}\begin{pmatrix} H_{11}+H_{22} & H_{12}-H_{21} & -F_{11}+F_{22} & -F_{12}-F_{21} \\ H_{21}-H_{12} & H_{22}+H_{11} & -F_{21}-F_{12} & -F_{22}+F_{11} \\ F_{11}-F_{22} & F_{12}+F_{21} & H_{11}+H_{22} & H_{12}-H_{21} \\ F_{21}+F_{12} & F_{22}-F_{11} & H_{21}-H_{12} & H_{22}+H_{11} \end{pmatrix} \tag{56}$$

is well defined, because $t(S) = \frac{1}{2}(S + JSJ') \in H^+(p, \mathbb{H})$. Since J is the matrix w.r.t. (47) for scalar multiplication by k, it follows that J commutes with all matrices of the form (48), and thus also that t commutes with the actions of $GL(p, \mathbb{H})$. Moreover, the residual

$$(57)\qquad S - t(S) = \frac{1}{2}\begin{Bmatrix} H_{11}-H_{22} & H_{12}+H_{21} & -F_{11}-F_{22} & -F_{12}+F_{21} \\ H_{21}+H_{12} & H_{22}-H_{11} & -F_{21}+F_{12} & -F_{22}-F_{11} \\ F_{11}+F_{22} & F_{12}-F_{21} & H_{11}-H_{22} & H_{12}+H_{21} \\ F_{21}-F_{12} & F_{22}+F_{11} & H_{21}+H_{12} & H_{22}-H_{11} \end{Bmatrix} \varepsilon\, \mathcal{S}(p, \mathbb{H}).$$

Lemma 8. Let ϕ be a positive definite hermitian left sesquilinear form on E and ψ a hermitian left ρ-sesquilinear form on E. Then there exists a basis for E such that the matrices of ϕ and ψ are I_p and Λ, respectively, where Λ is defined in Lemma 1.

Proof. (See Bourbaki [9], p. 120.) Since ϕ is positive definite there exists an additive map u of E into E such that $\psi(x,y) = \phi(u(x),y)$, $x, y \in E$. For $q \in \mathbb{H}$, $\phi(u(xq),y) = \psi(xq,y) = \rho(q)\psi(x,y) = \rho(q)\phi(u(x),y) = \phi(u(x)\overline{\rho(q)},y)$. Hence

$$(58)\qquad u(xq) = u(x)\overline{\rho(q)},\ q \in \mathbb{H},\ x \in E.$$

Moreover, $\phi(u(x),y) = \psi(x,y) = \rho(\psi(y,x)) = \rho(\phi(u(y),x))$, and it follows that the orthogonal complement w.r.t. ϕ of any u-invariant $\mathbb{H}$-subspace S of E (i.e., $u(S) \subseteq S$) is also u-invariant. E is therefore a direct orthogonal sum of minimal u-invariant subspaces. Let S be one of these subspaces. Since $\overline{\rho(q)} = q$ for $q \in \mathbb{C} \equiv \{a + jc \mid a,c \in \mathbb{R}\}$, it follows from (58) that the restriction of u to S is a $\mathbb{C}$-linear map of S into S, and it must have an eigenvector. Therefore there exists $\lambda \in \mathbb{C}$ and $x \in S$ such that $\phi(x,x) = 1$ and $u(x) = x\lambda$. Hence $u(x\mathbb{H}) \subseteq x\mathbb{H}$ and, since $x\mathbb{H} \subseteq S$, $S = x\mathbb{H}$. Since $\overline{\lambda}\phi(x,x) = \phi(x\lambda,x) = \phi(u(x),x) = \psi(x,x) = \rho(\psi(x,x)) = \rho(\phi(u(x),x)) = \rho(\overline{\lambda}\phi(x,x)) = \phi(x,x)\rho(\overline{\lambda})$, it follows that $\rho(\overline{\lambda}) = \overline{\lambda}$, and, since $\lambda \in \mathbb{C}$, $\lambda \in \mathbb{R}$. This shows that there exists an orthonormal basis $e_1,\ldots,e_p$ for E and $\lambda_1,\ldots,\lambda_p \in \mathbb{R}$ such that $u(e_\alpha) = e_\alpha\lambda_\alpha$, $\alpha = 1,\ldots,p$; i.e. the matrix of ϕ is the identity matrix and the matrix of ψ is $\mathrm{diag}(\lambda_1,\ldots,\lambda_p)$. It can be assumed that each $\lambda_\alpha \geq 0$, for if $\lambda_\alpha < 0$ we can replace e_α by $e_\alpha i$ and thus replace λ_α by $-\lambda_\alpha > 0$. $\square$

Since the matrices of $\mathrm{Re}\circ\phi$ and $\mathrm{Re}\circ\psi$ transform according to (50) and (53), the lemma has an equivalent formulation in terms of $4p \times 4p$ real matrices. Let $T \in H^+(p,\mathbb{H})$ and $R \in S(p,\mathbb{H})$. Then there exists an $M \in GL(p,\mathbb{H})$ such that

(59) $$MTM' = I_{4p} \quad \text{and} \quad MRM' = \mathrm{diag}(\Lambda,-\Lambda,\Lambda,-\Lambda).$$

It is seen from (59) that $\pm\lambda_1,\ldots,\pm\lambda_p$ are uniquely determined as the eigenvalues of R w.r.t. T, each with multiplicity two.

For $T = t(S)$ and $R = S - t(S)$ it now follows that there exists an $M \in GL(p,\mathbb{H})$, such that

(60) $$MSM' = \mathrm{diag}(I_p+\Lambda,\ I_p-\Lambda,\ I_p+\Lambda,\ I_p-\Lambda)$$

and that (54) can be represented by

(61) $$\Pi: H^+(2p,\mathbb{C}) \to \Lambda_p,$$

where $\Pi(S)$ is the ordered family of non-negative eigenvalues of $S - t(S)$ w.r.t. $t(S)$. As in section 2 it is seen that this representation is also topological.

Theorem 3. The maximum likelihood estimator of Σ under H_0 is $t(S)$ and the likelihood ratio statistic for testing H_0 is

(62) $$\prod_{\gamma=1}^{p} (1-\lambda_\gamma^2)^N,$$

where $(\lambda_1,\ldots,\lambda_p) = \Pi(S)$. Under the hypothesis H_0 the statistics $t(S)$ and $\Pi(S)$ are independently distributed. The distribution of $t(S)$ has the density

(63) $$\left(\frac{\det T}{\det \Sigma}\right)^{N/2} \exp(-N/2\,\mathrm{tr}(\Sigma^{-1}T)),\ T \in H^+(p,\mathbb{H})$$

w.r.t. a unique measure $\nu_{\mathbb{H},p,N}$ which is invariant under the action (51). The distribution of $\Pi(S)$ has density (62) w.r.t. a measure κ on Λ_p which is uniquely defined by

(64) $$(t,\Pi)(\nu_{\mathbb{C},2p,N}) = \nu_{\mathbb{H},p,N} \otimes \kappa.$$

Furthermore, κ has the density

(65) $$\prod_{1\le\alpha<\beta\le p} (\lambda_\alpha^2 - \lambda_\beta^2)^2 \prod_{\gamma=1}^{p} \lambda_\gamma^2(1-\lambda_\gamma^2)^{-2p}$$

w.r.t. a Lebesgue measure on Λ_p.

Remark. The distribution given by the density (63) is called the quaternion Wishart distribution with N degrees of freedom and parameter $\frac{1}{N}\Sigma \in H^+(p,\mathbb{H})$.

Proof. If in Section 2 one replaces $GL(p,\mathbb{C})$ by $GL(p,\mathbb{H})$, $H^+(2p,\mathbb{R})$ by $H^+(2p,\mathbb{C})$, $H^+(p,\mathbb{C})$ by $H^+(p,\mathbb{H})$, $S(p,\mathbb{C})$ by $S(p,\mathbb{H})$, $\nu_{\mathbb{R},2p,N}$ by $\nu_{\mathbb{C},2p,N}$, and $\nu_{\mathbb{C},p,N}$ by $\nu_{\mathbb{H},p,N}$, then the proof is completely analogous to the proof of Theorem 1 except for the following changes: The determinant of the right hand side of (60) is $\Pi(1-\lambda_\gamma^2)^2$, which gives (62). The invariant measure $\nu_{\mathbb{C},2p,N}$ has density $|\det S|^{-2p/2}$

w.r.t. a Lebesgue measure on $H^+(2p, \mathbb{C})$ (Bourbaki [11], p. 93) which gives the factor $\Pi(1-\lambda_\gamma^2)^{-2p}$ in (65). Finally the Jacobian of

$$dR \to \sum_{j=1}^{m} a_j \sum_{k=0}^{2j} R_0^k (dR) R_0^{2j-k},$$

where dR has the form (52) and $R_0 = \mathrm{diag}(\Lambda, -\Lambda, \Lambda, -\Lambda)$, is

$$\prod_{\gamma=1}^{p} Dr(\lambda_\gamma)\left(\frac{r(\lambda_\gamma)}{\lambda_\gamma}\right)^2 \prod_{1\le\alpha<\beta\le p}\left(\frac{r(\lambda_\alpha)^2 - r(\lambda_\beta)^2}{\lambda_\alpha^2 - \lambda_\beta^2}\right)^2 \tag{66}$$

when $1 > \lambda_1 > \ldots > \lambda_p > 0$, which gives the remaining factors in (65). The proof of (66) is analogous to the proof of (29). $\square$

5. TESTING THE HYPOTHESIS THAT A COVARIANCE MATRIX WITH QUATERNION STRUCTURE HAS COMPLEX STRUCTURE.

Let E be as in Section 2, f a $\mathbb{C}$-linear map of E into E, $\phi : E \times E \to \mathbb{C}$ an antisymmetric bilinear complex form on E, and $C + iD = (\phi(e_\alpha, e_\beta))$ the matrix of ϕ w.r.t. $e_1, \ldots, e_p$. Then C and D are antisymmetric. Moreover $\mathrm{Re} \circ \phi$ is an antisymmetric bilinear real form on E, the matrix of $\mathrm{Re} \circ \phi$ w.r.t. (1) is given by (7), $(\phi \circ (f \times f))$ is also antisymmetric bilinear, and the matrix of $\mathrm{Re} \circ (\phi \circ (f \times f)) = (\mathrm{Re} \circ \phi) \circ (f \times f)$ w.r.t. (1) is given by

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix}' \begin{pmatrix} C & -D \\ -D & -C \end{pmatrix} \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \in A(p, \mathbb{C}), \tag{67}$$

where $A(p, \mathbb{C})$ is the set of all antisymmetric matrices of the form (7).

Let $x_1,\ldots,x_N$, $N \geq p$, be i.i.d. observations from a normal distribution on $\mathbb{R}^{4p}$ with mean vector 0 and unknown covariance matrix

$$\Sigma = \begin{pmatrix} \Gamma & -\Psi_1 & -\Psi_2 & -\Psi_3 \\ \Psi_1 & \Gamma & -\Psi_3 & \Psi_2 \\ \Psi_2 & \Psi_3 & \Gamma & -\Psi_1 \\ \Psi_3 & -\Psi_2 & \Psi_1 & \Gamma \end{pmatrix}.$$

It follows from Sections 2 and 3 that the empirical covariance matrix transformed by the mappings (12) (with p replaced by $2p$) followed by (56) is a minimal sufficient statistic, which follows a quaternion Wishart distribution with N degrees of freedom and parameter $\frac{1}{N}\Sigma$. This distribution has the density (63) w.r.t. $\nu_{\mathbb{H},p,N}$.

For $p > 1$ we shall consider the hypothesis H_0 that $\psi_2 = \psi_3 = 0$, i.e., that Σ has a complex structure. The statistical problem of testing H_0 is invariant under the restriction of the action (51) to the subgroup $GL(p,\mathbb{C}) \otimes I_2 = \{\mathrm{diag}(A,A) \mid A \in GL(p,\mathbb{C})\}$ of $GL(p,\mathbb{H})$. The problem is to find a representation of the orbit projection

(68) $$\Pi : H^+(p,\mathbb{H}) \to H^+(p,\mathbb{H}) / GL(p,\mathbb{C}) \otimes I_2$$

and, when H_0 is true, the distribution of Π.

The group $GL(p,\mathbb{C}) \otimes I_2$ acts on $H^+(p,\mathbb{C}) \otimes I_2 \equiv \{\mathrm{diag}(H,H) \mid H \in H^+(p,\mathbb{C})\}$ by restriction of (51). The linear map

$$t : H^+(p,\mathbb{H}) \to H^+(p,\mathbb{C}) \otimes I_2$$

$$S = \begin{pmatrix} H & -F \\ F & H \end{pmatrix} \to \begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix}$$

commutes with the actions of $GL(p, \mathbb{C}) \otimes I_2$, and the residual

$$S - t(S) = \begin{pmatrix} 0 & -F \\ F & 0 \end{pmatrix} \varepsilon \ A(p, \mathbb{C}) \otimes J_1,$$

where

$$A(p, \mathbb{C}) \otimes J_1 = \left\{ \begin{pmatrix} 0 & -F \\ F & 0 \end{pmatrix} \mid F \ \varepsilon \ A(p, \mathbb{C}) \right\}.$$

<u>Lemma 9</u>. Let ϕ be a positive definite hermitian left sesquilinear complex form and ψ an antisymmetric bilinear complex form on a p-dimensional complex vector space. Then there exists a basis such that the matrices for ϕ and ψ are I_p and Λ respectively, where Λ is defined in (39).

<u>Proof</u>. Bourbaki [9], p. 123. $\square$

Since the matrices for $Re \circ \phi$ and $Re \circ \psi$ transform according to (4) and (67), respectively, an equivalent formulation of the lemma is that for every $H \varepsilon H^+(p, \mathbb{C})$ and $F \varepsilon A(p, \mathbb{C})$, there exists $A \varepsilon GL(p, \mathbb{C})$ such that $AHA' = I_p$ and $AFA' = \Lambda$. This can be reformulated in terms of $4p \times 4p$ real matrices. For

$$T \equiv \begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix} \varepsilon \ H^+(p, \mathbb{C}) \otimes I_2 \quad \text{and} \quad R \equiv \begin{pmatrix} 0 & -F \\ F & 0 \end{pmatrix} \varepsilon \ A(p, \mathbb{C}) \otimes J_1,$$

there exists

$$M \equiv \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \varepsilon \ GL(p, \mathbb{C}) \otimes I_2$$

such that

(69) $$MTM' = I_{4p} \quad \text{and} \quad MRM' = \begin{pmatrix} 0 & 0 & -\Lambda & 0 \\ 0 & 0 & 0 & \Lambda \\ \Lambda & 0 & 0 & 0 \\ 0 & -\Lambda & 0 & 0 \end{pmatrix}.$$

It is seen from (69) that $\pm\lambda_1,\ldots,\pm\lambda_{[p/2]}$ are uniquely determined as the eigenvalues of R w.r.t. T, each with multiplicity four. (When p is odd 0 is always an eigenvalue with multiplicity four.)

For $T = t(S)$ and $R = S - t(S)$ it now follows that there exists an $M \varepsilon GL(p, \mathbb{C}) \otimes I_2$ such that

$$MSM' = \begin{pmatrix} I_p & 0 & -\Lambda & 0 \\ 0 & I_p & 0 & \Lambda \\ \Lambda & 0 & I_p & 0 \\ 0 & -\Lambda & 0 & I_p \end{pmatrix} \tag{70}$$

and that (68) can be represented by

$$\Pi : H^+(p, \mathbb{H}) \to \Lambda_{[p/2]}, \tag{71}$$

where $\Pi(S)$ is the ordered family of non-negative eigenvalues of $S - t(S)$ w.r.t. $t(S)$. As in section 2 it is seen that this representation is also topological.

<u>Theorem 4</u>. The maximum likelihood estimator of Σ under H_0 is $t(S)$ and the likelihood ratio statistic for testing H_0 is

$$\prod_{\gamma=1}^{[p/2]} (1 - \lambda_\gamma^2)^{2N}, \tag{72}$$

where $(\lambda_1,\ldots,\lambda_{[p/2]}) = \Pi(S)$. Under the hypothesis H_0 the statistics $t(S)$ and $\Pi(S)$ are independently distributed. The distribution of $t(S)$ has the density

$$\left(\frac{\det T}{\det \Sigma}\right)^{N/2} \exp(-N/2 \ \mathrm{tr}(\Sigma^{-1}T)), \ T \,\varepsilon\, H^+(2p, \mathbb{C}) \otimes I_2 \tag{73}$$

w.r.t. a unique measure $\nu'_{\mathbb{C},p,N}$ which is invariant under the action of $GL(p,\mathbb{C})\otimes I_2$ on $H^{+}(p,\mathbb{C})\otimes I_2$. The distribution of $\Pi(S)$ has density (72) w.r.t. a measure κ on $\Lambda_{[p/2]}$ which is uniquely defined by

$$(t,\Pi)(\nu_{\mathbb{H},p,N}) = \nu'_{\mathbb{C},p,N} \otimes \kappa.$$

Furthermore, κ has the density

$$(74) \qquad \prod_{1\le\alpha<\beta\le p} (\lambda_\alpha^2-\lambda_\beta^2)^4 \prod_{\gamma=1}^{[p/2]} \lambda_\gamma^{1+4\varepsilon}\,(1-\lambda_\gamma^2)^{-2p+1},$$

where $\varepsilon = p-2[p/2]$, w.r.t. a Lebesgue measure on $\Lambda_{[p/2]}$.

Remark. The maximum likelihood estimator of $\begin{pmatrix}\Gamma & -\Psi_1\\ \Psi_1 & \Gamma\end{pmatrix}$ under H_0 is H. Since the mapping $T=\mathrm{diag}(H,H)\to H$ from $H^{+}(p,\mathbb{C})\otimes I_2$ onto $H^{+}(p,\mathbb{C})$ transforms $\nu'_{\mathbb{C},p,N}$ into $\nu_{\mathbb{C},p,2N}$ it is seen from (73) that H follows a complex Wishart distribution with $2N$ degrees of freedom and parameter $\frac{1}{2N}\begin{pmatrix}\Gamma & -\Psi_1\\ \Psi_1 & \Gamma\end{pmatrix}$.

Proof. If in section 2 one replaces $GL(p,\mathbb{C})$ by $GL(p,\mathbb{C})\otimes I_2$, $H^{+}(2p,\mathbb{R})$ by $H^{+}(p,\mathbb{H})$, $H^{+}(p,\mathbb{C})$ by $H^{+}(p,\mathbb{C})\otimes I_2$, $S(p,\mathbb{C})$ by $A(p,\mathbb{C})\otimes J_1$, $\nu_{\mathbb{R},2p,N}$ by $\nu_{\mathbb{H},p,N}$, $\nu_{\mathbb{C},p,N}$ by $\nu'_{\mathbb{C},p,N}$ and $(\lambda_1,\ldots,\lambda_p)$ by $(\lambda_1,\ldots,\lambda_{[p/2]})$ then the proof is completely analogous to the proof of Theorem 1 except for the following changes: The determinant of the right hand side of (70) is $\Pi(1-\lambda_\gamma^2)^4$, which gives (72). The invariant measure $\nu_{\mathbb{H},p,N}$ has density $|\det S|^{-(p-\frac{1}{2})/2}$ w.r.t. a Lebesgue measure on $H^{+}(p,\mathbb{H})$ (Bourbaki [11], p. 93) which gives the factor $\Pi(1-\lambda_\gamma^2)^{-2p+1}$ in (74). Finally the Jacobian of

$$dR \to \sum_{j=1}^{m} a_j \sum_{k=0}^{2j} R_0^k(dR)R_0^{2j-k},$$

where

$$dR = \begin{pmatrix} 0 & 0 & -C & D \\ 0 & 0 & D & C \\ C & -D & 0 & 0 \\ -D & -C & 0 & 0 \end{pmatrix} \varepsilon\ A(p, \mathbb{C}) \otimes J_1 \text{ and } R_0 = \begin{pmatrix} 0 & 0 & -\Lambda & 0 \\ 0 & 0 & 0 & \Lambda \\ \Lambda & 0 & 0 & 0 \\ 0 & -\Lambda & 0 & 0 \end{pmatrix},$$

is

$$(75) \qquad \prod_{\gamma=1}^{[p/2]} Dr(\lambda_\gamma)\frac{r(\lambda_\gamma)}{\lambda_\gamma} \prod_{1\le\alpha<\beta\le p} \left[\frac{r(\lambda_\alpha)^2 - r(\lambda_\beta)^2}{\lambda_\alpha^2 - \lambda_\beta^2}\right]^4 \prod_{\gamma=1}^{[p/2]} \left(\frac{r(\lambda_\gamma)}{\lambda_\gamma}\right)^{4\varepsilon}$$

when $1 > \lambda_1 > \ldots > \lambda_{[p/2]} > 0$, which gives the remaining factors in (74). The proof of (75) is completely analogous to the proof of (46). $\square$

6. TESTING INDEPENDENCE OF TWO SETS OF VARIATES WHERE THE SIMULTANEOUS COVARIANCE MATRIX HAS REAL, COMPLEX, OR QUATERNION STRUCTURE.

Let $\mathbb{D}$ denote $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$ and set $\delta = \dim_{\mathbb{R}}\mathbb{D}$, i.e., $\delta = 1,2$, or 4. Let $M(p_2, p_1, \mathbb{D})$ denote the set of all $p_2 \times p_1$ real matrices when $\mathbb{D} = \mathbb{R}$; the set of all $2p_2 \times 2p_1$ real matrices of the form (2), where A and B are $p_2 \times p_1$ matrices, when $\mathbb{D} = \mathbb{C}$; and the set of all $4p_2 \times 4p_1$ real matrices of the form (48), where A, B_1, B_2 and B_3 are $p_2 \times p_1$ matrices, when $\mathbb{D} = \mathbb{H}$. Let E_1 and E_2 be right vector spaces over $\mathbb{D}$ of dimensions p_1 and p_2, respectively, $g: E_1 \to E_2$ a $\mathbb{D}$-linear map, and $\psi: E_2 \times E_1 \to \mathbb{D}$ a left sesquilinear $\mathbb{D}$-form. The $\delta p_2 \times \delta p_1$ real matrices of g and $\mathrm{Re} \circ \psi$ w.r.t. bases for E_1 and E_2 considered as vector spaces over $\mathbb{R}$ are both of the form (2) when $\mathbb{D} = \mathbb{C}$ and of the form (48) when $\mathbb{D} = \mathbb{H}$, hence belong to $M(p_2, p_1, \mathbb{D})$. If $f_1: E_1 \to E_1$ and $f_2: E_2 \to E_2$ are $\mathbb{D}$-linear maps, then $\psi \circ (f_2 \times f_1)$ is a left sesquilinear $\mathbb{D}$-form, and since $\mathrm{Re} \circ (\psi \circ (f_2 \times f_1)) = (\mathrm{Re} \circ \psi) \circ (f_2 \times f_1)$, it is seen that

$$M_2 \Psi M_1' \in M(p_2, p_1, \mathbb{D})$$

for $M_1 \in GL(p_1, \mathbb{D})$, $M_2 \in GL(p_2, \mathbb{D})$, and $\Psi \in M(p_2, p_1, \mathbb{D})$.

Let x be an observation from a normal distribution on $\mathbb{R}^{\delta p}$ with covariance matrix $\Sigma \in H^+(p, \mathbb{D})$ and let $p = p_1 + p_2$ with $0 < p_2 \leq p_1$. Partition x into δ p-dimensional subvectors x_α, $\alpha = 1, \ldots, \delta$, corresponding to the partition of Σ into $p \times p$ submatrices given by its "$\mathbb{D}$-structure" (see (3) and (49)). Furthermore, partition each x_α into x_α^1 and x_α^2 consisting of the first p_1 and the last p_2 coordinates of x_α, respectively, $\alpha = 1, \ldots, \delta$. Permuting these 2δ subvectors into the order $x_1^1, \ldots, x_\delta^1, x_1^2, \ldots, x_\delta^2$, one transforms Σ into the form

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \tag{76}$$

where $\Sigma_{11} \in H^+(p_1, \mathbb{D})$, $\Sigma_{22} \in H^+(p_2, \mathbb{D})$, $\Sigma_{21} \in M(p_2, p_1, \mathbb{D})$, and $\Sigma_{12} = \Sigma_{21}'$. The set of positive definite matrices of the form (76) is here denoted $H^+(p_1, p_2, \mathbb{D})$. We shall discuss the problem of testing the independence of $(x_1^1, \ldots, x_\delta^1)$ and $(x_1^2, \ldots, x_\delta^2)$.

Let $x_1, \ldots, x_N$, $N \geq p > 1$, be i.i.d. observations from a normal distribution on $\mathbb{R}^{\delta p}$ with mean vector 0 and unknown covariance matrix $\Sigma \in H^+(p_1, p_2, \mathbb{D})$. The maximum likelihood estimate S of Σ also has the form (76),

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix},$$

and it follows from the preceding sections that S is $\mathbb{D}$-Wishart distributed with N degrees of freedom and parameter $\frac{1}{N}\Sigma$.

The statistical problem under consideration is that of testing the hypothesis H_0 that $\Sigma_{12} = \Sigma_{21}' = 0$. This problem is invariant under the action

$$GL(p_1, \mathbb{D}) \oplus GL(p_2, \mathbb{D}) \times H^+(p_1, p_2, \mathbb{D}) \to H^+(p_1, p_2, \mathbb{D})$$

(77)
$$\left(\begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix}, \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}\right) \to \begin{pmatrix} M_1S_{12}M_1' & M_1S_{12}M_2' \\ M_2S_{21}M_1' & M_2S_{22}M_2' \end{pmatrix},$$

where $GL(p_1, \mathbb{D}) \oplus GL(p_2, \mathbb{D}) = \{\mathrm{diag}(M_1, M_2) \mid M_1 \varepsilon GL(p_1, \mathbb{D}), M_2 \varepsilon GL(p_2, \mathbb{D})\}$. If one permutes the coordinates as above it is seen that $GL(p_1, \mathbb{D}) \oplus GL(p_2, \mathbb{D})$ becomes a subgroup of $GL(p_1 + p_2, \mathbb{D})$ and thus (77) is a restriction of the actions (6), (5), or (51) in the cases $\mathbb{D} = \mathbb{R}, \mathbb{C}$, or $\mathbb{H}$, respectively.

The linear map

$$t: H^+(p_1, p_2, \mathbb{D}) \to H^+(p_1, \mathbb{D}) \oplus H^+(p_2, \mathbb{D})$$

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \to \begin{pmatrix} S_{11} & 0 \\ 0 & S_{22} \end{pmatrix}$$

commutes with the action (77) and the transitive action

(78)
$$GL(p_1, \mathbb{D}) \oplus GL(p_2, \mathbb{D}) \times H^+(p_1, \mathbb{D}) \oplus H^+(p_2, \mathbb{D}) \to H^+(p_1, \mathbb{D}) \oplus H^+(p_2, \mathbb{D})$$
$$(M, T) \to MTM',$$

where $H^+(p_1, \mathbb{D}) \oplus H^+(p_2, \mathbb{D}) = \{\mathrm{diag}(\Sigma_1, \Sigma_2) \mid \Sigma_1 \varepsilon H^+(p_1, \mathbb{D}), \Sigma_2 \varepsilon H^+(p_2, \mathbb{D})\}$.

Lemma 10. Let $\phi_1: E_1 \times E_1 \to \mathbb{D}$ and $\phi_2: E_2 \times E_2 \to \mathbb{D}$ be positive definite hermitian left sesquilinear forms and $g: E_2 \to E_1$ a $\mathbb{D}$-linear map. Then there exist

$\mathbb{D}$-bases of E_1 and E_2 such that the matrices of ϕ_1 and ϕ_2 are I_{p_1} and I_{p_2} respectively, and the matrix of g is

$$\Lambda = \begin{pmatrix} \lambda_1 \ldots 0 & \ldots 0 \\ \ddots \ \vdots & \vdots \\ 0 \ \ldots \vdots \lambda_{p_2} & \ldots 0 \end{pmatrix},$$

where $\lambda_1 \geqq \ldots \geqq \lambda_{p_2} \geqq 0$.

<u>Proof</u>. It follows from Lemma 11 in the next section that there exists a basis $f_1, \ldots, f_{p_2}$ of E_2 such that the matrix of ϕ_2 is I_{p_2} and the matrix of $\phi_1 \circ (g \times g)$ is $\mathrm{diag}(\lambda_1^2, \ldots, \lambda_{p_2}^2)$, where $\lambda_1 \geqq \ldots \geqq \lambda_{p_2} \geqq 0$. Set $e_i = g(f_i)\lambda_i^{-1}$ for $\lambda_i > 0$; these e_i's are orthonormal w.r.t. ϕ_1 and they can therefore be extended to a basis $e_1, \ldots, e_{p_1}$ for E_1 which is orthonormal w.r.t. ϕ_1, i.e., the matrix of ϕ_1 is I_{p_1}. Since $g(f_i) = e_i\lambda_i$, $i = 1, \ldots, p_2$, the lemma follows. $\square$

In terms of real matrices the lemma states that for $\Phi_1 \varepsilon H^+(p_1, \mathbb{D})$, $\Phi_2 \varepsilon H^+(p_2, \mathbb{D})$, and $G \varepsilon M(p_2, p_1, \mathbb{D})$, there exists an $A_1 \varepsilon GL(p_1, \mathbb{D})$ and an $A_2 \varepsilon GL(p_2, \mathbb{D})$ such that $A_1 \Phi_1 A_1' = I_{p_1} \otimes I_\delta$, $A_2 \Phi_2 A_2' = I_{p_2} \otimes I_\delta$, and $(A_2')^{-1} G A_1' = \Lambda \otimes I_\delta$.

For $\Phi_1 = S_{11}$, $\Phi_2 = S_{22}^{-1}$, $G = S_{21}$, $M_1 = A_1$, and $M_2 = (A_2')^{-1}$, one obtains that there exists an $M \varepsilon GL(p\ , \mathbb{D}) \otimes GL(p\ , \mathbb{D})$ such that

$$\text{(79)} \qquad MSM' = \begin{Bmatrix} I_{p_1} \otimes I_\delta & \Lambda' \otimes I_\delta \\ \Lambda \otimes I_\delta & I_{p_2} \otimes I_\delta \end{Bmatrix} .$$

Hence the orbit projection corresponding to the action (77) can be represented by

$$\Pi : H^+(p_1, p_2, \mathbb{D}) \to \Lambda_{p_2} ,$$

where $\Pi(S)$ is the ordered family of the first p_2 non-negative eigenvalues of $S - t(S)$ w.r.t. $t(S)$ each with multiplicity δ. (0 is always an eigenvalue with multiplicity at least $\delta(p_1 - p_2)$).

Theorem 5. The maximum likelihood estimator of Σ under H_0 is $t(S)$ and the likelihood ratio statistic for testing H_0 is

$$\prod_{\gamma=1}^{p_2} (1 - \lambda_\gamma^2)^{\delta N/2} , \tag{80}$$

where $(\lambda_1, \ldots, \lambda_{p_2}) = \Pi(S)$. Under the hypothesis H_0 the statistics $t(S)$ and $\Pi(S)$ are independently distributed. The distribution of $t(S)$ has density

$$\left(\frac{\det T}{\det \Sigma}\right)^{N/2} \exp(-N/2 \operatorname{tr} (\Sigma^{-1} S)),\ S \in H^+(p_1, \mathbb{D}) \oplus H^+(p_2, \mathbb{D}) \tag{81}$$

w.r.t. a unique measure $\nu_{\mathbb{D},p_1,N} \otimes \nu_{\mathbb{D},p_2,N}$ which is invariant under the action of $GL(p_1, \mathbb{D}) \oplus GL(p_2, \mathbb{D})$ on $H^+(p_1, \mathbb{D}) \oplus H^+(p_2, \mathbb{D})$. The distribution of $\Pi(S)$ has density (80) w.r.t. a measure κ on Λ_{p_2} which is uniquely defined by

$$(t, \Pi)(\nu_{\mathbb{D},p,N}) = (\nu_{\mathbb{D},p_1,N} \otimes \nu_{\mathbb{D},p_2,N}) \otimes \kappa.$$

Furthermore, κ has density

$$\prod_{\gamma=1}^{p_2} (1 - \lambda_\gamma^2)^{-\frac{\delta}{2}(p - 1 + \frac{2}{\delta})} \lambda_\gamma^{\delta(p_1 - p_2) + \delta - 1} \prod_{1 \le \alpha < \beta \le p_2} (\lambda_\alpha^2 - \lambda_\beta^2)^\delta \tag{82}$$

w.r.t. a Lebesgue measure on Λ_{p_2}.

Remark. The maximum likelihood estimators for Σ_{11} and Σ_{22} are S_{11} and S_{22}, respectively, and it is seen from (81) that these are independently distributed and that S_{ii} follows a $\mathbb{D}$-Wishart distribution with N degrees of freedom and parameter $\frac{1}{N}\Sigma_{ii}$, $i = 1,2$.

Proof. The proof is analogous to the proof of Theorem 1. The determinant of (79) is $\Pi(1-\lambda_\gamma^2)^\delta$, which gives (80). The invariant measure $\nu_{\mathbb{D},p,N}$ on $H^+(p,\mathbb{D})$ has density $|\det|\, S^{-(p-1+\frac{2}{\delta})/2}$ w.r.t. a Lebesgue measure, and that gives the first factor in (82). The Jacobian of

$$(83) \qquad dR \to \sum_{j=1}^{m} a_j \sum_{k=o}^{2j} R_o^k(dR)R_o^{2j-k} ,$$

where $dR = \begin{pmatrix} 0 & G' \\ G & 0 \end{pmatrix}$, $G \in M(p_1, p_2, \mathbb{D})$ and $R_o = \begin{pmatrix} 0 & \Lambda' \otimes I_\delta \\ \Lambda \otimes I_\delta & 0 \end{pmatrix}$, is

$$(84) \qquad \prod_{\gamma=1}^{p_2} Dr(\lambda_\gamma)\left(\frac{r(\lambda_\gamma)}{\lambda_\gamma}\right)^{\delta-1+\delta(p_1-p_2)} \prod_{1\le\alpha<\beta\le p_2}\left(\frac{r(\lambda_\alpha)^2-r(\lambda_\beta)^2}{\lambda_\alpha^2-\lambda_\beta^2}\right)^\delta$$

when $1 > \lambda_1 > \ldots > \lambda_{p_2} > 0$, which gives the remaining factors in (82).

We shall indicate the proof of (84) in the case $\mathbb{D} = \mathbb{H}$: Recall that G is of the form (48). Under the mapping (83) the (γ,γ)'th element of A is multiplied by $Dr(\lambda_\gamma)$, the (γ,γ)'th elements of B_1, B_2, and B_3 are multiplied by $r(\lambda_\gamma)/\lambda_\gamma$, and for $\beta > p_2$ the (γ,β)'th elements of A, B_1, B_2, and B_3 are multiplied by $r(\lambda_\gamma)/\lambda_\gamma$. For $1 \leqq \alpha < \beta \leqq p_2$ the pair of elements $(t_{\alpha\beta}, t_{\beta\alpha})$ of A, B_1, B_2, and B_3 are mapped into the pair $(at_{\alpha\beta} + bt_{\beta\alpha}, at_{\beta\alpha} + bt_{\alpha\beta})$, where a and b are as in the proof of Theorem 2; this mapping has the determinant

$$(a^2-b^2)^\delta = [(r(\lambda_\alpha)^2 - r(\lambda_\beta)^2)/(\lambda_\alpha^2-\lambda_\beta^2)]^\delta . \qquad \square$$

7. TESTING THE HYPOTHESIS THAT TWO COVARIANCE MATRICES WITH REAL, COMPLEX, OR QUATERNION STRUCTURE ARE IDENTICAL.

As in section 6 we denote by $\mathbb{D}$ either $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and $\delta = \dim_{\mathbb{R}} \mathbb{D}$. Let S_1 and S_2 be independent observations from $\mathbb{D}$-Wishart distributions with N_1 and N_2 degrees of freedom and parameters $\Sigma_1, \Sigma_2 \in H^+(p, \mathbb{D})$ respectively, $p \leq N_1$ and $p \leq N_2$. Since $\nu_{\mathbb{D},p,N}$ is invariant and $\det(NS) = N^{\delta p} \det(S)$ for $S \in H^+(p, \mathbb{D})$, it follows from the preceding sections that the distribution of (S_1, S_2) has density

$$\left(\frac{\det S_1}{\det \Sigma_1}\right)^{N_1/2} \cdot \left(\frac{\det S_1}{\det \Sigma_1}\right)^{N_2/2} \exp(-1/2 \operatorname{tr}(\Sigma_1^{-1} S_1 + \Sigma_2^{-1} S_2)), (S_1, S_2) \in H^+(p, \mathbb{D})^2$$

w.r.t. $\nu^o_{\mathbb{D},p,N_1} \otimes \nu^o_{\mathbb{D},p,N_2}$, where $\nu^o_{\mathbb{D},p,N_i} = (1/N_i)^{\delta p N_i/2} \nu_{\mathbb{D},p,N}$, $p \leq N_i$, $i=1,2$.

Let H_o denote the hypothesis that $\Sigma_1 = \Sigma_2$. The statistical problem of testing H_o is invariant under the action

$$(85) \qquad \begin{aligned} GL(p, \mathbb{D}) \times H^+(p, \mathbb{D})^2 &\to H^+(p, \mathbb{D})^2 \\ (M, (S_1, S_2)) &\to (MS_1M', M S_2M'). \end{aligned}$$

The linear map

$$\begin{aligned} t: H^+(p, \mathbb{D})^2 &\to H^+(p, \mathbb{D}) \\ (S_1, S_2) &\to S_1 + S_2 \end{aligned}$$

commutes with the actions (85) and (6), (5), or (51), respectively.

Lemma 11. Let ϕ be a positive definite hermitian left sesquilinear $\mathbb{D}$-form and ψ a hermitian left sesquilinear $\mathbb{D}$-form on a p-dimensional vector space over $\mathbb{D}$. Then there exists a basis such that the matrix of ϕ is I_p and the matrix of ψ is $\Lambda=\mathrm{diag}(\lambda_1,\ldots,\lambda_p)$, where $\lambda_1 \geqq \ldots \geqq \lambda_p$.

Proof. Bourbaki [9], p. 123. □

In terms of $\delta p \times \delta p$ real matrices the lemma implies that there exists an $M \varepsilon GL(p, \mathbb{D})$ such that $M(S_1+S_2)M' = I_{\delta p}$, and $MS_1M' = \Lambda \otimes I_\delta$. Then $MS_2\, M' = (I - \Lambda) \otimes I_\delta$, and it is seen that the orbit projection corresponding to the action (85) can be represented by

$$\Pi : H^+(p, \mathbb{D})^2 \to \Lambda_p,$$

where $\Pi(S_1, S_2) = (\lambda_1,\ldots,\lambda_p)$ and $1 > \lambda_1 \geq \ldots \geq \lambda_p > 0$ are the eigenvalues of S_1 w.r.t. $S_1 + S_2$, each with multiplicity δ.

Theorem 6. The maximum likelihood estimator for Σ (the common value of Σ_1 and Σ_2) under H_o is $\frac{1}{(N_1+N_2)}\, t(S_1, S_2)$ and the likelihood ratio test statistic for testing H_o is

$$(86) \qquad \prod_{\gamma=1}^{p} \lambda_\gamma^{\delta N_1/2}(1-\lambda_\gamma)^{\delta N_2/2} \quad ,$$

where $(\lambda_1,\ldots,\lambda_p) = \Pi(S_1, S_2)$. Under the hypothesis H_o the statistics $t(S_1, S_2)$ and $\Pi(S_1, S_2)$ are independently distributed. The distribution of $t(S_1, S_2)$ has density

$$\left(\frac{\det T}{\det \Sigma}\right)^{(N_1+N_2)/2} \exp(-1/2 \;\mathrm{tr}\; \Sigma^{-1} T), \; T \,\varepsilon\, H^+(p, \mathbb{D})$$

w.r.t. $\nu^{o}_{\mathbb{D},p,N_1+N_2}$. The distribution of $\Pi(S_1, S_2)$ has density (86) w.r.t. a measure κ on Λ_p which is uniquely defined by

$$(t,\Pi)(\nu^{o}_{\mathbb{D},p,N_1} \otimes \nu^{o}_{\mathbb{D},p,N_2}) = \nu^{o}_{\mathbb{D},p,N_1+N_2} \otimes \kappa.$$

Furthermore, κ has density

$$(87) \qquad \prod_{\gamma=1}^{p} (\lambda_\gamma(1-\lambda_\gamma))^{-\delta(p-1+\frac{2}{\delta})/2} \prod_{1\le\alpha<\beta\le p} (\lambda_\alpha - \lambda_\beta)^{\delta}$$

w.r.t. a Lebesgue measure on Λ_p.

<u>Proof</u>. The proof is analogous to the proofs for the other theorems. For (87), one defines $\tilde{r}(S_1, S_2) = (r(S_1, S_1+S_2), S_1+S_2 - r(S_1, S_1+S_2))$, where

$$r(R,T) = \sum_{j=1}^{m} a_j (RT^{-1})^{2j} R.$$

The determinant of

$$dR \to \sum_{j=1}^{m} a_j \sum_{k=o}^{2j} R_o^{k}(dR)R_o^{2j-k},$$

where $dR \in H^{+}(p, \mathbb{D})$ and $R_o = \Lambda \otimes I_\delta$, is

$$\prod_{\gamma=1}^{p} Dr(\lambda_\gamma) \prod_{1\le\alpha<\beta\le 1} \left(\frac{r(\lambda_\alpha)-r(\lambda_\beta)}{\lambda_\alpha-\lambda_\beta}\right)^{\delta}$$

when $1 > \lambda_1 > \ldots > \lambda_p > 0$. $\square$

8. THE NORMING CONSTANTS.

In Theorems 1-6 we have represented the distribution of the orbit projection Π by a density w.r.t. a Lebesgue measure on Λ_q for appropriate q. This density is the product of the likelihood ratio statistic g and the density f of a measure κ with respect to that Lebesgue measure. (In Theorem 1, for example, the density of Π is the product of (18) and (21).) In order to find the density of Π w.r.t. the *usual* Lebesgue measure $d\mu(\lambda) \equiv \Pi d\lambda_i$, one must evaluate the norming constant $(\int g(\lambda)f(\lambda)d\mu(\lambda))^{-1}$. We shall first evaluate the norming constants associated with Theorems 5 and 6 by a simultaneous recursion argument. The remaining norming constants are easily obtained from these.

We shall use the same notation as in sections 6 and 7. The distribution of Π in Theorem 6 has density

$$(88) \qquad b(\delta,p,N_1,N_2) \prod_{\gamma=1}^{p} \lambda_\gamma^{\delta(N_1-m)/2} (1-\lambda_\gamma)^{\delta(N_2-m)/2} \prod_{1\le\alpha<\beta\le p} (\lambda_\alpha-\lambda_\beta)^{\delta}$$

w.r.t. μ, where $m = p-1+\frac{2}{\delta}$ and $b(\delta,p,N_1,N_2)$ is the norming constant. In fact, b is defined for all real values of $N_1, N_2 \in [m, \infty)$. Let $\nu_{\mathbb{D},p}$ be an arbitrary invariant measure on $H^+(p,\mathbb{D})$. Then $\nu^o_{\mathbb{D},p,N} = c(\delta,p,N)\,\nu_{\mathbb{D},p}$, where $c(\delta,p,N)$ is a constant. Since $(t,\Pi)(\nu^o_{\mathbb{D},p,N_1} \otimes \nu^o_{\mathbb{D},p,N_2}) = \nu^o_{\mathbb{D},p,N_1+N_2} \otimes b(\delta,p,N_1,N_2)\, f\mu$, where f is given by (87), we obtain $(t,\Pi)(\nu_{\mathbb{D},p} \otimes \nu_{\mathbb{D},p}) = \nu_{\mathbb{D},p} \otimes \left[\dfrac{b(\delta,p,N_1,N_2)\;c(\delta,p,N_1+N_2)}{c(\delta,p,N_1)\quad c(\delta,p,N_2)}\right] f\mu$. It follows that

$$(89) \qquad b(\delta,p,N_1,N_2) = \frac{c(\delta,p,N_1)c(\delta,p,N_2)}{c(\delta,p,N_1+N_2)}\; k_1(\delta,p),$$

where $k_1(\delta,p)$ is a constant only depending on δ and p.

In Theorem 5, Π has density

$$(90) \qquad a(\delta,p_1,p_2,N) \prod_{\gamma=1}^{p_2} \lambda_\gamma^{\delta(p_1-p_2+1)-1} (1-\lambda_\gamma^2)^{\delta(N-p+1)/2-1} \prod_{1\le\alpha<\beta\le p_2} (\lambda_\alpha^2-\lambda_\beta^2)^\delta$$

w.r.t. μ, where $p = p_1 + p_2$ and $a(\delta,p_1,p_2,N)$ is the norming constant. In the same manner as in the derivation of (89), it is shown that

$$(91) \qquad a(\delta,p_1,p_2,N) = \frac{c(\delta,p_1+p_2,N)}{c(\delta,p_1,N)c(\delta,p_2,N)} k_2(\delta,p_1,p_2),$$

where $k_2(\delta,p_1,p_2)$ is a constant only depending on δ,p_1, and p_2. (Recall that $\nu_{\mathbb{D},p,N} = N^{\delta p N/2} \nu^o_{\mathbb{D},p,N} = N^{\delta p N/2} c(\delta,p,N)\nu_{\mathbb{D},p}$.)

Since $a(\delta,p,1,N) = 2\Gamma(\delta N/2)/(\Gamma(\delta p/2)\Gamma(\delta(N-p)/2))$ and $c(\delta,1,N) = c_o(\delta/2)^{\delta N/2}/\Gamma(\delta N/2)$, where c_o is a constant which depends on $\nu_{\mathbb{D},1}$, it follows by induction from (91) that

$$(92) \qquad c(\delta,p,N) = k_o(\delta,p)(\delta/2)^{\delta Np/2}/d(\delta,p,N),$$

where

$$(93) \qquad d(\delta,p,N) = \prod_{\gamma=1}^{p} \Gamma(\delta(N-\gamma+1)/2)$$

and $k_o(\delta,p)$ is a constant only depending on δ and p.

If (92) is substituted into (89) one obtains

$$(94) \qquad b(\delta,p,N_1,N_2) = \frac{k(\delta,p)d(\delta,p,N)}{d(\delta,p,N_1)d(\delta,p,N_2)} ,$$

where $k(\delta,p) = k_o(\delta,p)k_1(\delta,p)$.

Since both sides of (94) are rational functions in N_1 and N_2 it follows that (94) in fact holds for all real values of $N_1, N_2 \in [m, \infty)$. (It then follows that $b(p,N_1,N_2)$ is defined for all real values of $N_1,N_2 \in (p-1, \infty)$.)

By making the substitutions $\lambda_1 = y_1$ and $\lambda_\gamma = y_1 y_\gamma$, $\gamma = 2\ldots,p$, in (88) and integrating over y_1 it is seen that $b(\delta,p,p-1+\frac{2}{\delta},p-1+\frac{2}{\delta})$ $= p\left(\frac{(p-1)}{2} + 1\right) b\left(\delta,p-1,p-2+\frac{2}{\delta}, p+\frac{2}{\delta}\right)$. If (94) is substituted into this expression one obtains after a reduction that $k(\delta,p) = k(\delta,p-1)\Gamma(\delta/2)/\Gamma(\delta p/2)$, and it follows that $k(\delta,p) = k_3(\delta)[\Gamma(\delta/2)]^p/d(p,p)$, where $k_3(\delta)$ is a constant only depending on δ. Since $b(\delta,1,N_1,N_2) = \Gamma(\delta(N_1+N_2)/2)/(\Gamma(\delta N_1/2)\Gamma(\delta N_2/2)$, it follows that $k_3(\delta) \equiv 1$, so that

$$(95) \qquad b(\delta,p,N_1,N_2) = \frac{[\Gamma(\delta/2)]^p d(\delta,p,N_1+N_2)}{d(\delta,p,p)d(\delta,p,N_1)d(\delta,p,N_2)} .$$

Finally, by making the substitution $y_\gamma = \lambda_\gamma^2$, $\gamma=1,\ldots,p_2$, in (90), one obtains that

$$(96) \qquad a(\delta,p_1,p_2,N) \equiv 2^{p_2}b(\delta,p_2,p_1,N-p_1)$$

and that $a(\delta,p_1,p_2,N)$ in fact is defined for all real values of $p_1 \in (p_2-1,\infty)$ and all real values of $N \in (p_1+p_2-1,\infty)$.

It is now a simple matter to obtain the norming constants associated with Theorems 1 - 4. These are:

$$a(1,p+1,p,N+1) = 2^p b(1,p,p+1,N-p),\ 2p \le N, \tag{97}$$

$$a(2,p-[\tfrac{p}{2}] - \tfrac{1}{2},\ [\tfrac{p}{2}],\ N - \tfrac{1}{2}) = 2^{[\frac{p}{2}]} b(2,[\tfrac{p}{2}],p-[\tfrac{p}{2}] - \tfrac{1}{2}, N-p + [\tfrac{p}{2}]),\ p \le N, \tag{98}$$

$$a(2,p+\tfrac{1}{2},p,N-\tfrac{1}{2}) = 2^p b(2,p+\tfrac{1}{2}, p, N-p),\ 2p \le N, \tag{99}$$

$$a(4,p - [\tfrac{p}{2}] - \tfrac{1}{2}, [\tfrac{p}{2}],\ N - \tfrac{1}{2}) = 2^{[\frac{p}{2}]} b(4,[\tfrac{p}{2}],\ p-[\tfrac{p}{2}] - \tfrac{1}{2}, N-p+[\tfrac{p}{2}]),\ p \le N. \tag{100}$$

For each of the ten testing problems the α'th moment, $\alpha \ge 0$, of the likelihood ratio test statistic is readily obtained from these norming constants. For the problem associated with Theorem 1 one obtains

$$\frac{a(1,\ p+1,\ p,\ N+1)}{a(1,\ p+1,\ p,\ N\alpha+N+1)} = \frac{d(1,\ p,\ N+1)d(1,\ p,\ N\alpha+N-p)}{d(1,\ p,\ N\alpha+N+1)d(1,\ p,\ N-p)} =$$

$$\prod_{\gamma=1}^{p} \frac{\Gamma((N-\gamma+2)/2)\Gamma((N\alpha+N-p-\gamma+1)/2)}{\Gamma((N\alpha+N-\gamma+2)/2)\Gamma((N-p-\gamma+1)/2)}, \tag{101}$$

and similarly for the problems associated with Theorems 2 - 6

$$\prod_{\gamma=1}^{[\frac{p}{2}]} \frac{\Gamma(N-\gamma+\frac{1}{2})\Gamma(N\alpha+N-p+[\frac{p}{2}]-\gamma+1)}{\Gamma(N\alpha+N-\gamma+\frac{1}{2})\Gamma(N-p+[\frac{p}{2}]-\gamma+1)} \tag{102}$$

$$\prod_{\gamma=1}^{p} \frac{\Gamma(N-\gamma+\frac{1}{2})\Gamma(N\alpha+N-p-\gamma)}{\Gamma(N\alpha+N-\gamma+\frac{1}{2})\Gamma(N-p-\gamma)} \tag{103}$$

$$(104)\qquad \prod_{\gamma=1}^{[\frac{p}{2}]} \frac{\Gamma(2N - 2\gamma+1)\Gamma(2(N\alpha+N-p+[\frac{p}{2}]-\gamma+1))}{\Gamma(2N\alpha + 2N - 2\gamma+1)\Gamma(2(N-p+[\frac{p}{2}]-\gamma+1))}$$

$$(105)\qquad \prod_{\gamma=1}^{p_2} \frac{\Gamma(\delta(N-\gamma+1)/2)\Gamma(\delta(N\alpha+N-p_1-\gamma+1)/2)}{\Gamma(\delta(N\alpha+N-\gamma+1)/2)\Gamma(\delta(N-p_1-\gamma+1)/2)}$$

and

$$(106)\qquad \prod_{\gamma=1}^{p} \frac{\Gamma(\delta(N_1+N_2-\gamma+1)/2)\Gamma(\delta(N_1\alpha+N_1-\gamma+1)/2)\Gamma(\delta(N_2\alpha+N_2-\gamma+1)/2)}{\Gamma(\delta(N_1\alpha+N_2\alpha+N_1+N_2-\gamma+1)/2)\Gamma(\delta(N_1-\gamma+1)/2)\Gamma(\delta(N_2-\gamma+1)/2)} .$$

Acknowledgement: We wish to thank M.D. Perlman for his help in improving the presentation of this paper.

REFERENCES

[1] Andersson, S.A. (1975). Invariant normal models. *Ann. Statist.* **3**, 132-154.

[2] Andersson, S.A. (1978). Invariant measures, Technical Report 129. Stanford University, Department of Statistics, Stanford, California.

[3] Andersson, S.A. (1978). Canonical correlations with respect to a complex structure, Preprint 13. Institute of Mathematical Statistics, University of Copenhagen.

[4] Andersson, S.A. (1981). Distributions of maximal invariants using quotient measures, Technical Report No. 7. Department of Statistics, University of Washington, Seattle.

[5] Andersson, S.A., Brøns, H.K., Jensen, S.T. (1982). An algebraic theory for normal statistical models. In preparation.

[6] Andersson, S.A., Perlman, M.D. (1982). Two testing problems relating the real and complex multivariate normal distribution. To appear in *J. Multivariate Analysis*.

[7] Andersson, S.A., Perlman, M.D. (1982). Two testing problems relating the complex and quarternion multivariate normal distribution. In preparation.

[8] Anderson, T.W. (1958). *An Introduction to Multivariate Statistical Analysis*, Wiley, New York.

[9] Bourbaki, N. (1959). Éléments de Mathématique. Algèbre, Chap. 9. Hermann, Paris.

[10] Bourbaki, N. (1960). Éléments de Mathématique. Topologie general, Chap. 3 á 4. Hermann, Paris.

[11] Bourbaki, N. (1963). Éléments de Mathématique. Integration. Chap. 7 á 8. Hermann, Paris.

[12] Fisher, R.A. (1939). The sampling distribution of some statistics obtained from non-linear equations, *Ann. Engin.* **9**, 238-249.

[13] Gabrielsen, G. (1975). Fordeling af egenvaerdier i reelle, komplekse og kvaternion - Wishart - fordelinger. Thesis in Danish. Institute of Mathematical Statistics, University of Copenhagen.

[14] Hsu, P.L. (1939). On the distribution of roots of certain determinantal equations. *Ann. Engin.* **9**, 250-258.

[15] Khatri, C.G. (1965). Classical statistical analysis based on certain multivariate complex distributions. *Ann. Math. Statist.* **36**, 98-114.

[16] Khatri, C.G. (1965). A test for reality of covariance matrix in a certain complex Gaussian Distribution. *Ann. Math. Statist.* **36**, 115-119.

[17] Mehta, M.L. (1967). *Random Matrices and the Statistical Theory of Energy Levels*. Academic Press, New York.

[18] Porter, C.E., (Ed.), (1965). *Statistical Theories of Spectra: Fluctuations*. Academic Press, New York and London.

[19] Roy, S.N. (1939). p-statistics or some generalizations in analysis of variance appropriate to multivariate problems. *Sankhya* **4**, 381-396.

PREPRINTS 1981

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Johansen, Søren: Asymptotic Inference in Random Coefficient Regression Models.

No. 2 Asmussen, Søren: On the Role of a Certain Eigenvalue in Estimating the Growth Rate of a Branching Process.

No. 3 Lauritzen, Steffen.: Time Series Analysis in 1880. A Discussion of Contributions made by T.N. Thiele.

No. 4 Asmussen, Søren: Conditioned Limit Theorems Relating a Random Walk to its Associate, with Applications to Risk Reserve Processes and the GI/G/1 Queue.

No. 5 Johansen, Søren: The Statistical Analysis of a Markov Branching Process.

No. 6 Asmussen, Søren: Time - Dependent Approximations in some Queueing Systems with Imbedded Markov Chains Related to Random Walks.

No. 7 Skovgaard, Ib M.: A Second - Order Investigation of Asymptotic Ancillarity.

No. 8 Rootzén, Holger: The Rate of Extremes of Stationary Normal Sequences.

No. 9 Skovgaard, Ib M.: Large Deviation Approximations for Maximum Likelihood Estimators.

No. 10 Jensen, Ulla Funck: A Stochastic Projection Model with Implications for Multistate Demography and Manpower Analysis.

No. 11 Johansen, Søren: An Extension of Cox's Regression Model.

No. 12 Skovgaard, Ib M.: Edgeworth Expansions in Statistics.

PREPRINTS 1982

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Holmgaard, Simon and Yu, Song Yu: Gaussian Markov Random Fields Applied to Image Segmentation.

No. 2 Andersson, Steen A., Brøns, Hans K. and Jensen, Søren Tolver: Distribution of Eigenvalues in Multivariate Statistical Analysis.