NORMAL MULTIVARIATE ANALYSIS AND THE ORTHOGONAL GROUP[1]

By A. T. James

Princeton University

1. Summary. New methods are introduced for deriving the sampling distributions of statistics obtained from a normal multivariate population. Exterior differential forms are used to represent the invariant measures on the orthogonal group and the Grassmann and Stiefel manifolds. The first part is devoted to a mathematical exposition of these. In the second part, the theory is applied; first, to the derivation of the distribution of the canonical correlation coefficients when the corresponding population parameters are zero; and secondly, to split the distribution of a normal multivariate sample into three independent distributions, (a) essentially the Wishart distribution, (b) the invariant distribution of a random plane which is given by the invariant measure on the Grassmann manifold, (c) the invariant distribution of a random orthogonal matrix. This decomposition provides derivations of the Wishart distribution and of the distribution of the latent roots of the sample variance covariance matrix when the population roots are equal.

2. Introduction. Much of the distribution theory of normal multivariate analysis can be deduced from, or is closely related to the fact that the distribution of a normal multivariate sample is invariant under orthogonal transformations.

Consider a set of n independent observations from a normal k-variate distribution ($n \geqq k$) with a nonsingular variance covariance matrix Σ. In most distribution problems one can eliminate the population means with the loss of 1 degree of freedom by a suitable orthogonal transformation. Assume this has already been done. Let the rows of the $n \times k$ matrix X be independent observations from a normal k-variate distribution with zero means;

$$dF(X) = \frac{|\Sigma|^{-\frac{1}{2}n}}{(2\pi)^{\frac{1}{2}nk}} e^{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}X'X)} \Pi\, dx_{ij}. \tag{2.1}$$

The distribution is clearly invariant under the transformation

$$H: X \to HX \tag{2.2}$$

where H is an $n \times n$ orthogonal matrix. The invariance is a fundamental property of dF, indeed, as Bartlett [1] has proved for the univariate case, and a similar proof holds for the multivariate case, the invariance under (2.2) together with

Received 10/16/49, revised 6/25/53.

[1] This work was initiated in the Section of Mathematical Statistics, Commonwealth Scientific and Industrial Research Organization, Australia and completed during the tenure of a C. S. I. R. O. studentship at Princeton University and subsequently with support from the Office of Naval Research.

the independence of the rows of X uniquely characterizes the distribution (2.1) of X.

With probability one, the columns of X, regarded as vectors in n-dimensional Euclidean space R^n, span a k-dimensional linear subspace (henceforth called k-plane). Hotelling [8] observed that the invariance of the distribution of X under the group of transformations (2.2) implies that the k-plane is invariantly distributed. (A formal proof of this result will be given in Section 6.) He also recognized that the problem of finding the distribution of the canonical correlation coefficients could be reduced to the problem of finding the distribution of the cosines of the critical angles between the plane spanned by the columns of X and a plane distributed independently of X, or a fixed plane. From these observations Hotelling went on to obtain the distribution of the canonical correlation coefficients for the special case of two canonical correlation coefficients (assuming the population correlations are zero). The general distribution was later derived by Fisher [5], Hsu [9], Roy [16], Girshick [6] and Mood [13], using different methods.

To complete the derivation of the general result along the lines followed by Hotelling, one requires a convenient analytic expression to represent the invariant distribution of a random plane. Such an expression would also be very useful in other connections. The most obvious way to obtain such an expression would be as follows. A k-plane in R^n can be specified by a system of $k(n - k)$ parameters, in fact, in many ways. The parameters will then have a certain distribution in $R^{k(n-k)}$ corresponding to the invariant distribution of the random plane. However, such methods lead to intractable expressions because, as we shall see later on, they destroy the symmetry of the space of k-planes.

Instead, we consider the k-planes in R^n as points of a space which is an analytic manifold, the *Grassmann manifold.* Blaschke [2] has constructed an exterior differential form on the Grassmann manifold which may be considered as the probability density for an invariantly distributed random plane. By a simple transformation the exterior differential form may be expressed in terms of the critical angles, and hence the distribution of the canonical correlation coefficients obtained. This will be carried out in Section 7.

Another analytic manifold, important in multivariate analysis, is the *Stiefel manifold.* A set of k orthonormal vectors in R^n is called a *k-frame.* The k-frames are the points of the Stiefel manifold. Both the Grassmann and the Stiefel manifolds are coset spaces of the orthogonal group which is also an analytic manifold.

The theory of Grassmann and Stiefel manifolds, exterior differential forms, etc. used in this derivation, is familiar to the differential geometer; but its literature is widely scattered, and not readily accessible to the statistician unless he is prepared to go far more deeply into these subjects than is required here. It therefore seems desirable to give an outline of those parts of the theory that we require, in a form suitable for immediate application to problems of multivariate statistics. Sections 3 to 5 are devoted to this.

Exterior differential forms on manifolds have evolved from a simple rule for transforming multiple and surface integrals in Euclidean space. It is based on an anticommutative multiplication of differentials (see Goursat [7] and Kähler [11]). As Chern [3] has pointed out, it has potential application in statistics. Although it is equivalent to the calculation of the Jacobian, it is usually simpler because it avoids the necessity of explicitly writing out bulky determinants.

Section 3.1 gives the definition of an analytic manifold. In Sections 3.2 to 3.4 the three analytic manifolds to be considered in this paper, namely the orthogonal group and the Grassmann and Stiefel manifolds, are defined and the relationship of the Grassmann and Stiefel manifolds as coset spaces of the orthogonal group is explained.

Exterior differential forms are introduced in Section 4.1 and their integrals defined in Section 4.2. The transformation of them is discussed in Section 4.3 and it is shown how an invariant differential form yields an invariant measure.

The exterior differential forms representing the invariant measures on the orthogonal group and the Grassmann and Stiefel manifolds are constructed in Sections 4.5 to 4.7 and their integrals are evaluated in Sections 5.1 to 5.3.

Sections 6 and 7 give the derivation of the distribution of the canonical correlation coefficients, as outlined above, and in Section 8 the results stated in the summary on the decomposition of the distribution of a normal multivariate sample, are proved. Olkin [14] has given this decomposition by what amounts to using parameters for the Grassmann and Stiefel manifolds based on the Cayley parameters for the orthogonal group.

3. The orthogonal group and its coset spaces.

3.1. *Analytic manifolds.* An n-dimensional *manifold*, $\mathfrak{M}$, is a Hausdorff topological space in which every point p has a neighbourhood $\mathfrak{O}_p$ with a system of coordinates $x_1^p, \cdots, x_n^p$, that is, such that the map $r \leftrightarrow x_1^p, \cdots, x_n^p$ $(r \in \mathfrak{O}_p)$ is a one-to-one bicontinuous map (homeomorphism) of $\mathfrak{O}_p$ on an open set in real Euclidean space, R^n. The coordinates $x_1^p, \cdots, x_n^p$ will be referred to as coordinates *centred* at p. They are also coordinates *centred* at any other point of $\mathfrak{O}_p$.

If $x_1^p, \cdots, x_n^p$ and $x_1^q, \cdots, x_n^q$ are the coordinates of a point $r \in \mathfrak{O}_p \cap \mathfrak{O}_q$ relative to coordinate systems centred at p and q respectively, then since the correspondences

$$x_1^p, \cdots, x_n^p \leftrightarrow r \leftrightarrow x_1^q, \cdots, x_n^q$$

are homeomorphisms, it follows that $x_1^p, \cdots, x_n^p$ and $x_1^q, \cdots, x_n^q$ are continuous (single valued) functions of each other. A manifold, together with a set of overlapping coordinate systems, which cover the entire manifold and have the property that the transformation between any two overlapping coordinate systems is analytic, is called an *analytic*[2] manifold. (A function defined on R^n

[2] It would be sufficient for the applications in this paper only to assume that the functions have continuous derivatives, thus defining *differentiable manifolds*. But as we are applying the theory to the orthogonal group, the Grassmann and the Stiefel manifolds which are not only differentiable but indeed analytic, we may as well assume analyticity.

is called *analytic* in a domain if it can be expanded as a convergent multiple power series in the neighbourhood of any point of that domain.) The systems of coordinates possessing the required properties are called *admissible*.

A familiar example of an analytic manifold is the surface of a unit sphere in Euclidean space, for example, in R^3. A system of coordinates, centred at any point p of the sphere, can be obtained by taking the orthogonal projection of the open hemisphere with p as pole on the plane tangent to the sphere at p. This is obviously a homeomorphism of the open hemisphere on the interior of the unit circle in the tangent plane. Introduce coordinate axes in the tangent plane and let (x_1^p, x_2^p) be the coordinates of the projection on the tangent plane of a point r in the hemisphere. Then (x_1^p, x_2^p) serve as admissible coordinates for r. The transformations between two such coordinate systems centred at p and q respectively can be shown to be analytic in their domain of overlap.

More generally, the construction of admissible coordinate systems by projection on the tangent plane can be applied to show that any algebraic variety which has a tangent plane at every point, is an analytic manifold. (An algebraic variety is a surface in Euclidean space determined by a system of algebraic equations). In particular, the orthogonal group and Stiefel manifold, which we shall now discuss, are analytic manifolds.

DEFINITION. A function f defined on an analytic manifold is an *analytic function* in the domain $\mathfrak{D}$ if, for any arbitrary coordinates $x_1, \cdots, x_n$ admissible in a domain $\mathfrak{O}$, f is an analytic function of $x_1, \cdots, x_n$ in the domain $\mathfrak{D} \cap \mathfrak{O}$.

3.2. *The orthogonal group* $O(n)$. An $n \times n$ matrix, A, satisfying the equation $A'A = I_n$ where I_n is the identity matrix and A' means the transpose of A, is called an orthogonal matrix. An equivalent definition is that A is the matrix of a linear transformation which leaves the quadratic form $x_1^2 + \cdots + x_n^2$ invariant. The set of all $n \times n$ orthogonal matrices with the operation of matrix multiplication is called the *orthogonal group*, $O(n)$.

There are $\frac{1}{2}n(n + 1)$ functionally independent conditions on the n^2 elements of an orthogonal matrix $A \; \varepsilon \; O(n)$; consequently, the elements of A can be regarded as the coordinates of a point on a $\frac{1}{2}n(n - 1)$-dimensional algebraic variety or surface in Euclidean n^2-space. Since $\Sigma_{i,j}a_{ij}^2 = n$, the group surface is a subset of the sphere of radius $\sqrt{n}$ in n^2-space.

In 1896, Hurwitz [10] pointed out that the element of area of the group surface is a two-sided invariant measure on $O(n)$, that is, invariant under *left* and *right translations*, by which we mean that the respective transformations

$$A \to HA \qquad (3.1)$$

$$A \to AH \qquad (3.2)$$

$$H \; \varepsilon \; O(n)$$

leave the element of area invariant. For, suppose X is an $n \times n$ matrix, regarded as a vector in an n^2 dimensional space, and transformed by $H \; \varepsilon \; O(n)$;

$$X \to HX \quad \text{or} \quad X \to XH. \qquad (3.3)$$

These are linear transformations of the n^2-space which leave the quadratic form $\Sigma_{i,j} x_{ij}^2$ invariant. Therefore they are orthogonal transformations (of order $n^2 \times n^2$). But the area of a surface in Euclidean space is invariant under orthogonal transformations. Hence the area of the group surface in n^2-space is invariant under (3.1) and (3.2).

The invariant measure is sometimes referred to as the "Haar" measure on the orthogonal group, named after Haar, who, in 1933, generalized Hurwitz's result by proving the existence of an invariant measure on any locally compact topological group. Herglotz and Blaschke [2] have derived an exterior differential form for the invariant measure on $O(n)$ which is, (apart from a scale factor), a convenient expression for the area of the surface. We shall derive it later on.

3.3. *The Stiefel manifold* $V_{k,n}$. Let us call a set of k orthonormal vectors in Euclidean n-space, a "*k-frame*". The k-frames are the points of the *Stiefel manifold*, $V_{k,n}$. Regarding the k vectors of a k-frame as the columns of a matrix A, we can represent the Stiefel manifold as the set of $n \times k$, $(k \leq n)$, rectangular matrices, A, satisfying the equation $A'A = I_k$. $V_{k,n}$ is an $\frac{1}{2}k(2n - k - 1)$-dimensional algebraic variety in nk-dimensional Euclidean space and an analytic manifold. The same argument as for the orthogonal group shows that the element of area of this surface is a measure invariant under (3.1).

A group of transformations is said to act *transitively* in a space if, given any two points of the space, there is an element of the group which transforms one into the other. Such a space is said to be *homogeneous* with respect to the group. If x_0 is any point of a homogeneous space $\mathfrak{X}$ (with respect to a group $\mathfrak{H}$) and $\mathfrak{H}_0$ is the subgroup consisting of all elements of $\mathfrak{H}$ which leave x_0 invariant, and if $h \,\varepsilon\, \mathfrak{H}$ transforms x_0 into x, then the set of all elements of $\mathfrak{H}$ which transform x_0 into x is the coset $h\mathfrak{H}_0$. Hence the points $x \,\varepsilon\, \mathfrak{X}$ are in one-to-one correspondence with the cosets $h\mathfrak{H}_0$. Thus a space, homogeneous with respect to a group of transformations, may be regarded as a space of cosets of the group.

The Stiefel manifold is obviously homogeneous with respect to the orthogonal group of transformations acting on $V_{k,n}$ according to (3.1). If $A_0 \,\varepsilon\, V_{k,n}$ say for simplicity

$$A_0 = \begin{bmatrix} I_k \\ \hline 0 \end{bmatrix}$$

then the group O_0 which leaves A_0 invariant is the set of square matrices of the form

$$\begin{bmatrix} I_k & 0 \\ 0 & H_{n-k} \end{bmatrix}$$

where H_{n-k} is any $n - k \times n - k$ orthogonal matrix and I_k is the unit matrix of order k.

Hence the coset corresponding to $A \,\varepsilon\, V_{k,n}$ is

$$[A \mid B] \begin{bmatrix} I_k & 0 \\ 0 & O(n-k) \end{bmatrix}$$

where B is any $n \times n - k$ matrix such that the partitioned matrix $[A \vdots B]$ is orthogonal and $O(n - k)$ is the group of orthogonal matrices of order $n - k$.[3]

3.4. *The Grassmann manifold.* The points of the *Grassmann manifold*, $G_{k,r}$, $(r = n - k)$ are the k-dimensional planes (passing through the origin) in Euclidean n-space, R^n. For our purposes, the following obviously equivalent definition is useful. Consider the set, $\mathfrak{X}$, of all $n \times k$ matrices $(k \leqq n)$ of rank k;

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1k} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nk} \end{bmatrix}$$

and the group of transformations $X \to XL$ where L is any nonsingular $k \times k$ matrix. The group defines an equivalence relation in $\mathfrak{X}$. Two elements of $\mathfrak{X}$ are equivalent if there is an element of the group which transforms one into the other. Such is possible if and only if the column vectors of the two matrices span the same k-plane in the Euclidean n-space, R^n, of column vectors. Hence the equivalence classes of $\mathfrak{X}$ are in one-to-one correspondence with the points of the Grassmann manifold, $G_{k,r}$.

$G_{k,r}$ is an analytic manifold. It has dimension $k(n - k)$, because X may be regarded as a point in Euclidean nk-space and for each fixed X the set of all elements XL in the equivalence class is a surface in R^{nk} of dimension k^2. Hence the dimension of $G_{k,r}$ is $nk - k^2 = k(n - k)$.

Like the Stiefel manifold, the Grassmann manifold can be regarded as a coset space of the orthogonal group $O(n)$. An orthogonal transformation of R^n transforms k-planes into k-planes; thus it *induces* a transformation of $G_{k,r}$. We shall use the same symbol for the induced transformation of $G_{k,r}$ as for the original transformation of R^n. In this sense, the orthogonal group $O(n)$ is a transitive group of transformations of $G_{k,r}$ because, given any two k-planes in R^n, there exists a rotation which transforms one into the other.

If $\mathfrak{p}_0$ is any fixed point of $G_{k,r}$ and O_0 is the subgroup of all elements of $O(n)$ which leave $\mathfrak{p}_0$ invariant, and if $H \varepsilon O(n)$ transforms $\mathfrak{p}_0$ into $\mathfrak{p} \varepsilon G_{k,r}$, then the set of all elements of $O(n)$ which transform $\mathfrak{p}_0$ into $\mathfrak{p}$ are the elements of the coset HO_0. Hence the cosets HO_0, $H \varepsilon O(n)$, are in one-to-one correspondence with the points $\mathfrak{p} \varepsilon G_{k,r}$.

Suppose, for simplicity, we let $\mathfrak{p}_0$ be the plane spanned by the first k coordinate axes. The cosets are then of the form

$$[A \mid B] \begin{bmatrix} O(k) & 0 \\ 0 & O(n-k) \end{bmatrix} \tag{3.4}$$

where the first k columns of the matrix on the left are orthonormal vectors spanning the plane $\mathfrak{p}$ and the last $n - k$ columns are likewise orthonormal vectors, but they span the orthogonal complement of $\mathfrak{p}$. The matrix is thus an

[3] The orthogonal group manifold can be expressed as a *fibre bundle* with the Stiefel manifold as the *base space* and the subgroup $O(n - k)$ as the *fibre*. Steenrod [18] discusses the Stiefel and Grassmann manifolds from this point of view.

element of $O(n)$. The matrix on the right denotes the subgroup O_0 of $O(n)$ consisting of all matrices which are a direct sum of orthogonal matrices of orders k and $n - k$ respectively. O_0 is the subgroup which leaves the plane spanned by the first k coordinate axes invariant.

4. Exterior differential forms on manifolds.

4.1. *Definition.* Consider a multiple integral over a domain Δ in Euclidean space R^n;

$$k = \int_\Delta f(x_1, \cdots, x_n)\, dx_1 \cdots dx_n. \tag{4.1}$$

On making a change of variables

$$\begin{aligned} x_1 &= x_1(u_1, \cdots, u_n) \\ &\vdots \\ x_n &= x_n(u_1, \cdots, u_n) \end{aligned} \tag{4.2}$$

we have

$$k = \int_{\Delta'} f(x(u)) \det\left(\frac{\partial x_i}{\partial u_j}\right) du_1 \cdots du_n. \tag{4.3}$$

To calculate the Jacobian, instead of writing out the matrix of partial derivatives $(\partial x_i / \partial u_j)$ and calculating its determinant, we can evaluate it in the following way. Differentiate the transformations (4.2);

$$dx_i = \sum_{j=1}^{n} \frac{\partial x_i}{\partial u_j} du_j \tag{4.4}$$

and substitute the linear differential forms (4.4) in (4.1;)

$$k = \int_{\Delta'} f(x(u)) \left(\sum_j \frac{\partial x_i}{\partial u_j} du_j\right) \cdots \left(\sum_j \frac{\partial x_n}{\partial u_j} du_j\right). \tag{4.5}$$

Now multiply out the differential forms in (4.5) in a formal manner using the associative and distributive laws, but instead of the commutative law use an anticommutative rule for multiplying differentials; that is, put

$$du_j\, du_i = -du_i\, du_j. \tag{4.6}$$

In particular $-du_i\, du_i = du_i\, du_i = 0$.

The justification for this formal procedure is that the rules are consistent and lead to the correct result as given in (4.3) (see Goursat [7] chap. 3). In fact, the formal procedure is equivalent to calculating the Jacobian as is shown by the following

LEMMA 4.1. *If du is a column vector of n differentials and if $dx = J\, du$, where J is an $n \times n$ matrix and thus dx is a column vector of linear differential, forms, then the anticommutative product of the elements of dx is the anticommutative product of the elements of du multiplied by $|J|$; that is*

$$\prod_{i=1}^{n} dx_i = |J| \prod_{\nu=1}^{n} du_\nu . \tag{4.7}$$

Proof. The left-hand side of (4.7) is clearly equal to $\prod_{\nu=1}^{n} du_\nu$ multiplied by a polynomial $p(J)$ in the elements of J, which is linear in each row of J. Interchanging the order of two factors, say dx_i and dx_j, reverses the sign of $\prod_{i=1}^{n} dx_i$. However, it is also equivalent to interchanging the ith and jth rows of the matrix J. Thus interchange of two rows of J reverses the sign of $p(J)$. Finally, if J is the identity matrix then $p(J) = 1$. Hence, according to the Weierstrass definition of a determinant, $p(J) = |J|$. The formal procedure may also be used to transform surface integrals.

An exterior differential form of degree r in Euclidean space R^n is a formal expression of the type

$$\sum_{i_1<i_2<\cdots<i_r} u_{i_1\cdots i_r}(x)\, dx_{i_1} \cdots dx_{i_r} \tag{4.8}$$

where $u_{i_1\cdots i_r}(x)$ are analytic functions of $x_1, \cdots, x_n$. It may be regarded as the integrand of an r-dimensional surface integral. The exterior product of a form of degree r with a form of degree s is defined as the form of degree $r + s$ which is obtained by formal multiplication of the two forms using the associative, distributive, and anticommutative laws for the multiplication of the symbols dx_i.

A form of degree n has only one term, namely $u(x)\, dx_1 \cdots dx_n$. A form of degree greater than n is zero because it has at least one of the symbols dx_i repeated in each term.

The definition may be extended to define an exterior differential form on an analytic manifold $\mathfrak{M}$. Relative to a system of admissible coordinates on the manifold, it is an expression of the type (4.8).

Definition. An *exterior differential form* $\omega(p)$, $p \in \mathfrak{M}$, on an analytic manifold is a system of expressions of type (4.8), one for each admissible coordinate system, such that if $x_1, \cdots, x_n$ and $y_1, \cdots, y_n$ are two coordinate systems, then in the domain of overlap of these coordinates the corresponding expressions of type (4.8) with coefficients $u_{i_1\cdots i_r}$ and $v_{i_1\cdots i_r}$ respectively, are related by the transformation

$$\sum_{i_1<\cdots<i_r} u_{i_1\cdots i_r}(x(y)) \left(\Sigma \frac{\partial x_{i_1}}{\partial y_j} dy_j\right) \cdots \left(\Sigma \frac{\partial x_{i_r}}{\partial y_j} dy_j\right) = \Sigma v_{i_1\cdots i_r}(y)\, dy_{i_1} \cdots dy_{i_r} . \tag{4.9}$$

The exterior product of two exterior differential forms ω and v is defined as the exterior differential form whose representation in any coordinate system is the product of the representations of ω and v in that coordinate system. One can check that the resulting product transforms according to the rule (4.9) on change of coordinates. Hence it constitutes an exterior differential form on the manifold.

If p_0 is a fixed point of $\mathfrak{M}$ with coordinates x^0, we shall call the expression

$$\omega(p_0) = \Sigma u_{i_1\cdots i_r}(x^0)\, dx_{i_1} \cdots dx_{i_r}$$

the *value* of $\omega(p)$ at p_0. The values

$$u_1(x^0)\, dx_1 + \cdots + u_n(x^0)\, dx_n$$

of linear differential forms at a fixed point p_0 of the manifold can be considered as vectors in an n dimensional vector space with $dx_1, \cdots, dx_n$ as its basic vectors. This space is called the *tangent space* to the manifold at p_0. It is the analogue of the tangent plane to a surface in Euclidean space.

DEFINITION. The *differential* of an analytic function f on a manifold is the linear exterior differential form represented in coordinates $x_1, \cdots, x_n$ by

$$df = \frac{\partial f}{\partial x_1}\, dx_1 + \cdots + \frac{\partial f}{\partial x_n}\, dx_n .$$

The ordinary rules of calculus show that such expressions transform correctly, that is, in accordance with (4.9).

4.2. *Integration of exterior differential forms.* It is, of course, possible to define the integral of a differential form of degree r over a submanifold of dimension r or a measurable subset thereof. However, for our applications we only require the integrals of differential forms of maximum degree, that is, of degree equal to the dimension of the manifold, n, and we shall restrict our definition to these.

Expressed in coordinates, an exterior differential form of maximum degree has only one term

$$U(x^p)\, dx_1^p \cdots dx_n^p . \tag{4.10}$$

As our integrals will be interpreted as probabilities, we require that, if $U(x^p)$ is not positive, it be replaced by its modulus. The domain of integration is divided into subdomains or cells C_i each contained in an admissible coordinate system. The admissible coordinates map the cell C_i into a domain $\tilde{C}_i$ in R^n and the differential form, expressed in these coordinates, is then regarded as the integrand of an ordinary volume integral over the domain $\tilde{C}_i$ in R^n and evaluated as such. Thus to integrate (4.10) over a subdomain C_i in an admissible coordinate system, we take the multiple integral

$$\int_{\tilde{C}_i} |\, U(x_1^p, \cdots, x_n^p)\, dx_1^p \cdots dx_n^p \,| . \tag{4.11}$$

The sum of the integrals over the (finite number of) subdomains, into which the total domain of integration has been divided, is defined as the integral of the differential form. The integral of the differential form does not depend on the subdivision of the domain of integration; that is, if a portion of the domain of integration has two admissible coordinate systems, say x_i^p and x_i^q, they give the same integral, namely

$$\int |\, U(x^p)\, dx_1^p \cdots dx_n^p \,| = \int \left|\, U(x^q) \det\left(\frac{\partial x_i^p}{\partial x_j^q}\right) dx_1^q \cdots dx_n^q \,\right|$$

according to the classical formula for the transformation of multiple integrals in Euclidean space.

Corresponding to an exterior differential form $\omega(p)$ of maximum degree on a manifold $\mathfrak{X}(p \varepsilon \mathfrak{X})$, there is a measure μ given by

$$\mu(\mathfrak{S}) = \int_{\mathfrak{S}} \omega(p) \qquad \mathfrak{S} \subset \mathfrak{X}. \tag{4.11a}$$

In the general theory of integration of exterior differential forms, a difficulty arises as to which sign should be assigned to the integrals (4.11) over the component domains before taking their sum. It is connected with the orientation of the domains. However, as we are only going to integrate exterior differential forms representing probability densities, we have been able to avoid the difficulty by defining only positive integrals. Changing the sign of an exterior differential form does not alter its integral, as defined above. Hence we may ignore the sign of an exterior differential form of maximum degree and ignore questions of orientability of the manifolds.

To integrate a function on the manifold with respect to the differential form, express it as a function of the coordinates $x_1^p, \cdots, x_n^p$ and include it under the integral sign in (4.11).

4.3. *Transformation of exterior differential forms.* A one-to-one map of a manifold $\mathfrak{X}$ on another $\mathfrak{Y}$, induces maps of the functions on $\mathfrak{X}$ to functions on $\mathfrak{Y}$, measures on $\mathfrak{X}$ to measures on $\mathfrak{Y}$ and differential forms on $\mathfrak{X}$ to differential forms on $\mathfrak{Y}$.

Suppose f is an analytic homeomorphism of the analytic manifold $\mathfrak{X}$ on another $\mathfrak{Y}$. By f being an *analytic homeomorphism*, we mean that the map f is one-to-one, and if $x_1, \cdots, x_n$ are coordinates of $p \varepsilon \mathfrak{X}$ and $y_1, \cdots, y_n$ are coordinates of the image point $q = f(p) \varepsilon \mathfrak{Y}$, then the y_i are analytic functions of $x_1, \cdots, x_n$ and the x_i are analytic functions of $y_1, \cdots, y_n$. Since f is one-to-one, q and p are functions of each other. Put $p = f^{-1}(q)$.

Definition. If $\varphi(p)$ is an analytic function on $\mathfrak{X}$, f *induces* a mapping of it to a function $\Phi(q)$ on $\mathfrak{Y}$ given by

$$\varphi(p) \xrightarrow{f} \Phi(q) = \varphi(f^{-1}(q)). \tag{4.12}$$

Definition. If μ is a measure on $\mathfrak{X}$, f *induces* a mapping of it to a measure $\bar{\mu}$ on $\mathfrak{Y}$ given by

$$\mu \xrightarrow{f} \bar{\mu} \tag{4.13}$$

where

$$\bar{\mu}(\mathfrak{T}) = \mu(f^{-1}(\mathfrak{T})) \qquad \mathfrak{T} \subset \mathfrak{Y}$$

where $f^{-1}(\mathfrak{T})$ denotes the *inverse image* of $\mathfrak{T}$, that is, the set of all points of $\mathfrak{X}$ mapped into $\mathfrak{T}$ by f.

Definition. The image of a differential form under the *mapping induced by* f

is obtained by replacing dx_i by $\Sigma_j(\partial x_i)/(\partial y_j)\, dy_j$, the coefficient functions by their images under map f, and using the rules for exterior products. Thus

$$\sum_{i_1,\cdots i_r} u_{i_1\cdots i_r}(x)\, dx_{i_1} \cdots dx_{i_r} \xrightarrow{f} \sum_{i_1,\cdots i_r} u_{i_1\cdots i_r}(f^{-1}(y)) \left(\sum_j \frac{\partial x_{i_1}}{\partial y_j} dy_j\right) \cdots \left(\sum_j \frac{\partial x_{i_r}}{\partial y_j} dy_j\right). \tag{4.14}$$

This mapping could be carried out using arbitrary coordinate systems $x_1, \cdots, x_n$ and $y_1, \cdots, y_n$ in $\mathfrak{X}$ and in $\mathfrak{Y}$ respectively. It can easily be shown that mappings carried out in different coordinates agree with one another. From the definition of the mapping, it is clear that the map of a product of two differential forms is the product of their maps. The map of a differential form of maximum degree gives the map of the corresponding measure, according to the

LEMMA 4.2. *If μ is a measure on $\mathfrak{X}$ given by the differential form $\omega(p)$,*

$$\mu(\mathfrak{S}) = \int_{\mathfrak{S}} \omega(p) \qquad \mathfrak{S} \subset \mathfrak{X}, \tag{4.15}$$

and f is an analytic homeomorphism of $\mathfrak{X}$ on $\mathfrak{Y}$ which induces maps of μ on $\bar{\mu}$ and $\omega(p)$ on $\bar{\omega}(q)$, then the measure $\bar{\mu}$ is given by the differential form $\bar{\omega}(q)$:

$$\bar{\mu}(\mathfrak{T}) = \int_{\mathfrak{T}} \bar{\omega}(q) \qquad \mathfrak{T} \subset \mathfrak{Y} \tag{4.16}$$

PROOF. By definition

$$\bar{\mu}(\mathfrak{T}) = \mu(f^{-1}(\mathfrak{T})) = \int_{f^{-1}(\mathfrak{T})} \omega(p). \tag{4.17}$$

Suppose $\omega(p) = \varphi(x)\, dx_1 \cdots dx_n$. Then

$$\begin{aligned} \bar{\omega}(q) &= \varphi(f^{-1}(y)) \left(\sum_j \frac{\partial x_1}{\partial y_j} dy_j\right) \cdots \left(\sum_j \frac{\partial x_n}{\partial y_j} dy_j\right) \\ &= \varphi(f^{-1}(y)) \det\left(\frac{\partial x_i}{\partial y_j}\right) dy_1 \cdots dy_n, \end{aligned} \tag{4.18}$$

and hence

$$\int_{f^{-1}(\mathfrak{T})} \omega(p) = \int_{\mathfrak{T}} \bar{\omega}(q) \tag{4.19}$$

by the classical formula for the transformation of multiple integrals. (4.17) and (4.19) imply (4.16). Q.E.D.

An important case is the mapping of a manifold $\mathfrak{X}$ on itself by an analytic homeomorphism f. If f maps a set $\mathfrak{S} \subset \mathfrak{X}$ onto a set $\mathfrak{T} = f(\mathfrak{S}) \subset \mathfrak{X}$ and an exterior differential form ω into $\bar{\omega}$, then (4.19) holds. The differential form ω is said to be *invariant* under f if $\bar{\omega}(q) = \omega(q)$. In this case, by (4.19), we have

$$\int_{f^{-1}(\mathfrak{T})} \omega(p) = \int_{\mathfrak{T}} \bar{\omega}(q) = \int_{\mathfrak{T}} \omega(q). \tag{4.20}$$

A measure μ is said to be *invariant* under a transformation f if $\mu(f^{-1}(\mathfrak{T})) = \mu(\mathfrak{T})$. (4.20) states that if a differential form is invariant, then the corresponding measure is invariant.

While we have used admissible coordinate systems in defining and establishing the fundamental properties of exterior differential forms, in the practical applications we shall avoid them as they are usually complicated and difficult to handle. Indeed, the main reason for introducing exterior differential forms in this paper is to deal with measures on manifolds without the explicit use of admissible coordinates. The map of a differential form has been defined by use of coordinates, but the following lemma enables us to map them without coordinates.

LEMMA 4.3. *Let f be an analytic homeomorphism of an analytic manifold $\mathfrak{X}$ upon another $\mathfrak{Y}$, which induces a map of an analytic function $\varphi(p)$ on $\mathfrak{X}$ to a function $\Phi(q)$ on $\mathfrak{Y}$. Then the differential form $d\varphi(p)$ is mapped on $d\Phi(q)$.*

PROOF. Let $x_1, \cdots, x_n$ be coordinates for p and $y_1, \cdots, y_n$ coordinates for q. From the definition of a map of a differential form

$$dx_i \xrightarrow{f} \sum_j \frac{\partial x_i}{\partial y_j} dy_j$$

and thus

$$d\varphi = \sum_i \frac{\partial \varphi}{\partial x_i} dx_i \xrightarrow{f} \sum_{i,j} \frac{\partial \varphi}{\partial x_i} \frac{\partial x_i}{\partial y_j} dy_j = \sum_j \frac{\partial \Phi}{\partial y_j} dy_j = d\Phi. \quad \text{Q.E.D.} \tag{4.21}$$

The exterior differential forms representing the measures that we require on the manifolds will be constructed in the following way. The differential of an analytic function (see Section 4.1) on an n-dimensional manifold is a linear differential form and so are linear combinations of differentials of functions, the coefficients being analytic functions on the manifold. The exterior product of n such linear differential forms is a differential form of maximum degree and thus represents a measure.

In this way we shall construct invariant measures on the orthogonal group, and the Grassmann, and the Stiefel manifolds. The invariance characterizes them uniquely (up to a multiplicative constant) according to the

THEOREM 4.1. *If $\mathfrak{X}$ is a topological space and $\mathfrak{H}$ is a transitive compact topological group of transformations of $\mathfrak{X}$ onto itself such that HX is a continuous function of H and X into $\mathfrak{X}$, then there exists a finite measure μ on $\mathfrak{X}$ invariant under $\mathfrak{H}$. μ is unique in the sense that any other invariant measure on $\mathfrak{X}$ is a constant finite multiple of μ.*

For our special applications we prove the existence of such invariant measures by actually constructing them. For a proof of uniqueness see Weil [19]. Chevalley [4] gives an account of analytic manifolds and exterior differential forms on them in chapters 3 and 5, but from a more advanced and abstract standpoint.

4.4. *Repeated integrals.* The topological product $\mathfrak{M} \times \mathfrak{N}$ of two analytic manifolks $\mathfrak{M}$ and $\mathfrak{N}$ is an analytic manifold. If $x_1, \cdots, x_m$ and $y_1, \cdots, y_n$ are admissible coordinates of points $r \varepsilon \mathfrak{M}$ and $s \varepsilon \mathfrak{N}$ in domains $\mathfrak{D}_1 \subset \mathfrak{M}$ and $\mathfrak{D}_2 \subset \mathfrak{N}$, then $x_1, \cdots, x_m, y_1, \cdots, y_n$ is an admissible coordinate system in $\mathfrak{D}_1 \times \mathfrak{D}_2 \subset \mathfrak{M} \times \mathfrak{N}$. Given differential forms $\omega_1(r)$ on $\mathfrak{M}$ and $\omega_2(s)$ on $\mathfrak{N}$ and a function $f(r, s)$ on $\mathfrak{M} \times \mathfrak{N}$, then the exterior product of ω_1 and ω_2 is a differential form on $\mathfrak{M} \times \mathfrak{N}$ and we have

$$\int_{\mathfrak{D}_1 \times \mathfrak{D}_2} f(r, s)\omega_1(r)\omega_2(s) = \int_{\mathfrak{D}_2} \omega_2(s) \int_{\mathfrak{D}_1} f(r, s)\omega_1(r),$$

for this reduces to the classical formula for repeated integrals when f and the differential forms are expressed in terms of the coordinates in $\mathfrak{D}_1$ and $\mathfrak{D}_2$. A similar result holds for the whole manifold or any subdomain $\Delta \subset \mathfrak{M} \times \mathfrak{N}$,

$$\int_{\Delta} f(r, s)\omega_1(r)\omega_2(s) = \int_{\Delta_2} \omega_2(s) \int_{\Delta_1(s)} f(r, s)\omega_1(r),$$

because Δ can be approximated by a union of product sets $\mathfrak{D}_1 \times \mathfrak{D}_2$. $\Delta_1(s) \subset \mathfrak{M}$ is the section of Δ at s and Δ_2 is the projection of Δ on $\mathfrak{N}$.

The importance, for us, of this result is that a manifold, which may be Euclidean space, is often homeomorphic to a topological product of manifolds, apart from a set of measure zero perhaps. And a differential form when transformed to a differential form on the product manifold often splits into the exterior product of differential forms on the component manifolds. Such transformations are useful for evaluating integrals of exterior differential forms and for deriving sampling distributions.

4.5. *The differential form for the invariant measure on the orthogonal group.* Let A be an orthogonal matrix;

$$A'A = I_n. \tag{4.22}$$

To keep the notation clear, we introduce an abstract group manifold isomorphic to the group of orthogonal matrices, and denote its elements by Greek letters. We then regard the elements of our orthogonal matrices as functions on this abstract group manifold. Indeed, they are analytic functions. Let $A(\alpha)$ be the orthogonal matrix corresponding to the abstract group element α. However, the symbol H will be used to denote both a fixed orthogonal matrix and the corresponding element of the abstract group; $A(H\alpha) = HA(\alpha)$.

The differential of a vector or a matrix (such as $A(\alpha)$) whose elements are analytic functions on the group manifold, is defined as the vector or matrix of differentials of the elements. Regarding A as a function of α and differentiating (4.22) we have

$$(dA(\alpha))'A(\alpha) + (A(\alpha))' dA(\alpha) = 0.$$

Thus $(A' dA)' = -A' dA$. Hence $A' dA$ is a skew symmetric matrix of linear differential forms. The exterior product of the super diagonal elements gives us a differential form

$$\omega(\alpha) = \prod_{i<j}^{n} (a_i' \, da_j) = \prod_{i<j}^{n} (a_{1i} \, da_{1j} + \cdots + a_{ni} \, da_{nj}) \tag{4.23}$$

where $a_i = a_i(\alpha)$ is the ith column vector of the matrix $A(\alpha)$. The differential form is of degree $\frac{1}{2}n(n-1) = N$ and is thus of maximum degree. Hence it defines a measure μ on $O(n)$ given by

$$\mu(\mathfrak{S}) = \int_{\mathfrak{S}} \prod a_i' \, da_j \qquad \mathfrak{S} \subset O(n). \tag{4.24}$$

THEOREM 4.2. *The differential forms* $a_i' \, da_j$ *and* $\omega(\alpha)$ *are invariant under the transformation*

$$\alpha \to \tilde{\alpha} = H\alpha \tag{4.25}$$

or equivalently

$$A(\alpha) \to A(\tilde{\alpha}) = HA(\alpha).$$

This transformation is called a left translation.

PROOF. Applying the definition given in (4.12) to the individual elements of the column vector $a_j(\alpha)$, we see that the map (4.25) induces a map of $a_j(\alpha)$ given by

$$a_j(\alpha) \to \hat{a}_j(\tilde{\alpha}) \tag{4.26}$$

where the elements of the column vector $\hat{a}_j(\tilde{\alpha})$ are functions of $\tilde{\alpha}$ defined by the equation $\hat{a}_j(\tilde{\alpha}) = a_j(H^{-1}\tilde{\alpha})$ and this equals $H'a_j(\tilde{\alpha})$. By Lemma 4.3 applied to the elements of $da_j(\alpha)$, we have

$$da_j(\alpha) \to d\hat{a}_j(\tilde{\alpha}) = d(H'a_j(\tilde{\alpha})) = H'da_j(\tilde{\alpha}). \tag{4.27}$$

Hence

$$\begin{aligned} a_i(\alpha)' \, da_j(\alpha) &\to \hat{a}_i(\tilde{\alpha})' \, d\hat{a}_j(\tilde{\alpha}) \\ &= a_i(\tilde{\alpha})'HH' \, da_j(\tilde{\alpha}) \\ &= a_i(\tilde{\alpha})' \, da_j(\tilde{\alpha}), \end{aligned} \tag{4.28}$$

and this is the value of the differential form $a_i(\alpha)' \, da_j(\alpha)$ at $\tilde{\alpha}$, that is, the differential form $a_i' \, da_j$ is invariant under (4.25).

Since the transform of the product of differential forms is the product of their transforms, it follows that $\omega(\alpha)$ is invariant. Q.E.D.

THEOREM 4.3. $\omega(\alpha)$ *is invariant under a right-translation*

$$\alpha \to \tilde{\alpha} = \alpha H. \tag{4.29}$$

PROOF. The transform of the matrix $A(\alpha)' \, dA(\alpha)$ of differential forms is calculated as follows:

$$A(\alpha) \to \hat{A}(\tilde{\alpha}) = A(\tilde{\alpha}H^{-1}) = A(\tilde{\alpha})H',$$

$$dA(\alpha) \to dA(\tilde{\alpha})H'$$

and

$$A(\alpha)'\, dA(\alpha) \to HA(\tilde{\alpha})'\, dA(\tilde{\alpha})H'.$$

The exterior product of the super diagonal elements of the matrix $HA(\tilde{\alpha})'\, dA(\tilde{\alpha})H'$ will be the transform of $\omega(\alpha)$. To evaluate it, consider the transformation

$$A(\tilde{\alpha})'\, dA(\tilde{\alpha}) \to HA(\tilde{\alpha})' dA(\tilde{\alpha})H'.$$

This is a linear transformation of the $N = \frac{1}{2}n(n-1)$ linear differential forms $a_i(\tilde{\alpha})'\, da_j(\tilde{\alpha})$. If a vector of linear differential forms undergoes a linear transformation, then by Lemma 4.1, the exterior product of them is multiplied by the determinant of the linear transformation. To complete the proof we require the

LEMMA 4.4. *If S is a skew symmetric matrix which we regard as a point in an $N = \frac{1}{2}n(n-1)$ dimensional vector space and if L is a fixed matrix, then the transformation*

$$S \to LSL' \tag{4.30}$$

is a linear transformation of the vector space whose determinant is a power of the determinant of L.

PROOF OF LEMMA. The determinant of the linear transformation (4.30) is a polynomial, say $p(L)$, in the elements of L. The transformation L_1L_2 is the same as L_2 and L_1 carried out successively. Therefore, by the multiplication theorem for determinants (applied to the Nth order determinants)

$$p(L_1L_2) = p(L_1)p(L_2). \tag{4.31}$$

But a polynomial $p(L)$ in the elements of a matrix L which satisfies the equation (4.31) for all matrices L_1 and L_2 is a power of the determinant of L (see MacDuffee [12] chap. 3). Therefore $p(L)$ is a power of $|L|$. Q.E.D.

Applying the lemma to the proof of the theorem, we see that the exterior product of the super-diagonal elements of $HA(\tilde{\alpha})'\, dA(\tilde{\alpha})H'$ is the product of the super-diagonal elements of $A(\tilde{\alpha})'\, dA(\tilde{\alpha})$ multiplied by some power of $|H|$, which is 1 apart from sign. Hence the transform of $\omega(\alpha)$ equals its value at $\tilde{\alpha}$; that is, $\omega(\alpha)$ is right-invariant. Q.E.D.

Since $\omega(A)$ is invariant under left and right translations, so is the measure μ which is given by (4.24).

LEMMA 4.5. *μ is invariant under the transformation $A \to A^{-1} = A'$.*

PROOF. Putting

$$\mu(\mathfrak{X}) = \int_{\mathfrak{X}} \omega(A), \qquad \mathfrak{X} \subset O(n)$$

introduce a new measure ν given by $\nu(\mathfrak{X}) = \mu(\mathfrak{X}^{-1})$ where by $\mathfrak{X}^{-1}$ we mean the set of elements of the orthogonal group whose inverses are in $\mathfrak{X}$. Then under the transformation $A \to HA$ we have $A' \to A'H'$. Thus when $\mathfrak{X}$ undergoes a left translation, $\mathfrak{X}^{-1}$ undergoes a right translation. But μ is invariant under right translations. Therefore ν is invariant under left translations. From the uniqueness of invariant measures, ν must be equal to μ apart from a multiplicative constant, which will be unity since $\nu(O(n)) = \mu(O(n)^{-1}) = \mu(O(n))$. The result holds for any compact group.

Of course it is necessary to prove that the invariant differential forms which we construct are not identically zero. This will become evident when we obtain their integrals over the whole space.

As an illustration, let us consider the invariant measure on the proper orthogonal group for the case $n = 2$.

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \qquad 0 \leqq \theta < 2\pi$$

and

$$\omega(A) = a_1' \, da_2 = \cos\theta \, d(\sin\theta) - \sin\theta \, d(\cos\theta) = d\theta.$$

4.6. *The invariant measure on the Grassmann manifold.* In the case of the orthogonal group the differential form for the invariant measure was given by a single expression defined on the whole manifold. For the Grassmann and Stiefel manifolds this is not possible. Instead we construct a system of differential forms each defined locally. Their domains of definition together cover the whole manifold, and wherever they overlap, the differential forms are equal. The *system* of local differential forms is then regarded as a *single* global differential form. It represents the invariant measure.

$G_{k,r}$ is the space of k-planes $\mathfrak{p}$ in R^n, as defined in Section 3.4. For points $\mathfrak{p}$ in a neighbourhood of a point $\mathfrak{p}_0 \varepsilon G_{k,r}$ let $a_1, \cdots, a_k$ and $b_1, \cdots, b_{n-k}$ be orthonormal column vectors spanning the plane $\mathfrak{p}$ and its orthogonal complement respectively, such that the elements of these vectors are all analytic functions of $\mathfrak{p}$. Such a system of vectors can only be constructed locally. The invariant measure is given by the differential form

$$\prod_{j=1}^{n-k} \prod_{i=1}^{k} b_j' \, da_i \tag{4.32}$$

in the domain where the vectors a_i and b_j are defined. The system of all such expressions is a global differential form, denoted by $v_k^n(\mathfrak{p})$, which represents the invariant measure on the Grassmann manifold.

There are three things to be proved; (1) that vectors such as a_i and b_j can be constructed in the neighbourhood of any point $\mathfrak{p}_0 \varepsilon G_{k,r}$; (2) that the differential form (4.32) does not depend on the choice of the a_i and b_j, that is, that any two expressions of type (4.32) are equal wherever their domains of definition overlap; (3) that v_k^n is invariant under the transformations of $G_{k,r}$ induced by orthogonal transformations of R^n.

(1) Take n fixed linearly independent column vectors in R^n the first k of which span $\mathfrak{p}_0$. Take the orthogonal projections of the first k vectors on $\mathfrak{p}$ and of the remaining $n - k$ vectors on the orthogonal complement of $\mathfrak{p}$. For $\mathfrak{p} \varepsilon \mathfrak{D}_{\mathfrak{p}_0}$,

where $\mathfrak{O}_{\mathfrak{p}_0}$ is the domain[4] of $G_{k,r}$ for which the set of projections are linearly independent, we can orthonormalize the projections on $\mathfrak{p}$ by the Gram-Schmidt process to give orthonormal column vectors $a_1, \cdots, a_k$ in $\mathfrak{p}$ and orthonormalize the projections on the orthogonal complement of $\mathfrak{p}$ to give orthonormal vectors $b_1, \cdots, b_{n-k}$ in the orthogonal complement of $\mathfrak{p}$. Then $a_1, \cdots, a_k, b_1, \cdots, b_{n-k}$ are the required set of orthonormal vectors spanning $\mathfrak{p}$ and its orthogonal complement respectively.

(2) Suppose that in a domain $\mathfrak{D} \subset G_{k,r}$ there are two expressions like (4.32) constructed from vectors a_i, b_j and $\tilde{a}_i$, $\tilde{b}_j$ respectively. Let A, B, $\tilde{A}$, $\tilde{B}$ be the respective matrices with these vectors as their columns. Then there exist orthogonal matrices H_1 and H_2 of respective orders k and $n - k$ whose elements are analytic functions of $\mathfrak{p} \varepsilon \mathfrak{D}$ such that

$$\tilde{A} = AH_1 \tag{4.33}$$

$$\tilde{B} = BH_2 . \tag{4.34}$$

We can carry out the transformations (4.33) and (4.34) in two stages.

Differentiate (4.33), $d\tilde{A} = (dA)H_1 + A\, dH_1$. Premultiply by B', $B'\, d\tilde{A} = (B'\, dA)H_1 + 0$ since the columns of A and B are orthogonal. Hence by Lemma 4.1

$$\textstyle\prod_{i=1}^{k} b_j'\, d\tilde{a}_i = |\, H_1\,| \prod_{i=1}^{k} b_j'\, da_i ,$$

and

$$\textstyle\prod_{j=1}^{n-k} \prod_{i=1}^{k} b_j'\, d\tilde{a}_i = |\, H_1\,|^{n-k} \prod_{j=1}^{n-k} \prod_{i=1}^{k} b_j'\, da_i = \prod_{j=1}^{n-k} \prod_{i=1}^{k} b_j'\, da_i .$$

In a similar way we can carry out the transformation (4.34) and we have

$$\textstyle\prod_{j=1}^{n-k} \prod_{i=1}^{k} \tilde{b}_j'\, d\tilde{a}_i = \prod_{j=1}^{n-k} \prod_{i=1}^{k} b_j'\, da_i . \tag{4.35}$$

Thus the differential form (4.32) does not depend on the choice of a_i and b_j. Hence the local differential forms (4.32) agree wherever their domains of definition overlap and they can be considered as local expressions for a single global differential form v_k^n.

(3) Proof of invariance of v_k^n. Let H be a fixed orthogonal transformation:

$$\mathfrak{p} \to \mathfrak{q} = H\mathfrak{p}; \qquad \mathfrak{p} = H^{-1}\mathfrak{q} \qquad \mathfrak{p}, \mathfrak{q}\ \varepsilon\ G_{k,r} . \tag{4.36}$$

Suppose $a_i(\mathfrak{p})$ and $b_j(\mathfrak{p})$ are the respective sets of orthonormal vectors which span $\mathfrak{p}$ and its orthogonal complement, the elements of $a_i(\mathfrak{p})$ and $b_j(\mathfrak{p})$ being functions of $\mathfrak{p}$. Then

$$\begin{aligned} a_i(\mathfrak{p}) \to \hat{a}_i(\mathfrak{q}) &= a_i(H^{-1}\mathfrak{q}), \qquad b_j(\mathfrak{p}) \to \hat{b}_j(\mathfrak{q}) = b_j(H^{-1}\mathfrak{q}), \\ da_i(\mathfrak{p}) \to d\hat{a}_i(\mathfrak{q}) &= da_i(H^{-1}\mathfrak{q}), \\ b_j(\mathfrak{p})'\, da_i(\mathfrak{p}) &\to b_j(H^{-1}\mathfrak{q})'\, da_i(H^{-1}\mathfrak{q}) \\ &= b_j(H^{-1}\mathfrak{q})'H'H\, da_i(H^{-1}\mathfrak{q}) = (Hb_j(H^{-1}\mathfrak{q}))'\, d(Ha_i(H^{-1}\mathfrak{q}). \end{aligned} \tag{4.37}$$

[4] $\mathfrak{O}_{\mathfrak{p}_0}$ is almost all of $G_{k,r}$.

Hence

(4.38) $\prod_{j=1}^{n-k} \prod_{i=1}^{k} b_j(\mathfrak{p})' \, da_i(\mathfrak{p}) \to \prod_{j=1}^{n-k} \prod_{i=1}^{k} (Hb_j(H^{-1}\mathfrak{q}))' \, d(Ha_i(H^{-1}\mathfrak{q}))$.

Since $a_i(H^{-1}\mathfrak{q}) = a_i(\mathfrak{p})$ and $b_j(H^{-1}\mathfrak{q}) = b_j(\mathfrak{p})$ span $\mathfrak{p}$ and its orthogonal complement respectively, it follows that $Ha_i(H^{-1}\mathfrak{q})$ and $Hb_j(H^{-1}\mathfrak{q})$ span $\mathfrak{q} = H\mathfrak{p}$ and its orthogonal complement respectively. Therefore the right-hand side of (4.38) is equal to $v_k^n(\mathfrak{q})$. Thus the differential form $v_k^n(\mathfrak{p})$ is transformed by H to a differential form which at $\mathfrak{q}$ is equal to $v_k^n(\mathfrak{q})$. Hence v_k^n is invariant. Q.E.D.

4.7. *The invariant measure on the Stiefel manifold.* $V_{k,n}$ is the space of orthonormal k-frames in Euclidean n-space, of which we denote the typical member by an $n \times k$ matrix, A, satisfying the equation $A'A = I_k$. As in the case of the orthogonal group, $A' \, dA$ is a skew-symmetric matrix. Choose an $n \times (n - k)$ matrix, B, whose columns are orthonormal vectors spanning the orthogonal complement of the plane spanned by the columns of A. As in the case of the Grassmann manifold, the elements of B must be analytic functions of admissible coordinates for A. The invariant measure on the Stiefel manifold is given by the differential form

$$\prod_{j=1}^{n-k} \prod_{i=1}^{k} b_j' \, da_i \prod_{i<j}^{k} a_j' \, da_i, \tag{4.39}$$

which is defined almost everywhere on $V_{k,n}$.

Expressions like (4.39) can be constructed in a set of domains which cover the *entire* manifold. They define a differential form ω_k^n which is of maximum degree, namely $\frac{1}{2}k(2n - k - 1)$, and therefore represents a measure. (4.39) does not depend on the choice of B, the proof being similar to the one for $G_{k,r}$, Section 4.6 (2). It is invariant under the transformation $A \to HA$ where H is an $n \times n$ orthogonal matrix. The proof is a combination of the proofs for $O(n)$ and $G_{k,r}$. The b_j transform like the b_j for the Grassmann manifold as in (4.37), while a_i and da_i transform like the corresponding quantities for the orthogonal group as in (4.26) and (4.27).

Finally (4.39) is invariant under the group of transformations $A \to AH$ where H is now a $k \times k$ orthogonal matrix. The proof is practically identical with that of Theorem 4.3.

5. Integrals of the invariant measures.

5.1. *Integration of the invariant measure over the Stiefel manifold.* We first consider the case $k = 1$ which is that of the unit sphere in Euclidean n-space. The column vector a is of unit length and can be regarded as a point on the unit sphere. $b_1, b_2, \cdots, b_{n-1}$ are orthonormal column vectors orthogonal to a. We have to integrate the exterior differential form

$$\prod_{j=1}^{n-1} b_j' \, da = \omega_1^n(a) \tag{5.1}$$

over the unit sphere. As we shall see, the differential form is really the element of area on the unit sphere. This can be shown by a direct transformation to spherical polar coordinates, as follows.

Let x_n be the unit vector lying along the last coordinate axis and let $\mathfrak{q}$ be the

$[n-1]$; plane perpendicular to it, which thus contains the first $n-1$ coordinate axes. Let θ_1 be the angle between a and x_n and α be the unit vector lying along the orthogonal projection of a on $\mathfrak{q}$. θ_1 and α are new "coordinates" for a, and

$$a = x_n \cos \theta_1 + \alpha \sin \theta_1 . \tag{5.2}$$

If we exclude the points x_n and $-x_n$ from the sphere, then θ_1 has the range $0 < \theta_1 < \pi$, and α ranges over the unit sphere in the Euclidean $n-1$ space, $\mathfrak{q}$. In fact, the unit sphere with x_n and $-x_n$ removed, is the topological product of the ranges of θ_1 and α.

To express the differential form, (5.1), in terms of θ_1 and α, choose b_1 in the 2-plane spanned by a and x_n and such that b_1 is perpendicular to a; thus put

$$b_1 = -x_n \sin \theta_1 + \alpha \cos \theta_1 . \tag{5.3}$$

Choose $b_2, \cdots, b_{n-1}$ in $\mathfrak{q}$, perpendicular to α. Differentiating (5.2) we have

$$da = (-x_n \sin \theta_1 + \alpha \cos \theta_1)\, d\theta_1 + d\alpha \sin \theta_1 . \tag{5.4}$$

Since α is a unit vector and is perpendicular to x_n, $\alpha'\alpha = 1$, $\alpha' x_n = 0$. Differentiating,

$$\alpha' d\alpha = 0, \qquad d\alpha' x_n = 0. \tag{5.5}$$

Therefore, from (5.3), (5.4) and (5.5)

$$b_1' \, da = d\theta_1 . \tag{5.6}$$

Since $b_2, \cdots, b_{n-1}$ are orthogonal to x_n and α,

$$b_j' \, da = b_j' \, d\alpha \sin \theta_1 \qquad j = 2, \cdots, n-1.$$

Hence, if we repeat the procedure,

$$\begin{aligned} \textstyle\prod_{j=1}^{n-1} b_j' \, da &= \sin^{n-2}\theta_1 \, d\theta_1 \textstyle\prod_{j=2}^{n-1} b_j' \, d\alpha \\ &= \sin^{n-2}\theta_1 \sin^{n-3}\theta_2 \cdots \sin \theta_{n-2} \, d\theta_1 \, d\theta_2 \cdots d\theta_{n-1} \\ & \qquad 0 < \theta_{n-1} < 2\pi,\ 0 < \theta_i < \pi,\ i = 1, \cdots, n-2. \end{aligned} \tag{5.7}$$

Hence the differential form (5.1) is simply the element of area on the unit sphere. Integrating (5.7), we have

$$\int \prod b_j' \, da = A(n) \tag{5.8}$$

where $A(n)$ is the integral of (5.7) which is the area of the unit sphere in R^n:

$$A(n) = \frac{2\pi^{\frac{1}{2}n}}{\Gamma(\frac{1}{2}n)} . \tag{5.9}$$

In the general case, the invariant measure ω_k^n on $V_{k,n}$ is represented almost everywhere by the differential form (4.39).

THEOREM 5.1.

$$\int_{V_{k,n}} \omega_k^n = \prod_{i=1}^{k} A(n - i + 1) \tag{5.10}$$

where $A(\nu)$ is the area of the unit sphere in R^ν given in (5.9).

PROOF. It is sufficient to prove that

$$\int_{V_{k,n}} \omega_k^n = A(n) \int_{V_{k-1,n-1}} \omega_{k-1}^{n-1}, \tag{5.11}$$

where ω_{k-1}^{n-1} is the invariant measure on the Stiefel manifold $V_{k-1,n-1}$ of $(k - 1)$-frames in R^{n-1}, because iteration of (5.11) gives (5.10).

Rewriting (4.39) in full, we have

$$\begin{aligned} \omega_k^n = a_2' da_1\, a_3'\, da_1 \cdots a_k' da_1\, b_1'\, da_1 \cdots b_{n-k}'\, da_1 \\ \cdot\, a_3'\, da_2 \cdots a_k' da_2\, b_1'\, da_2 \cdots b_{n-k}'\, da_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot\, b_1'\, da_k \cdots b_{n-k}'\, da_k\, . \end{aligned} \tag{5.12}$$

The differential form in the first row depends only upon a_1. In fact, by an argument similar to the proof of (4.35) one can show that the first row remains unaltered if the vectors $a_2, \cdots, a_k, b_1, \cdots, b_{n-k}$ in it are replaced by any set of orthonormal vectors orthogonal to a_1. Comparison with formula (5.1) shows that the first row of (5.12) is the element of area on the unit sphere $V_{1,n}$ given by the equation $a_1'a_1 = 1$ in R^n. Denote it by $\omega_1^n(a_1)$.

For a fixed a_1, $(a_2, \cdots, a_k)$ range over all $(k - 1)$ frames in the $(n - 1)$-plane perpendicular to a_1. Denote this set of $(k - 1)$-frames by $V_{k-1,n-1}(a_1)$. The integral (5.10) can then be written as a repeated integral

$$\int_{V_{k,n}} \omega_k^n = \int_{V_{1,n}} \omega_1^n(a_1) \int_{V_{k-1,n-1}(a_1)} \tilde{\omega} \tag{5.13}$$

where $\tilde{\omega}$ is the differential form consisting of the last $k - 1$ rows of (5.12). Although $\tilde{\omega}$ and the range $V_{k-1,n-1}(a_1)$ over which it has to be integrated both depend on a_1, nevertheless the integral

$$\varphi(a_1) = \int_{V_{k-1,n-1}(a_1)} \tilde{\omega} \tag{5.14}$$

does not depend on a_1, as we shall now prove.

Let H be any fixed $n \times n$ orthogonal matrix. Then

$$\begin{aligned} \varphi(a_1) &= \int_{V_{k-1,n-1}(a_1)} a_3'\, da_2 \cdots b_1'\, da_2 \cdots \\ &= \int_{V_{k-1,n-1}(a_1)} a_3' H'H\, da_2 \cdots b_1' H'H\, da_2 \cdots \end{aligned}$$

Put

$$\tilde{a}_1 = Ha_1$$
$$\tilde{a}_i = Ha_i \qquad i = 2, \cdots, k$$
$$\tilde{b}_j = Hb_j \qquad j = 1, \cdots, n - k.$$

As $(a_2, \cdots, a_k)$ ranges over $V_{k-1,n-1}(a_1)$, $(\tilde{a}_2, \cdots, \tilde{a}_k)$ ranges over

$$V_{k-1,n-1}(Ha_1) = V_{k-1,n-1}(\tilde{a}_1).$$

Hence

$$\varphi(a_1) = \int_{V_{k-1,n-1}(\tilde{a}_1)} \tilde{a}_3' \, d\tilde{a}_2 \cdots \tilde{b}_1' \, d\tilde{a}_2 \cdots = \varphi(\tilde{a}_1). \tag{5.15}$$

Any unit vector $\tilde{a}_1$ can be obtained from a_1 by a suitable choice of H. Therefore $\varphi(a_1)$ does not depend on a_1 and is simply a constant φ.

Choose H so that $\tilde{a}_1$ is a vector with its last coordinate unity and all others zero. Since $\tilde{a}_2, \cdots, \tilde{b}_1 \cdots$ are orthogonal to $\tilde{a}_1$, each will then have zero as its last coordinate. Hence φ can be seen to be the integral of the invariant measure on the space of $k - 1$ frames in R^{n-1}. Since φ is a constant, we can integrate over the unit sphere in the right-hand side of (5.13) giving (5.11), and the theorem follows.

5.2. *Integration of the invariant measure on the orthogonal group.* It is a special case of the integral for the Stiefel manifold. If A is an $n \times n$ orthogonal matrix then from Theorem 5.1 we have

$$\int \omega_k^k = \int \prod_{i<j}^{k} a_i' \, da_j = \prod_{i=1}^{k} A(\nu) = \prod_{i=1}^{k} \frac{2\pi^{\nu/2}}{\Gamma^{(\nu/2)}}. \tag{5.16}$$

(5.16) gives the integral over the whole (improper) orthogonal group, that is, including the orthogonal matrices with negative determinant. The formula, (5.16), is consistent if we take the area of the unit sphere in R^1, which consists of only two points, namely ± 1, to be 2.

5.3. *Integration of the invariant measure on the Grassmann manifold.* The differential form, (4.39), representing the invariant measure on the Stiefel manifold looks like a product of differential forms representing the invariant measures on the Grassmann manifold and the orthogonal group. It suggests that the integral of the invariant measure on the Stiefel manifold should be the product of the integrals of the invariant measures on the Grassmann manifold and the orthogonal group, and such is, indeed, the case. Having evaluated the integrals of the invariant measures on the Stiefel manifold and the orthogonal group, we can thus find the integral of the invariant measure on the Grassmann manifold.

If $A \varepsilon V_{k,n}$ is an $n \times k$ matrix with orthonormal column vectors, the column vectors of A span a k-plane in R^n which can be regarded as a point, $\mathfrak{p}$, in the Grassmann manifold $G_{k,r}(r = n - k)$. The k-frame is determined uniquely by the specification of the plane, $\mathfrak{p}$, and the orientation of the k-frame in $\mathfrak{p}$. To specify the orientation, introduce another "reference" k-frame, represented by the columns of an $n \times k$ matrix, H, in the plane $\mathfrak{p}$, the elements of H being analytic functions of $\mathfrak{p}$ for almost all $\mathfrak{p}$. Then

$$A = HC \tag{5.17}$$

where C is a $k \times k$ orthogonal matrix. $\mathfrak{p}$ and C are functions of A and the transformation $A \leftrightarrow \mathfrak{p}$, C is one to one where A ranges over almost all the Stiefel manifold, $\mathfrak{p}$ over almost all the Grassmann manifold and C over the orthogonal group of order k.

Differentiating (5.17)

$$dA = H\,dC + dHC.$$

It can be assumed that the $n \times (n - k)$ matrix, B, introduced in Section 4.7 to construct the invariant measure on the Stiefel manifold, is a function of $\mathfrak{p}$ alone. Since $B'H = 0$,

$$B'\,dA = B'\,dHC \tag{5.18}$$

$$A'\,dA = C'\,dC + C'H'\,dHC. \tag{5.19}$$

Therefore

$$\prod\prod b_j'\,da_i = |\,C\,|^{n-k}\prod\prod b_j'\,dh_i = \prod\prod b_j'\,dh_i \tag{5.20}$$

$$\textstyle\prod_{i<j}^{k} a_j'\,da_i = \prod_{i<j}^{k} c_j'\,dc_i + *dH \tag{5.21}$$

where $*dH$ signifies differential forms involving the elements of dH. The right-hand side of (5.20) is a differential form defined on the Grassmann manifold and is of maximum degree, while H is a function defined on the same space. Therefore, the product of any differential of dH with (5.20) is zero.

Hence

$$\prod\prod b_j'\,da_i\prod a_j'\,da_i = \prod\prod b_j'\,dh_i\prod c_j'\,dc_i \tag{5.22}$$

and

$$K = \int_{G_{k,r}} v_k^n = \int\prod\prod b_j'\,dh_i = \int\omega_k^n \Big/ \int\omega_k^k = \sum_{\nu=n-k+1}^{n} A(\nu) \Big/ \sum_{\nu=1}^{k} A(\nu), \tag{5.23}$$

where $A(\nu)$ is given by (5.9).

6. Measures invariant under an induced group of transformations.

6.1. *Definitions.* If f is a map of a space $\mathfrak{X}$ on a space $\mathfrak{Y}$ then $f^{-1}y$ (or $f^{-1}\mathfrak{T}$) for $y \in \mathfrak{Y}$ (or $\mathfrak{T} \subset \mathfrak{Y}$) denotes the *inverse image* of y (or $\mathfrak{T}$) that is, the set of all points of $\mathfrak{X}$ mapped by f into y (or $\mathfrak{T}$). We say that a measure μ on $\mathfrak{X}$ is mapped by f on a measure $\bar{\mu}$ on $\mathfrak{Y}$ if, for every (measurable) set $\mathfrak{T} \subset \mathfrak{Y}$, $\bar{\mu}(\mathfrak{T}) = \mu(f^{-1}\mathfrak{T})$.

A many-to-one map f of a space $\mathfrak{X}$ on a space $\mathfrak{G}$ divides $\mathfrak{X}$ into a system of equivalence classes, each equivalence class being the inverse image $f^{-1}p$ of a point $p \varepsilon \mathfrak{G}$. The set of equivalence classes is thus in one-to-one correspondence with the points of $\mathfrak{G}$. Now if $\mathfrak{H}$ is a group of transformations of $\mathfrak{X}$ on itself each of whose elements transforms each equivalence class onto an equivalence class, then $\mathfrak{H}$ may be said to *induce* a group of transformations of the space of equivalence classes. Since each equivalence class corresponds to a point in $\mathfrak{G}$, $\mathfrak{H}$ thus induces a group of transformations of $\mathfrak{G}$.

DEFINITION. Let $\mathfrak{H}$ be a group of one-to-one transformations of a space $\mathfrak{X}$ onto itself and f be a map of $\mathfrak{X}$ on another space $\mathfrak{G}$. If for each $p \varepsilon \mathfrak{G}$ and each $H \varepsilon \mathfrak{H}$ there exists a point $p_1 \varepsilon \mathfrak{G}$ such that

$$H(f^{-1}p) = f^{-1}p_1 \tag{6.1}$$

then we define the transformation H_I of $\mathfrak{G}$ by the equation $p_1 = H_I p$ and say that the transformation $\mathfrak{H}$ acting on $\mathfrak{X}$ *induces* the transformation H_I on $\mathfrak{G}$ and that the group $\mathfrak{H}$ induces a group $\mathfrak{H}_I$.

LEMMA 6.1. *Suppose that f maps a space $\mathfrak{X}$ on a space $\mathfrak{G}$ and that a group $\mathfrak{H}$ of transformations of $\mathfrak{X}$ onto itself yields an induced group $\mathfrak{H}_I$ of transformations of $\mathfrak{G}$ onto itself. Let μ be a measure on $\mathfrak{X}$ mapped by f to a measure $\bar{\mu}$ on $\mathfrak{G}$. Then if μ is invariant under $\mathfrak{H}$, $\bar{\mu}$ is invariant under $\mathfrak{H}_I$.*

PROOF. For a measurable subset $\mathfrak{T} \subset \mathfrak{G}$

$$\bar{\mu}\{H_I\mathfrak{T}\} = \mu\{f^{-1}H_I\mathfrak{T}\} = \mu\{Hf^{-1}\mathfrak{T}\} = \mu\{f^{-1}\mathfrak{T}\} = \bar{\mu}\{\mathfrak{T}\}.$$

As a very simple illustration of the lemma, let $\mathfrak{X}$ be the Euclidean 2-plane and $\mathfrak{H}$ the group of rotations of it. Let μ be a finite measure on $\mathfrak{X}$ invariant under $\mathfrak{H}$, that is, circularly symmetrical, for example $e^{-(x_1^2+x_2^2)}dx_1dx_2$. Introduce polar coordinates (r, θ) in the plane. Let $\mathfrak{G}$ be the unit circle with θ as its coordinate and let f map a point (r, θ) in the plane, on θ. f maps the measure μ on a measure $\bar{\mu}$ in $\mathfrak{G}$. Then the group $\mathfrak{H}$ induces a group of rotations of the unit circle, under which, by Lemma 6.1, the measure $\bar{\mu}$ must be invariant. Thus μ is the uniform measure on the circle. We give several applications of the lemma.

6.2. *Distribution of the plane spanned by a set of random vectors.* Let X be an $n \times k$ matrix whose rows are n independent observations from a normal k-variate distribution with means zero, that is, with the distribution (2.1). As pointed out in Section 2 the distribution is invariant under the orthogonal group of transformations (2.2). Consider the columns of X as k vectors in Euclidean n-space namely, $x_1, \cdots, x_k$, and let $\mathfrak{p} = f(x)$ be the plane spanned by them. As, with probability one, $x_1, \cdots, x_k$ will be linearly independent, the plane will be k dimensional. Thus f is a map from the space of $n \times k$ matrices, $\mathfrak{X}$, to the Grassmann manifold $G_{k,r}$ $(r = n - k)$. The orthogonal group of transformations of $\mathfrak{X}$ induces a group of transformations of $G_{k,r}$. Hence, by Lemma 6.1, the distribution of $\mathfrak{p}$ is invariant under the induced group of transformations. According to Sections 4.3 and 4.6 the invariance characterizes the distribution of $\mathfrak{p}$ uniquely as the invariant measure on the Grassmann manifold, and the probability density is given by the differential form

$$K^{-1} \prod_{j=1}^{n-k} \prod_{i=1}^{k} b_j' da_i . \tag{6.2}$$

The invariance of the distribution of $\mathfrak{p}$ was recognized by Hotelling [8]. The type of argument, given above, to prove the invariance has been used by T. W. Anderson in other connections.

6.3. *Relation to the invariant measure on the orthogonal group.* As a second application of the lemma we show how the invariant measures on the Grassmann and Stiefel manifolds may be derived from the invariant measure on the orthogonal group. It was shown in Sections 3.3 and 3.4 that the Grassmann and Stiefel manifolds may be regarded as coset spaces of the orthogonal group. Let $\mathfrak{p}_0$ be a fixed k-plane in R^n, (thus $\mathfrak{p}_0 \varepsilon G_{k,r}$) and $A \varepsilon \mathfrak{A}$ an invariantly distributed orthogonal matrix. The matrix A transforms R^n into itself and induces a transformation of $G_{k,r}$ into itself. The mapping

$$A \rightarrow A\mathfrak{p}_0 = \mathfrak{p} \tag{6.3}$$

from $\mathfrak{A}$ to $G_{k,r}$ maps the invariant measure on $\mathfrak{A}$ to a measure on $G_{k,r}$ which, by Lemma 6.1 must be invariant.

The representation of a homogeneous space in terms of the group is very useful because the group has more symmetry, namely, a group element can be transformed from both left and right by other group elements and also the inverse can be taken. The representation of the invariant measure on the Grassmann manifold in terms of the invariant measure on the orthogonal group will be used in deriving the distribution of the canonical correlation coefficients.

The invariant distribution on the Grassmann manifold was obtained above by a random transformation of a fixed plane $\mathfrak{p}_0$ by an invariantly distributed orthogonal matrix. The result still holds if $\mathfrak{p}_0$ has an arbitrary probability distribution provided it is independent of A.

THEOREM 6.1. *If $\mathfrak{p}_0$ is a random point in the Grassmann manifold with an arbitrary probability distribution and A is an independently invariantly distributed orthogonal matrix and if*

$$\mathfrak{p} = A\mathfrak{p}_0 \tag{6.4}$$

then $\mathfrak{p}$ is invariantly distributed in the Grassmann manifold.

PROOF. Suppose $\mathfrak{p}_0 \varepsilon G_0$, $\mathfrak{p} \varepsilon G_{k,r}$ and $A \varepsilon \mathfrak{A}$. (6.4) is a map of $\mathfrak{A} \times G_0$ onto $G_{k,r}$. The joint distribution of the pair $(A, \mathfrak{p}_0)$ in $\mathfrak{A} \times G_0$ is invariant under the group of transformations

$$(A, \mathfrak{p}_0) \rightarrow (HA, \mathfrak{p}_0) \tag{6.5}$$

where H is an orthogonal matrix or transformation. But the transformation (6.5) induces the transformation $\mathfrak{p} \rightarrow H\mathfrak{p}$ in $G_{k,r}$. Hence by Lemma 6.1, $\mathfrak{p}$ is invariantly distributed. Q.E.D.

In a similar way one can show that the invariant measure on the Stiefel manifold can be obtained by a random orthogonal transformation of a fixed k-frame, or even of a random k-frame provided it is distributed independently of the orthogonal matrix.

6.4. *Critical angles between two planes.*

THEOREM 6.2. *If $\mathfrak{p}$ and $\mathfrak{q}$ are planes of dimension p and q respectively, $p \leqq q$, in Euclidean space R^n and if θ is the angle between an arbitrary vector a in $\mathfrak{p}$ and an arbitrary vector α in $\mathfrak{q}$, then as a and α vary over $\mathfrak{p}$ and $\mathfrak{q}$ respectively, θ has p stationary values, $\frac{1}{2}\pi \geqq \theta_1 \geqq \theta_2 \geqq \cdots \geqq \theta_p \geqq 0$ corresponding to pairs of vectors say $a_1, \alpha_1, \cdots, a_p, \alpha_p$. The stationary angles θ_i are uniquely determined by $\mathfrak{p}$ and $\mathfrak{q}$ and if no two of them are equal, the corresponding vectors are uniquely determined apart from length and a simultaneous reversal of direction of a_i* and *α_i. a_i is orthogonal to a_j and $\alpha_j (i \neq j)$.*

The angle between a_i and α_i is, of course, θ_i. These angles are called the *critical angles* between the planes $\mathfrak{p}$ and $\mathfrak{q}$. For a proof of the theorem see Hotelling [8] or Roy [17].

7. Application to the distribution of the canonical correlation coefficients and the roots of certain determinantal equations. If the rows of the matrix $[X \vdots Y] = [x_1 \cdots x_k y_1 \cdots y_q]$ are n independent samples from a $(k + q)$-variate distribution, with means all zero, and if $\mathfrak{p}$ and $\mathfrak{q}$ are the planes spanned by the column vectors $x_1, \cdots, x_k$ and $y_1, \cdots, y_q$ respectively, then Hotelling [8] (see also Roy [17]) showed that the sample canonical correlations between X and Y are the cosines of the critical angles $\theta_1, \cdots, \theta_l$ between $\mathfrak{p}$ and $\mathfrak{q}$ where $l = \min(k, q)$. Denote $(\theta_1, \cdots, \theta_l)$ by $\angle(\mathfrak{p}, \mathfrak{q})$.

The canonical correlation coefficients are often expressed as the roots of a determinantal equation. Let X_1 be the $n \times k$ matrix whose ith column is the orthogonal projection of the ith column of X on the plane spanned by the column vectors of Y. Then the roots of the determinantal equation $| X_1'X_1 - \lambda X'X | = 0$ are the squares of the canonical correlation coefficients, $\cos \theta_i$. The same problem also arises from multivariate analysis of variance (at least, in the null case). For this problem Y is fixed instead of random, for example, Y can be taken as the matrix whose column vectors represent the first q coordinate axes.

The distribution of the canonical correlations for samples from normal populations in the null case, was found simultaneously by Fisher [5], Hsu [9], Roy [16] and Mood [13], in 1939. Let us illustrate the application of Grassmann and Stiefel manifolds by giving yet another derivation!

For the null case, that is, when X and Y are independent, Fisher has pointed out that the assumption that Y is normally distributed can be dropped. Thus assume that the rows of X are independent samples from a k-variate normal distribution (with means zero) and that Y has any arbitrary distribution independent of that of X. In Section 6.2 it was proved that the plane $\mathfrak{p}$ is invariantly distributed. Since Y is independent of X, $\mathfrak{q}$ is independent of $\mathfrak{p}$. From the joint distribution of $\mathfrak{p}$ and $\mathfrak{q}$ we derive the distribution of the critical angles, $\angle(\mathfrak{p}, \mathfrak{q})$.

The distribution of $\angle(\mathfrak{p}, \mathfrak{q})$ remains the same if the random plane $\mathfrak{q}$ is replaced by a fixed plane. We prove this by representing the distribution of $\mathfrak{p}$ in terms of the distribution of an invariantly distributed orthogonal matrix as shown in Section 6.3. Let $A \varepsilon \mathfrak{A}$ be a random orthogonal matrix distributed invariantly and independently of $\mathfrak{q}$. Let $\mathfrak{p}_0$ be an arbitrary fixed k-plane. Then $A\mathfrak{p}_0$ is a random k-plane distributed invariantly and independently of $\mathfrak{q}$ and thus the joint distribution of $A\mathfrak{p}_0$ and $\mathfrak{q}$ is the same as the joint distribution of $\mathfrak{p}$ and $\mathfrak{q}$. Hence $\angle(Ap_0, \mathfrak{q})$ has the same distribution as $\angle(\mathfrak{p}, \mathfrak{q})$.

Since the critical angles between two planes are invariant under their simultaneous orthogonal transformation, we have, on multiplying $A\mathfrak{p}_0$ and $\mathfrak{q}$ by A^{-1}

$$\angle(A\mathfrak{p}_0, \mathfrak{q}) = \angle(\mathfrak{p}_0, A^{-1}\mathfrak{q}). \tag{7.1}$$

But by Lemma 4.5, A^{-1} has the same distribution as A; therefore $\angle(A\mathfrak{p}_0, \mathfrak{q})$ has the same distribution as $\angle(\mathfrak{p}_0, A\mathfrak{q})$. Since, by Theorem 6.1, $A\mathfrak{q}$ is invariantly distributed, it follows that $A\mathfrak{q}$ has the same distribution as $A\mathfrak{q}_0$, $\mathfrak{q}_0$ being an arbitrary but fixed q-plane

We have now proved that the distribution of $\angle(\mathfrak{p}, \mathfrak{q})$ is the same as the distribution of $\angle(\mathfrak{p}_0, A\mathfrak{q}_0)$ which, again, is the same as that of $\angle(A\mathfrak{p}_0, \mathfrak{q}_0)$. We can use whichever is the more convenient.

Case 1. $n \geqq k + q$. Let us choose the plane of the smaller number of dimensions as the random plane. Suppose it is $A\mathfrak{p}_0$ (hence $k \leqq q$) which we shall now denote by $\mathfrak{p}$. Were $k > q$, we could simply use $\angle(\mathfrak{p}_0, A\mathfrak{q}_0)$ instead of $\angle(A\mathfrak{p}_0, \mathfrak{q}_0)$. Choose $\mathfrak{q}_0$ as the plane spanned by the first q coordinate axes.

If we take arbitrary orthonormal vectors $a_1, \cdots, a_k$ in $\mathfrak{p}$ and $n - k$ arbitrary orthonormal vectors $b_1, \cdots, b_{n-k}$ in the orthogonal complement of $\mathfrak{p}$, then according to Sections 4.6 and 6.2 the distribution of $\mathfrak{p} = A\mathfrak{p}_0$ is given by the differential form (6.2).

We shall see that apart from a set of measure zero, the Grassmann manifold is analytically homeomorphic to the topological product of a simplex in R^p, over which the critical angles range, and two Stiefel manifolds. By transforming the differential form to a differential form on the product space and integrating over the two Stiefel manifolds, we find the distribution of the critical angles.

According to Theorem 6.2 the vectors in $\mathfrak{p}$ which make the critical angles with $\mathfrak{q}_0$ are uniquely determined by $\mathfrak{p}$ apart from length and reversal of direction, provided no two critical angles are equal and no critical angle is 0 or $\frac{1}{2}\pi$. Let us exclude these exceptional cases as they have measure zero.

The orthonormal column vectors $a_1, \cdots, a_k$ in (6.2) can be chosen arbitrarily in $\mathfrak{p}$. Let them be the vectors which make the respective critical angles with $\mathfrak{q}_0$ and such that the first component of each a_i is positive. Such conditions determine $(a_1, \cdots, a_k)$ uniquely as analytic functions of $\mathfrak{p}$ for almost all $\mathfrak{p}$, that is, on a set $\tilde{G}$ of $\mathfrak{p}$ excluding merely a set of measure zero. Let $\alpha_1, \cdots, \alpha_k$ be unit vectors in $\mathfrak{q}_0$ which make the respective critical angles with $\mathfrak{p}$. Then $\alpha_1, \cdots, \alpha_k$ are mutually orthogonal and lie along the respective projections of $a_1, \cdots, a_k$ on $\mathfrak{q}_0$. Let $\beta_1, \cdots, \beta_k$ be the orthonormal vectors lying along the respective projections of $a_1, \cdots, a_k$ on the orthogonal complement of $\mathfrak{q}_0$. Thus

$$\begin{aligned} &\alpha_i'\alpha_j = \delta_{ij}, \quad \beta_i'\beta_j = \delta_{ij}, \quad \alpha_i'\beta_j = 0, \quad && i, j = 1, \cdots, k \\ &a_i = \alpha_i \cos\theta_i + \beta_i \sin\theta_i && i = 1, \cdots, k. \end{aligned} \tag{7.2}$$

Since $(\alpha_1, \cdots, \alpha_k)$ is a k-frame in the q-space $\mathfrak{q}_0$ it can be regarded as a point in a Stiefel manifold. Let $\tilde{V}_{k,q}$ denote the part of this Stiefel manifold over which $(\alpha_1, \cdots, \alpha_k)$ ranges. From the condition imposed above upon $a_1, \cdots, a_k$, it follows that this will be the set of all orthonormal $(\alpha_1, \cdots, \alpha_k)$ such that the first component of each vector is positive. On the other hand, $(\beta_1, \cdots, \beta_k)$ will range over the whole Stiefel manifold $V_{k,n-q}$. Let Θ be the set of $(\theta_1, \cdots, \theta_k)$ such that $\frac{1}{2}\pi > \theta_1 > \theta_2 > \cdots > \theta_k > 0$.

Thus $\mathfrak{p} \,\varepsilon\, \tilde{G}$ determines $(\theta_1, \cdots, \theta_k) \,\varepsilon\, \Theta$, $(\alpha_1, \cdots, \alpha_k) \,\varepsilon\, \tilde{V}_{k,q}$ and $(\beta_1, \cdots, \beta_k) \,\varepsilon\, V_{k,n-q}$ uniquely, and conversely $\mathfrak{p}$ is determined by these, because by (7.2) they determine a set of vectors $a_1, \cdots, a_k$ which span $\mathfrak{p}$. The transformations are not only one-to-one but also analytic; hence $\tilde{G}$ is analytically homeomorphic to the topological product of Θ, $\tilde{V}_{k,q}$ and $V_{k,n-q}$. By a suitable choice of $b_1, \cdots, b_{n-k}$ we express the differential form (6.2) on $\tilde{G}$ as a differential form on the product space.

Differentiating (7.2), we have the relations

$$\alpha_i' \, d\alpha_i = 0, \qquad \beta_i' \, d\beta_i = 0 \qquad i = 1, \cdots, k$$

$$(7.3) \quad \alpha_i' \, d\alpha_j = -\alpha_j' \, d\alpha_i, \qquad \beta_i' \, d\beta_j = -\beta_j' \, d\beta_i \qquad i \neq j \quad i, j = 1, \cdots, k$$

$$da_i = (-\alpha_i \sin \theta_i + \beta_i \cos \theta_i) \, d\theta_i + d\alpha_i \cos \theta_i + d\beta_i \sin \theta_i \qquad i = 1, \cdots, k.$$

Since α_i and β_j lie in fixed mutually orthogonal planes

$$(7.4) \qquad \alpha_i' \, d\beta_j = \beta_j' \, d\alpha_i = 0 \qquad i, j = 1, \cdots, k.$$

Now $(b_1, \cdots, b_{n-k})$ is an arbitrary set of orthonormal vectors in the orthogonal complement of $\mathfrak{p}$. By choosing

$$(7.5) \qquad b_i = -\alpha_i \sin \theta_i + \beta_i \cos \theta_i \qquad i = 1, \cdots, k.$$

We have

$$(7.6) \qquad b_i' \, da_i = d\theta_i \qquad i = 1, \cdots, k$$

and

$$(7.7) \qquad b_i' \, da_j = -\alpha_i' \, d\alpha_j \sin \theta_i \cos \theta_j + \beta_i' \, d\beta_j \cos \theta_i \sin \theta_j \qquad i \neq j \quad i, j = 1, \cdots, k.$$

By using the relations (7.3) and remembering to change the sign when reversing the order of two linear differential forms, and remembering that any term containing a repeated linear form is zero, for example, $(\alpha_i' \, d\alpha_j)(\alpha_j' \, d\alpha_i) = 0$, we calculate that

$$(7.8) \qquad (b_i' da_j)(b_j' da_i) = (\alpha_j' d\alpha_i)(\beta_j' d\beta_i) \, (\cos^2 \theta_j - \cos^2 \theta_i) \qquad i \neq j \quad i, j = 1, \cdots, k.$$

Choose the orthonormal vectors $b_{k+1}, \cdots, b_q$ in $\mathfrak{q}_0$ perpendicular to $\alpha_1, \cdots, \alpha_k$ and choose $b_{q+1}, \cdots, b_{n-k}$ in the orthogonal complement of $\mathfrak{q}_0$ perpendicular to $\beta_1, \cdots, \beta_k$. Then $b_1, \cdots, b_{n-k}$ are orthonormal and span the orthogonal complement of $\mathfrak{p}$. Furthermore

$$(7.9)\qquad b_j' \, da_i = \begin{cases} b_j' \, d\alpha_i \cos\theta_i & i = 1, \cdots, k \quad j = k+1, \cdots, q \\ b_j' \, d\beta_i \cos\theta_i & i = 1, \cdots, k \quad j = q+1, \cdots, n-k. \end{cases}$$

Multiplying (7.6), (7.8) and (7.9) according to the rule of the exterior product, we see that (6.2) becomes

$$(7.10)\qquad \begin{aligned} K^{-1} \prod_{i=1}^{k} \prod_{j=1}^{n-k} b_j' \, da_i = K^{-1} \prod_{i<j}^{k} \alpha_j' \, d\alpha_i \prod_{i=1}^{k} \prod_{j=k+1}^{q} b_j' \, d\alpha_i \prod_{i<j}^{k} \beta_j' \, d\beta_i \prod_{i=1}^{k} \prod_{j=q+1}^{n-k} \\ \cdot b_j' \, d\beta_i \left(\prod_{i=1}^{k} \cos\theta_i \right)^{q-k} \left(\prod_{i=1}^{k} \sin\theta_i \right)^{n-q-k} \prod_{i<j}^{k} (\cos^2\theta_j - \cos^2\theta_i) \, d\theta_1 \cdots d\theta_k . \end{aligned}$$

Thus $\alpha_1, \cdots, \alpha_k$ and $\beta_1, \cdots, \beta_k$ are invariantly distributed k-frames in $\mathfrak{q}_0$ and the orthogonal complement of $\mathfrak{q}_0$ respectively. Hence the invariant distribution on the Grassmann manifold can be transformed into three independent distributions, namely, the distribution of the critical angles that the plane makes with a fixed plane, and two invariant distributions in Stiefel manifolds. Since the restriction on the a_i implies that the first component of each α_i is positive, the $\alpha_1, \cdots, \alpha_k$ range over the (2^{-k})th part of the Stiefel manifold while $\beta_1, \cdots, \beta_k$ ranges over the whole Stiefel manifold.

Therefore by Theorem 5.1

$$(7.11)\qquad \int \prod_{i<j}^{k} \alpha_j' \, d\alpha_i \prod_{i=1}^{k} \prod_{j=k+1}^{q} b_j' \, d\alpha_i = 2^{-k} \prod_{i=1}^{k} A(q - i + 1)$$

and

$$(7.12)\qquad \int \prod_{i<j}^{k} \beta_j' \, d\beta_i \prod_{i=1}^{k} \prod_{j=q+1}^{n-k} b_j' \, d\beta_i = \prod_{i=1}^{k} A(n - q - i + 1).$$

Hence the distribution of the critical angles is

$$(7.13)\qquad K(n, k, q) \left(\prod_{i=1}^{k} \cos\theta_i \right)^{q-k} \left(\prod_{i=1}^{k} \sin\theta_i \right)^{n-q-k} \prod_{i<j}^{k} (\cos^2\theta_j - \cos^2\theta_i) \, d\theta_1 \cdots d\theta_k$$

where

$$(7.14)\qquad K(n, k, q) = \prod_{i=1}^{k} \frac{A(k - i + 1) A(q - i + 1) A(n - q - i + 1)}{2A(n - i + 1)}$$

and

$$A(n) = \frac{2\pi^{n/2}}{\Gamma\left(\dfrac{n}{2}\right)}.$$

The distribution of the canonical correlations is found by putting $r_i = \cos\theta_i$.

CASE 2. $q < n < k + q$. The planes $\mathfrak{p}$ and $\mathfrak{q}$ must intersect in a $k + q - n$ dimensional space; therefore $k + q - n$ of the critical angles will be identically zero, leaving only $\theta_1, \cdots, \theta_{n-q}$ different from zero.

CASE 2a. $k \leqq q$.

Step I. According to (7.1) we can take a fixed plane $\mathfrak{p}_0$ and a random plane $\mathfrak{q}$ instead of $\mathfrak{p}$ and $\mathfrak{q}_0$. If $a_1, \cdots, a_q$ are chosen in $\mathfrak{q}$ and $b_1, \cdots, b_{n-q}$ in the orthogonal complement of $\mathfrak{q}$, then the distribution of $\mathfrak{q}$ is given by

$$K^{-1} \prod_{i=1}^{q} \prod_{j=1}^{n-q} b_j' \, da_i .$$

Step II. But this differential form equally represents the distribution of the orthogonal complement, $\mathfrak{q}^*$, of $\mathfrak{q}$, which has dimensions $n - q < k$. Indeed

$$K^{-1} \prod_{i=1}^{q} \prod_{j=1}^{n-q} b_j' \, da_i = K^{-1} \prod_{i=1}^{n-q} \prod_{j=1}^{q} a_j' \, db_i . \tag{7.15}$$

The critical angles $\theta_1^*, \cdots, \theta_{n-q}^*$ between $\mathfrak{q}^*$ and $\mathfrak{p}_0$ are the complements of the nonzero critical angles between $\mathfrak{q}$ and $\mathfrak{p}_0$, that is

$$\theta_i^* = \frac{\pi}{2} - \theta_i \qquad i = 1, \cdots, n - q. \tag{7.16}$$

We can carry out the preceding analysis, interchanging the roles of a_i and b_j. Using the correspondences

Old	*New*	
n	n	dimension of space
k	$n - q$	dimension of random plane
q	k	dimension of fixed plane

from (7.13) we obtain the distribution of the θ_i^*:

$$K(n, n - q, k) \left(\prod_{i=1}^{n-q} \cos \theta_i^* \right)^{k+q-n} \left(\prod_{i=1}^{n-q} \sin \theta_i^* \right)^{q-k} \cdot \prod_{i<j}^{n-q} (\cos^2 \theta_i^* - \cos^2 \theta_j^*) \, d\theta_1^* \cdots d\theta_{n-q}^*$$

and putting $\theta_i^* = \frac{1}{2}\pi - \theta_i$, the distribution of θ_i:

$$K(n, n - q, k) \left(\prod_{i=1}^{n-q} \sin \theta_i \right)^{k+q-n} \left(\prod_{i=1}^{n-q} \cos \theta_i \right)^{q-k} \cdot \prod_{i<j}^{n-q} (\cos^2 \theta_j - \cos^2 \theta_i) \, d\theta_1 \cdots d\theta_{n-q}. \tag{7.17}$$

CASE 2b. $k \geqq q \qquad k < n < k + q$. We only require Step II. That is we take the orthogonal complement of $\mathfrak{p}$, instead of $\mathfrak{p}$, as the random plane. Hence the correspondences are

Old	*New*	
n	n	dimension of space
k	$n - k$	dimension of random plane
q	q	dimension of fixed plane.

The distribution of θ_i^* is

$$K(n, n-k, q)\left(\prod_{i=1}^{n-k}\cos\theta_i^*\right)^{k+q-n}\left(\prod_{i=1}^{n-k}\sin\theta_i^*\right)^{k-q} \cdot \prod_{i<j}^{n-k}(\cos^2\theta_j^* - \cos^2\theta_i^*)\, d\theta_1^* \cdots d\theta_{n-k}^*$$

and hence of the θ_i is

$$K(n, n-k, q)\left(\prod_{i=1}^{n-k}\sin\theta_i\right)^{k+q-n}\left(\prod_{j=1}^{n-k}\cos\theta_i\right)^{k-q} \cdot \prod_{i<j}^{n-k}(\cos^2\theta_j - \cos^2\theta_i)\, d\theta_1 \cdots d\theta_{n-k}. \tag{7.18}$$

As before, the distribution of the canonical correlation coefficients is obtained by putting $r_i = \cos\theta_i$.

8. Decomposition of the distribution of a normal multivariate sample.

8.1. *Introduction.* The distribution of n independent observations from a univariate normal population with zero mean and unit variance can be split into two independent distributions by transformation to spherical polar coordinates; namely, the χ^2 distribution and the invariant distribution of a vector. The latter can be expressed in terms of the element of area on the unit sphere in R^n. This result is useful in deriving the various sampling distributions.

By using the exterior differential forms (see Sections 4.5 to 5.2), for the invariant measures on the Grassmann and Stiefel manifolds, we shall derive the multivariate analogue of this decomposition. Let X be an $n \times k$ matrix ($k \leqq n$) whose rows are n independent observations from a normal k-variate population with means zero and variance covariance matrix Σ. X is distributed as in (2.1).

THEOREM 8.1. *The distribution* (2.1) *of a normal k-variate sample can be decomposed into three independent distributions*

a. *essentially the Wishart distribution,*
b. *the invariant distribution of the plane spanned by the vectors* $x_1, \cdots, x_k$,
c. *the invariant distribution of the orthogonal* $k \times k$ *matrix which determines the orientation of* $x_1, \cdots, x_k$ *in the plane they span.*

The process of decomposition yields, incidentally, the distribution of the latent roots of the sample variance covariance matrix, a distribution found by Fisher [5], in the special case that all the population latent roots are equal.

Let X be an $n \times k$ matrix distributed as in (2.1). We can put

$$X = A\,L\,G' \tag{8.1}$$

where

1. A is an $n \times k$ matrix such that

$$A'A = I_k, \tag{8.2}$$

2. L is a $k \times k$ diagonal matrix with $l_1 > l_2 > \cdots > l_k > 0$ down the diagonal, $l_1^2, \cdots, l_k^2$ being the latent roots of $X'X$,

3. G is a $k \times k$ orthogonal matrix with the elements in the first row positive.

Equation (8.1) holds for almost all X and determines A, L, and G uniquely. To obtain (8.1), let G be the matrix satisfying Condition 3, which reduces $X'X$ to diagonal form L^2, that is, such that $X'X = G L^2 G'$. Putting $A = X G L^{-1}$ yields (8.1) with A satisfying (8.2). (8.1) implies that the Euclidean space R^{nk} of matrices X is, apart from a set of measure zero, analytically homeomorphic to the topological product of a Stiefel manifold $V_{k,n}$, a simplex in R^k and part of an orthogonal group manifold, over which A, L and G range respectively. We now express the volume element $\prod dx_{ij}$ of R^{nk} as an exterior product of differential forms on these manifolds.

Differentiate (8.1)

$$dX = dA\, L\, G' + A\, dL\, G' + A\, L\, dG'. \tag{8.3}$$

Choose an $n \times n - k$ matrix B such that the partitioned matrix $[A \mid B]$ is orthogonal. Premultiply (8.3) by the transpose of $[A \mid B]$ and post-multiply by G;

$$\begin{aligned} \left[\frac{A'}{B'}\right] dXG &= \left[\frac{A'\, dA}{B'\, dA}\right] L + \left[\frac{dL}{0}\right] + \left[\frac{L\, dG'G}{0}\right] \\ &= \left[\frac{A'dA\, L + dL - L\, G'\, dG}{B'\, dA\, L}\right] \end{aligned} \tag{8.4}$$

since $G'\, dG$, like $A'\, dA$, is skew symmetric (cf. Section 4.5).

To evaluate the exterior product of the left-hand side of (8.4), consider first a single column dx_j of dX. By Lemma 4.1 the exterior product of the elements of the transformed vector $[A \mid B]'dx_j$ is $\mid A \mid B \mid \prod_{i=1}^{n} dx_{ij}$. Hence the exterior product of the elements of the matrix $[A \mid B]'dX$ is $\mid A \mid B \mid^k \prod_{i,j} dx_{ij}$. In the left-hand side of (8.4) the row vectors of the matrix $[A \mid B]'\, dX$ are transformed by G. Hence the exterior product of the elements in a single row will be multiplied by $\mid G \mid$ and since $[A \mid B]'dX$ has n rows, the exterior product of all the elements is multiplied by $\mid G \mid^n$. Therefore the exterior product of the elements in the left-hand side of (8.4) is

$$\mid A \mid B \mid^k \mid G \mid^n \prod_{i,j} dx_{ij} \tag{8.5}$$

which equals $\prod_{i,j} dx_{ij}$ since $[A \mid B]$ and G are orthogonal matrices.

The exterior product of the (ij)th and (ji)th elements of the matrix on the right-hand side of (8.4) is

$$(a_i'\, da_j l_j - l_i g_i'\, dg_j)(a_j'\, da_i l_i - l_j g_j'\, dg_i) = (a_i'\, da_j)(g_i'\, dg_j)(l_j^2 - l_i^2). \tag{8.6}$$

$$i \neq j \quad i, j = 1, \cdots, k$$

By considering that the row vectors of the matrix $B' \, dA$ are transformed by L we see that the alternating product of the elements of the matrix $B' \, dA \, L$ is

$$| L |^{n-k} \prod_{i=1}^{k} \prod_{j=1}^{n-k} b_j' \, da_i \tag{8.7}$$

where b_j is the jth column of B and a_i is the ith column of A. Therefore, from (8.5), (8.6) and (8.7)

$$\prod dx_{ij} = \left(\prod_{i=1}^{k} l_i\right)^{n-k} \prod_{i<j}^{k} (l_i^2 - l_j^2) \, dl_1 \cdots dl_k \prod_{i<j}^{k} g_i' \, dg_j \cdot \prod_{i<j}^{k} a_i' \, da_j \prod_{j=1}^{n-k} \prod_{i=1}^{k} b_j' \, da_i . \tag{8.8}$$

This is an interesting decomposition of the volume element of the nk-dimensional Euclidean sample space.[5] The differential form

$$\prod_{i<j}^{k} g_i' \, dg_j \tag{8.9}$$

is the invariant measure on the orthogonal group, which was discussed in Section 4.4, and the differential form

$$\prod_{i<j}^{k} a_j' \, da_i \prod_{i=1}^{k} \prod_{j=1}^{n-k} b_j' \, da_i \tag{8.10}$$

is the invariant measure on the Stiefel manifold $V_{k,n}$ of k-frames in R^n (see Section 4.6). The decomposition leads immediately to the distribution of the latent roots of the sample variance covariance matrix.

8.2. *Latent roots of the variance covariance matrix.* The distribution of the latent roots, $l_1^2, \cdots, l_k^2$, is found by integrating over the Stiefel manifold and over the group of orthogonal matrices. Since the density function in (2.1) does not depend upon A, the integral over the Stiefel manifold can be evaluated separately and is given by (5.10).

Let C be the $k \times k$ orthogonal matrix which reduces Σ to diagonal form

$$C'\Sigma C = \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} . \tag{8.11}$$

The columns of C are linear functions which give the extremal variances in the population. Since the columns of G are the linear functions which give the extremal variances in the sample, G is the maximum likelihood estimate of C. From (8.11) and (8.1)

$$\sigma^{ij} = \sum_{\nu=1}^{k} \frac{C_{i\nu} C_{j\nu}}{\lambda_\nu} , \qquad || \sigma^{ij} || = \Sigma^{-1} \tag{8.12}$$

$$x_i' x_j = \sum_{\mu} g_{i\mu} g_{j\mu} l_\mu^2 . \tag{8.13}$$

[5] This corresponds to the result given by Olkin [14] p. 29.

Hence from (2.1), (8.8), (5.10), (8.12) and (8.13) we get the joint distribution of the latent roots $l_1^2, \cdots, l_k^2$ of $X'X$ and the linear functions G;

$$dF(l_1^2, \cdots, l_k^2; G) = \frac{\prod_{i=1}^{k} A(n-i+1)}{(2\pi)^{nk/2} 2^k \prod_{i=1}^{k} \lambda_i^{n/2}} \prod_{i<j}^{k} (l_i^2 - l_j^2) \tag{8.14}$$
$$\cdot \left(\prod_{i=1}^{k} l_i^2\right)^{\frac{1}{2}(n-k-1)} dl_1^2 \cdots dl_k^2 \, e^{-\frac{1}{2}\sum_{i,j,\nu,\mu}^{k} c_{i\nu} g_{i\mu} c_{j\nu} g_{j\mu} l_\mu^2 \lambda_\nu^{-1}} \prod_{i<j}^{k} g_j' \, dg_i$$

where $A(\nu)$ is given in (5.9).

To obtain the distribution of the latent roots of $X'X$ alone, put $S = C'G$. Being the invariant measure,

$$\prod_{i<j}^{k} g_i' \, dg_j = \prod_{i<j}^{k} s_i' \, ds_j \, .$$

By dividing (8.14) by 2^k, we can drop the restriction that the elements in the first row of G must be positive and thus let G and hence S range over the whole (improper) orthogonal group. The distribution of the latent roots of $X'X$ is then

$$dF(l_1^2, \cdots, l_k^2) = \frac{\prod_{i=1}^{k} A(n-i+1)}{(2\pi)^{\frac{1}{2}nk} 2^{2k} \prod_{i=1}^{k} \lambda_i^{\frac{1}{2}n}} \left(\prod_{i=1}^{k} l_i^2\right)^{\frac{1}{2}(n-k-1)} \prod_{i<j}^{k} (l_i^2 - l_j^2)$$
$$\cdot \left(\int e^{-\frac{1}{2}\sum_{\nu,\mu}^{k} l_\mu^2 \lambda_\nu^{-1} s_{\nu\mu}^2} \prod_{i<j}^{k} s_i' \, ds_j\right) dl_1^2 \cdots dl_k^2$$

the integral being taken over the orthogonal group.

In the special cases

a. When the population latent roots, λ_i, are all equal, the exponential term in the integral becomes

$$e^{-(1/2\lambda)\sum_{j=1}^{k} l_j^2}$$

that is, independent of S, and we only have the invariant measure on the orthogonal group whose integral is given by (5.16). This distribution was found by Fisher.

b. When $n = 2$, we can put

$$S = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \qquad 0 < \theta < 2\pi$$

$$s_2' \, ds_1 = d\theta$$

and the integral is expressible as an imaginary Bessel function of zero order as given by Girshick.

8.3. *The Wishart distribution.* The variables $l_1, \cdots, l_k, G$, can be expressed in terms of $X'X$.

From (8.1) $X'X = G L^2 G'$. Differentiating and pre- and post-multiplying by G' and G respectively

$$(8.15) \qquad \begin{aligned} G' \, d(X'X) \, G &= G' \, dG \, L^2 + L^2 \, dG' \, G + 2L \, dL \\ &= G' \, dG \, L^2 - L^2 \, G' \, dG + 2L \, dL. \end{aligned}$$

The (i, j)th element of the matrix in the right-hand side of (8.15) is

$$(8.16) \qquad \begin{aligned} &g_i' \, dg_j \, (l_j^2 - l_i^2) \qquad && i \neq j \\ &2l_i \, dl_i && i = j. \end{aligned}$$

To evaluate the alternating product of the diagonal and super diagonal elements of the left-hand side of (8.15), notice that

$$d(X'X) \to G' \, d(X'X) \, G$$

is a linear transformation of $d(X'X)$, regarded as a vector in a space of dimension $\frac{1}{2}k(k + 1)$. The coefficient of $\prod_{i \leq j}^k d(x_i'x_j)$ will be the determinant of this linear transformation, which, by an argument similar to the proof of Lemma 4.4, is proved to be a power of $| G |$ which is 1. Hence the exterior product of the diagonal and super diagonal elements of the matrix on the left-hand side of (8.15) is

$$(8.17) \qquad \prod_{i \leq j}^k d(x_i'x_j).$$

(8.17) and (8.16) give

$$(8.18) \qquad \prod_{i \leq j}^k d(x_i' x_j) = 2^k \left(\prod_{i=1}^k l_i \right) \prod_{i<j}^k (l_i^2 - l_j^2) \prod_{i<j}^k g_i' \, dg_j \, dl_1 \cdots dl_k.$$

Using (8.18) to substitute for $\prod_{i<j}^k g_i' \, dg_j$ in (8.8) yields

$$(8.19) \qquad \prod_{i,j} dx_{ij} = 2^{-k} \, | X'X |^{\frac{1}{2}(n-k-1)} \prod_{i \leq j}^k d(x_i'x_j) \prod_{i<j}^k a_i' \, da_j \prod_{j=1}^{n-k} \prod_{i=1}^k b_j' \, da_i .$$

Adjoining of the density factor of (2.1) and integration over the Stiefel manifold (given by (5.10)) yield the distribution of $X'X$ which is essentially the Wishart distribution:

$$(8.20) \qquad \begin{aligned} dF(X'X) = \frac{| \Sigma |^{-\frac{1}{2}n}}{(2\pi)^{\frac{1}{2}nk} \, 2^k} \prod_{\nu=1}^k A(n - \nu + 1) \, | X'X |^{\frac{1}{2}(n-k-1)} \\ e^{-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}X'X)} \prod_{i \leq j} d(x_i'x_j) \end{aligned}$$

where $A(\nu)$ is given in (5.9).

8.4. *The general decomposition.* Formula (8.19) shows that the distribution (2.1) of a normal multivariate sample can be decomposed into two independent distributions, a Wishart distribution and an invariant distribution on the Stiefel manifold.

To split off the distribution of the plane spanned by the columns of X, we decompose the differential form

$$\prod a'_j \, da_i \prod \prod b'_j da_i \tag{8.21}$$

for the invariant measure in the Stiefel manifold into two differential forms representing independent distributions, namely, the invariant measure on the Grassmann manifold and the invariant measure on an orthogonal group of order k. The decomposition is given by equation (5.22).

From (5.22) and (8.19) we have the complete decomposition of the distribution (2.1):

$$dF(X) = \frac{|\Sigma|^{-\frac{1}{2}n}}{(2\pi)^{\frac{1}{2}nk} 2^k} e^{-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}X'X)} |X'X|^{\frac{1}{2}(n-k-1)} \prod_{i \le j}^{k} d(x'_i x_j) \cdot \prod_{j=1}^{n-k} \prod_{i=1}^{k} b'_j \, dh_i , \prod_{i<j}^{k} c'_j \, dc_i , \tag{8.22}$$

which is the result stated in Theorem 8.1.

In the univariate case, the Wishart distribution becomes the χ^2, the invariant measure on the Grassmann manifold becomes the element of area on the unit n-sphere, and the third factor disappears.

9. Acknowledgements. The author expresses his gratitude to Dr. H. Schwerdtfeger, of Melbourne University, to Dr. E. A. Cornish, of the C. S. I. R. O. Adelaide, and to Professors S. S. Wilks and W. Feller of Princeton University for their help and encouragement.

REFERENCES

[1] M. S. Bartlett, "The vector representation of a sample," *Proc. Cambridge Philos. Soc.*, Vol. 30 (1934), pp. 327–340.

[2] W. Blaschke, *Integralgeometrie*, Actualités Scientifique et Industrielles, Vol. 252, Hermann, Paris, 1935.

[3] S. S. Chern, "On Grassmann and differential rings and their relation to the theory of multiple integrals," *Sankhyā*, Vol. 7 (1945), pp. 1–8.

[4] C. Chevalley, *The Theory of Lie Groups*, Princeton University Press, 1946.

[5] R. A. Fisher, "The sampling distribution of some statistics obtained from non-linear equations," *Ann. Eugenics*, Vol. 9 (1939), pp. 238–249.

[6] M. A. Girshick, "On the sampling theory of the roots of determinantal equations," *Ann. Math. Stat.*, Vol. 10 (1939), pp. 203–224.

[7] E. Goursat, *Leçons sur le Problème de Pfaff*, Hermann, Paris, 1922, chap. 3.

[8] H. Hotelling, "Relations between two sets of variates," *Biometrika*, Vol. 28, (1936), pp. 321–377.

[9] P. L. Hsu, "On the distribution of the roots of certain determinantal equations," *Ann. Eugenics*, Vol. 9 (1939), pp. 250–258.

[10] A. Hurwitz, Über die Erzeugung der Invarianten durch Integration (1897) Mathematische Werke II, Birkhäuser, Basel (1933), pp. 546–564.

[11] E. Kähler, *Einführung in die Theorie der Systeme von Differential-gleichungen*, Chelsea, Berlin, 1934.

[12] C. C. MacDuffee, *Vectors and Matrices*, Math. Assoc. Am., 1943.

[13] A. M. Mood, "On the distribution of the characteristic roots of normal second-moment matrices," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 266–273.

[14] I. Olkin, "On Distribution Problems in Multivariate Analysis," Institute of Statistics, Mimeo. Series No. 43, 1951.

[15] I. Olkin and W. L. Deemer, "The Jacobians of certain matrix transformations useful in multivariate analysis," *Biometrika*, Vol. 38 (1951), p. 345.
[16] S. N. Roy, "p-statistics or some generalizations in analysis of variance appropriate to multivariate problems," *Sankhyā*, Vol. 4 (1939), pp. 381–396.
[17] S. N. Roy, "A note on critical angles between two flats in hyperspace with certain statistical applications," *Sankhyā*, Vol. 8 (1947), pp. 177–194.
[18] N. Steenrod, *The Topology of Fibre Bundles*, Princeton University Press, 1951.
[19] A. Weil, *L'Intégration dans les Groupes Topologiques et ses Applications*, Actualités Scientifique et Industrielles, Hermann, Paris, 1953.