The Annals of Statistics
1990, Vol. 18, No. 2, 685–716

VOLUMES OF TUBULAR NEIGHBORHOODS OF SPHERICAL POLYHEDRA AND STATISTICAL INFERENCE[1]

BY DANIEL Q. NAIMAN

The Johns Hopkins University

For statistical procedures including Scheffé-type simultaneous confidence bounds for response surfaces and likelihood ratio tests for an additional regressor with unspecified parameters in a regression model, the confidence level or size can be expressed in terms of probabilities of the form $P[U \in D(\Gamma, \theta)]$, where Γ is a subset of S^m (the unit sphere in R^{m+1}), U is uniformly distributed in S^m and $D(\Gamma, \theta)$ denotes the tubular neighborhood of Γ of angular radius θ, the set of points in S^m whose angular distance from Γ is at most θ. Consequently, determining critical points involves the calculation of the *volumes of tubes*. For the case when Γ is the diffeomorphic image of an r-dimensional convex polytope, an upper bound is given for the volume of its tubular neighborhood when the tube radius is sufficiently small, and which is exact in some special cases. Even if the tubular radius is moderate in size, the expression can be used to approximate the volume. The volume expression is a sum of r-fold integrals, one corresponding to each face of the polytope, and is derived using a result of Weyl (1939), which gives the volume of a tubular neighborhood of a k-dimensional submanifold of the unit sphere. Use of the expression leads to conservative statistical procedures when the desired error probability is sufficiently small and to asymptotically valid procedures as the error probability goes to zero.

1. Introduction. For certain statistical procedures that arise in the multiple regression setting, critical probabilities can be expressed in terms of probabilities of the form $P[U \in D(\Gamma, \theta)]$, where Γ is a given subset of S^m (the unit sphere in R^{m+1}), U is uniformly distributed in S^m and $D(\Gamma, \theta)$ is the so-called *tubular neighborhood* of Γ of angular radius θ, that is,

$$D(\Gamma, \theta) \equiv \left\{u \in S^m \colon \inf_{\gamma \in \Gamma} \cos^{-1}(\gamma^t u) \le \theta\right\},$$

the set of points in S^m whose angular distance from Γ is at most θ. Finding critical probabilities reduces to the *volume of tubes problem* for Γ, that of determining the m-dimensional volume of the tubular neighborhoods of Γ for varying tube radii.

There has been a considerable amount of interest in the volume of tubes problem going back to Steiner (1840) and Bertrand and Diguet (1848), who obtained partial results for the case when Γ is a curve. Hotelling (1939), motivated by statistical applications, proved that the volume of the tubular neighborhood of a closed curve in S^m or R^m is the product of the length of the

Received November 1987; revised July 1989.
[1]Research supported in part by The Office of Naval Research, Grant No. N00014-79-C-0801.
AMS 1980 *subject classifications*. Primary 60E15, 62J01; secondary 62J05, 62F25, 60D05.
Key words and phrases. Simultaneous confidence bounds, response surfaces, likelihood ratio tests, differential geometry, convex geometry, volumes of tubes.

curve and its cross-sectional $(m - 1)$-dimensional volume, provided the curve does not overlap itself and the tube radius is sufficiently small. Weyl (1939) generalized the Hotelling result by giving an analogous formula for the volume of the tubular neighborhood of a submanifold of arbitrary dimension in S^m or R^m. His formula is an r-fold integral, where r is the dimension of the embedded submanifold, which is somewhat surprising since one would naively expect the volume to be an m-fold integral. This reduction in the computational complexity of the problem can be an important consideration for the statistical applications described below where m, the dimension of the ambient space, grows with the number of observations.

The statistical procedures described in Section 2 have type I error probabilities which can be evaluated once the volume of tubes problem is solved for a corresponding set Γ. In Section 2.1, Γ represents the set of potential settings of the predictor variables in a regression model and the problem is to construct simultaneous confidence intervals for the regression function evaluated at each point in Γ. For the problem described in Section 2.2, one wishes to test whether an additional term that depends on some unknown vector-valued parameter ought to appear in a regression model, and Γ represents the set of possible values for the portion of the expected response vector corresponding to the hypothesized term, for a given design. In either case, the formulas of Hotelling (1939) and Weyl (1939) may not be immediately applicable because additional constraints on Γ give it a boundary and it is no longer a submanifold. One may choose to ignore this constraint information, but incorporating it can lead to smaller volumes, hence to narrower confidence intervals or more powerful tests.

A solution to the volume of tubes problem for more complicated sets Γ is given in Section 3, where convex geometry and differential geometry are used to give a volume formula analogous to the one in Weyl (1939). The sets Γ dealt with here are those which can be represented as the image of an r-dimensional polytope $Q \subseteq R^r$ under an infinitely differentiable transformation Φ, where Φ extends to an open set U containing Γ and has nonvanishing Jacobian. The formula, like Weyl's turns out to depend on the embedding Φ, only via the induced Riemannian metric on U, that is, on the positive definite matrix $\langle \partial\Phi/\partial u_i, \partial\Phi/\partial u_j \rangle$ defined at each point in U. The formula is guaranteed to give an upper bound for the volume when the tube radius is sufficiently small. In many situations the formula is exact although general conditions guaranteeing this are difficult to give. In any case, for the applications below upper bounds for the volume lead to upper bounds for type I error probabilities and hence to conservative procedures. These bounds are iterated integrals over simplices of smooth functions whose Monte Carlo evaluation should lead to better coverage probability estimates than those based on the naive Monte Carlo hit-or-miss method, for small tube radii. This is because in the hit-or-miss approach the function being integrated is the indicator function of a small subset of Euclidean space, which gives the estimate a high degree of variability. Methods for estimating the critical tube radius when the formula breaks down are discussed in Section 4. A fairly easy to calculate upper bound for this critical radius is given.

Formula (3.8) combined with Theorem 3.3 of Section 3 generalizes a weakened form of the inequality in Naiman (1986), whose proof has been simplified by Johnstone and Siegmund (1989). There Γ is the image under a piecewise differentiable map of an interval, and the tubular neighborhood contains two additional *caps* at the endpoints of the curve. The inequality states that when the Hotelling formula is modified by adding the additional cap volume, the resulting formula gives an upper bound for *all* tube radii. The complete generalization of the result in Naiman (1986) to higher-dimensional sets Γ, that is, the problem of finding a formula which yields an upper bound for the volume for *all* tube radii, remains open.

For additional statistical applications of the volume of tubes problem besides the ones mentioned in Section 2, the reader is referred to the work of Johansen and Johnstone (1990), Knowles (1987) and Knowles and Siegmund (1990). The Weyl (1939) formula has also been used to give probabilistic analyses of certain numerical analysis problems by Demmel (1988).

2. Applications to statistical inference.

2.1. *Simultaneous confidence bounds for regression functions.* The use of predictor variable constraints for obtaining improved Scheffé-type simultaneous confidence bounds for a response surface, under the usual multiple regression model assumptions, has been of interest for some time, since the Working and Hotelling (1929) and Scheffé (1953, 1959) procedures have been recognized as overly conservative. Indeed, progressively more and more complex types of predictor variable constraint sets have been found to lead to Scheffé-type bounds with tractable coverage probabilities since Scheffé introduced his method, which is appropriate when the predictor variables are constrained to lie in a linear subspace.

Consider the typical multiple regression model where one observes

$$Y = A\beta + e, \tag{2.1}$$

where β is $k \times 1$ and unknown, A is a known full-rank $n \times k$ matrix, $\nu \equiv n - k > 0$ and $e \sim N_n(0, \sigma^2 I_n)$, with σ unknown. The results described below have obvious analogues under the weaker assumption that ϵ has a spherically symmetric distribution. Suppose one is interested in investigating the behavior of the regression function $E(y) = x^t\beta$ for various settings of the predictor variables, that is, for x ranging throughout a given constraint set $X \subseteq R^k$. One may construct confidence intervals for $x^t\beta$ of the form

$$J_x = \left[x^t\hat{\beta} - c\hat{\sigma}\sqrt{x^t(A^tA)^{-1}x}\,,\, x^t\hat{\beta} + c\hat{\sigma}\sqrt{x^t(A^tA)^{-1}x}\right], \quad \forall\, x \in X, \tag{2.2}$$

where $\hat{\beta}$ denotes the least squares estimator of β, $\hat{\sigma}$ denotes the usual root mean squared error estimator of σ and c is a positive constant. Since one can, if necessary, replace A by AP^{-1}, β by $P\beta$ and X by $(P^t)^{-1}X$, where P is a $k \times k$ matrix with $P(A^tA)^{-1}P^t = I_k$, it will be assumed without loss of generality that the design is orthogonal, i.e., $A^tA = I_k$.

Collections of confidence intervals of the form (2.2) are referred to as Scheffé (1953, 1959)-type simultaneous confidence intervals (SCI's) over X. One-sided confidence intervals, obtained by setting the left (right) endpoint of the intervals J_x to $-\infty$ $(+\infty)$ may also be considered, and the results of this paper apply to that situation as well. One important measure of overall performance for such a collection is their *simultaneous coverage probability*, namely,

$$p(X, c) = P\left[x^t\beta \in J_x, \forall\, x \in X\right]. \tag{2.3}$$

The family (2.2) is said to achieve a *simultaneous confidence level* of $1 - \alpha$ if $p(X, c) \geq 1 - \alpha$. This paper gives a lower bound for the simultaneous coverage probability of Scheffé-type SCI's which is appropriate for a large class of constraint sets. This bound allows for the construction of narrower Scheffé-type SCI's over given X with at least some prescribed confidence level.

The Scheffé (1953, 1959) method for constructing simultaneous confidence intervals over given X with prescribed confidence level $1 - \alpha$ is based on the following observation. When X is a linear subspace of dimension q in R^k, the coverage probability (2.3) is given by $F_{q, n-k}(c^2/q)$, so that one may use $c = \sqrt{qF^{-1}_{q, n-k}(1 - \alpha)}$ to obtain a family of SCI's over X with coverage probability exactly $1 - \alpha$. (This result is a simple consequence of Lemma 2.1.) Even if X is not a linear subspace, one may replace X by its linear span and use Scheffé's method. In many applications, when X is replaced by its linear span the resulting coverage probability calculation is simplified, but one pays a price, since the larger constraint set leads to wider confidence intervals. A great deal of research effort has been directed toward finding other constraint regions for which $p(X, c)$ remains tractable. See Halperin and Gurian (1968), Bohrer (1967), Bohrer and Francis (1972), Wynn and Bloomfield (1971), Cassella and Strawderman (1980), Uusipaikka (1984), Wynn (1975), Naiman (1986, 1987) and Knafl, Sacks and Ylvisaker (1985), for example.

Constraint sets for predictor variables arise naturally in experimental work. For example, it is often the case that the experimenter can specify the practical range of values for each predictor variable by giving a lower and upper bound for each, so that the resulting constraint set is a rectangular region. Sometimes the predictor variables are fractions of components in a mixture and the constraint set is simplex [see Cornell (1981)]. A large class of constraint sets, which can be handled using the method described in this paper, arises in the following manner. Suppose there are several predictor variables $x_1, \ldots, x_r$ and the regression function is a linear combination of known functions $f_1(x_1, \ldots, x_r), \ldots, f_k(x_1, \ldots, x_r)$. For example, these might be polynomials or trigonometric functions. If each predictor x_j is constrained to lie in an interval I_j, then the constraint set for bounding the regression function is the deformed hyperrectangle

$$\{(f_1(x_1,\dots,x_r),\dots,f_k(x_1,\dots,x_r)) \colon x_j \in I_j,\ j = 1,\dots,r\}.$$

Much more complicated constraint sets might be envisioned for a complex experiment, but even most of the simplest examples thus far have been considered unwieldy.

The connection between the simultaneous inference problem and the volume of tubes is made in the following lemma. A proof may be found in Naiman (1986).

LEMMA 2.1 [Uusipaikka (1984)]. *For the model* (2.1), *define* $\Gamma \equiv \{\|x\|^{-1}x\colon x \in X\} \subseteq S^{k-1}$ *and let* $-\Gamma$ *denote* $\{-\gamma\colon \gamma \in \Gamma\}$. *Then for the family* (2.2) *of confidence intervals, the coverage probability is given by*

$$p(X,c) = 1 - \int_{\theta=0}^{\pi/2} P[U \in D(\Gamma \cup -\Gamma,\theta)]g(\theta)\,d\theta, \tag{2.4}$$

where U *has a uniform distribution in* S^{k-1} *and*

$$g(\theta) = f_{\nu,k}\left(\frac{k\cos^2\theta}{c^2}\right)\frac{2k\cos\theta\sin\theta}{c^2},$$

where $f_{\nu,k}$ *denotes the density function of the* $F_{\nu,k}$ *distribution.*

The Scheffé (1953, 1959) method for bounding *all* linear combinations of the parameters corresponds to the case when $\Gamma = S^{k-1}$, so that the exact coverage probability is obtained by replacing $P[U \in D(\Gamma \cup -\Gamma,\theta)]$ in the integrand in (2.4) by 1. For smaller constraint sets Γ, following Naiman (1986), a lower bound for $p(X,c)$ is obtained by ignoring the intersection of the antipodal sets $D(\Gamma,\theta)$ and $D(-\Gamma,\theta)$ and replacing the term $P[U \in D(\Gamma \cup -\Gamma,\theta)]$ in the integrand in (2.4) by $2P[U \in D(\Gamma,\theta)]$. Then

$$\begin{aligned} p(X,c) \geq 1 &- \int_{\theta=0}^{\zeta^*} 2P[U \in D(\Gamma,\theta)]g(\theta)\,d\theta \\ &- \int_{\theta=\zeta^*}^{\pi/2} P[U \in D(\Gamma \cup -\Gamma,\theta)]g(\theta)\,d\theta \end{aligned} \tag{2.5}$$

for any $\zeta^* \in [0,\pi/2]$. If Γ is regular in the sense described in Section 3.1 and if ζ^* denotes the critical radius in Section 4, then the first integrand can be replaced using (3.8) and the volume formula of Theorem 3.3 and the second can be replaced by a crude upper bound to yield a lower bound for $p(X,c)$. Even if the upper bound of 1 is used for the probability in the second integral the resulting procedure is guaranteed to improve on Scheffé's. Asymptotically, as the desired coverage probability converges to 1 the weight function g is concentrated more in the range $[0,\zeta^*]$, so the improvement becomes more pronounced.

2.2. *Testing for additional structure in a regression model.* The connection described in this section between the problem of testing for additional terms in a regression model and the volume tubes was pointed out by Hotelling (1939). For recent related work see Knowles and Siegmund (1990). Consider the regression model

$$Y_i = x_i^t\beta + cf(x_i, \theta) + e_i, \qquad i = 1, \dots, n, \tag{2.6}$$

where $\beta \in R^k$, $c \in R$, $\theta \in \Theta \subset R^p$ are unknown parameters, the design points x_i are assumed to be known and lie in a set $X \subseteq R^k$, f is a known function defined in $X \times \Theta$ and the e_i are iid $N(0, \sigma^2)$ with σ unknown.

The problem is to find critical points for the likelihood ratio test of H: $c = 0$ vs. A: $c \neq 0$. For example, Hotelling (1939) considered testing whether there is a sinusoidal component of unknown amplitude, phase and frequency, in time series data, assuming a model of the form $Y_i = \beta_0 + c\sin(\omega t_i + \gamma) + e_i$. Here $\theta = (\omega, \gamma)$ is an unknown point in R^2.

Define column n-vectors $X_j \equiv [x_{1j}, \dots, x_{nj}]^t$, for $j = 1, \dots, k$, $T_\theta \equiv [f(x_1, \theta), \dots, f(x_n, \theta)]^t$, $Y \equiv [Y_1, \dots, Y_n]^t$ and let V denote the subspace of R^n spanned by the X_j. Geometrically, the null hypothesis states that the expectation of Y lies in V, while under the alternative hypothesis, we modify this expectation by adding a vector of the form cT_θ. In order to make the parameters identifiable, it is assumed that $T_\theta \notin V$ for all $\theta \in \Theta$ and that the vectors T_θ and $T_{\theta'}$ are not positive multiples of one another for $\theta \neq \theta'$. The problem can be put into a *canonical form* by subtracting from each T_θ its projection onto V, and it will be assumed without loss of generality that $T_\theta \in V^\perp$ for all $\theta \in \Theta$, where $V^\perp$ denotes the orthogonal complement of V.

The likelihood ratio test rejects H when the ratio of error sums of squares

$$\frac{\inf_{\beta, c, \theta}\|Y - \sum_{j=1}^k \beta_j X_j - cT_\theta\|^2}{\|Y - \sum_{j=1}^k \hat{\beta}_j X_j\|^2} \tag{2.7}$$

is sufficiently small, where $\hat{\beta}$ denotes the maximum likelihood estimator of β under the null hypothesis. Using the assumptions above the test statistic (2.7) becomes

$$\begin{aligned}
&\inf_{\beta, c, \theta}\left\{\left\|Y - \sum_{j=1}^k \beta_j X_j\right\|^2 - 2\langle Y, cT_\theta\rangle + \|cT_\theta\|^2\right\}/\|\hat{e}\|^2 \\
&= \inf_{c, \theta}\left\{\left\|Y - \sum_{j=1}^k \hat{\beta}_j X_j\right\|^2 - 2\langle Y, cT_\theta\rangle + \|cT_\theta\|^2\right\}/\|\hat{e}\|^2 \\
&= \inf_{c, \theta}\left\{\left\|Y - \sum_{j=1}^k \hat{\beta}_j X_j\right\|^2 - 2\left\langle Y - \sum_{j=1}^k \hat{\beta}_j X_j, cT_\theta\right\rangle + \|cT_\theta\|^2\right\}/\|\hat{e}\|^2 \\
&= \inf_{c, \theta}\|\hat{e} - cT_\theta\|^2/\|\hat{e}\|^2 = 1 - \sup_{\gamma \in \Gamma}\left(\gamma^t U\right)^2,
\end{aligned}$$

where $\hat{e}$ denotes the residual vector $\hat{e} \equiv Y - \sum_{j=1}^{k}\hat{\beta}_j X_j$, $U \equiv \hat{e}/\|\hat{e}\|$ and $\Gamma \equiv \{T_\theta/\|T_\theta\|: \theta \in \Theta\} \subseteq S^{n-1}$. Thus, the test rejects H when $U \in D(\Gamma, \theta)$ for some $\theta > 0$.

To determine critical points for the test, one must be able to evaluate the volume of $D(\Gamma, \theta)$ for given θ. If Γ is *regular* as in Definition 3.1, then Section 3 gives a method for bounding this volume for all sufficiently small θ and this allows for the possibility of constructing a conservative test. The condition that θ be sufficiently small corresponds to the requirement that the type I error be sufficiently small.

3. Volume of tubular neighborhoods. Many terms from convex geometry including convex hull, conical hull, convex polyhedron, convex cone, polar cone and cone of feasible directions are used below. Although the definitions are fairly standard, they and some basic results are reviewed in Appendix A for the reader who is unfamiliar with them. Stoer and Witzgall (1970) give a thorough treatment of the subject matter that appears in this appendix. It is assumed that the reader is familiar with basic concepts in differential geometry, for which the author found Hicks (1971), Boothby (1986) and Spivak (1979) to be particularly helpful.

3.1. *A class of constraint sets.* A class of constraints sets is defined now. To fix some terminology, for a manifold M and any point $x \in M$, the tangent space of M at x will be denoted by M_x. If f: $M \to N$ is a C^∞ mapping between manifolds and $x \in M$, the differential of f, which is the linear mapping f induces from M_x to $N_{f(x)}$, will be denoted by f_*. The reader who is unfamiliar with differential geometry is warned that while this mapping and others (e.g., vector fields) implicitly depend on a "base point," typically this is not indicated so as to simplify the notation.

Definition 3.1. $\Gamma \subseteq S^m$ is called *regular* if for some r there exists a *bounded* r-dimensional polyhedron $Q \subseteq R^r$, an open set $U \supseteqq Q$ and a C^∞ mapping Θ: $U \to S^m$ such that $\Phi_*: U_x \to S^m_{\Phi(x)}$ is one-to-one for every $x \in U$ (so that Φ is locally one-to-one) and $\Phi(Q) = \Gamma$.

It is necessary to introduce more terminology. The natural identification of the tangent space U_x with R^r and the corresponding identification of $(S^m)_u$ with a subspace of R^{m+1} will be used repeatedly. This allows one to view $\Phi_*: U_x \to (S^m)_{\Phi(x)}$ as a mapping from R^r into R^{m+1}. The Riemannian metric on S^m is taken to be the one it inherits as a submanifold of R^{m+1} with its usual metric, so the inner product between vectors in $(S^m)_{\Phi(x)}$ is the usual inner product between vectors in R^{m+1}. Φ_* defines a vector space isomorphism between U_x and its image for each $x \in U$. The reader may find explicit use of the usual coordinate systems in R^r and R^{m+1} illuminating, and their use will be important in actual calculations. The coordinate mappings in R^r will be denoted by $x_1, \ldots, x_r$, so that the associated tangent vectors $\partial/\partial x_i$, $i = 1, \ldots, r$, form a basis of U_x.

3.2. *A decomposition for tubular neighborhoods.* Fix a regular set $\Gamma \subseteq S^m$ for the remainder of Section 3 and fix Q, U and Φ as in Definition 3.1. Since U can be taken to be contractible and this property will be needed in Section 3.5, it will be assumed that this is indeed the case. Assume

$$Q = \{x \in R^r: \langle v(i), x\rangle \le c_i, i = 1, \dots, q\},$$

where $v(1), \dots, v(q) \in R^r$ and $c_1, \dots, c_q \in R$ and $\langle\ ,\ \rangle$ denotes the usual inner product in R^r. Each $x \in Q$ lies in an open face of Q, that is, a set of the form

$$F_I = \{x \in R^r: \langle v(i), x\rangle = c_i, \forall\, i \in I \text{ and } \langle v(i), x\rangle < c_i, \forall\, i \notin I\},$$

for some $I \subseteq \{1, \dots, q\}$ [see Appendix A (A.3)]. Two subsets of indices may define the same face. Let $\mathscr{F}$ denote the set of nonempty faces F_I. If $x \in F_I$, then by Proposition A.3(a) (Appendix A), the cone of feasible directions at x relative to Q, that is, the set of vectors that point into Q from x, is the polyhedral cone

$$C_{I,x} = \{u \in R^r: \langle v(i), u\rangle \le 0, \forall\, i \in I\}.$$

Identical cones are obtained as x varies throughout F_I, so at times there should be no confusion if they are denoted by C_I.

DEFINITION 3.2. A point $u = \Phi(x)$, for $x \in F_I$, is said to be a *projection of* $v \in S^m$ *onto* Γ *associated with the face* F_I if u minimizes the angular distance $\rho(u, v) \equiv \cos^{-1}(u^t v)$ among all points in $\Phi(Q \cap B(x, \varepsilon))$, for some $\varepsilon > 0$, where $B(x, \varepsilon)$ denotes the open ball centered at x with radius ε. The set of projections of v onto Γ associated with F_I is denoted by $\pi_{F_I}(v)$.

REMARK 3.1. Every point $v \in S^m$ has a projection associated with some face. To see this, let u be any closest point in Γ to v. By (A.3), $u \in \Phi(F_I)$, for some $I \subseteq \{1, \dots, q\}$, and it follows that $u \in \pi_{F_I}(v)$. Furthermore, v may possess multiple projections which may be associated with either the same face or with different faces.

The tubular neighborhood $D(\Gamma, \theta)$ decomposes into a union of "facial neighborhoods." To be precise, define

$$D^*(F_I, \theta) = \left\{v \in D(\Gamma, \theta): \inf_{u \in \pi_{F_I}(v)} \rho(u, v) \le \theta\right\}$$

for $I \subseteq \{1, \dots, q\}$. Then

$$D(\Gamma, \theta) = \bigcup_{F_I \in \mathscr{F}} D^*(F_I, \theta), \tag{3.1}$$

by Remark 3.1. The sets making up this union need not be disjoint. As a consequence of (3.1), the m-dimensional volume of $D(\Gamma, \theta)$ satisfies

$$\operatorname{Vol}(D(\Gamma, \theta)) \le \sum_{F_I \in \mathscr{F}} \operatorname{Vol}(D^*(F_I, \theta)). \tag{3.2}$$

The uniform probability measure of a given subset of S^m is obtained by dividing its volume by $\mathrm{Vol}(S^m)$. As a matter of fact, $\mathrm{Vol}(S^0) = 2$, $\mathrm{Vol}(S^1) = 2\pi$ and the recursion formula $\mathrm{Vol}(S^{m+2}) = (2\pi/(m + 1))\mathrm{Vol}(S^m)$ is satisfied for $m \geq 2$.

The reader may find (3.2) more illuminating by examining Figure 1, which illustrates the tubular neighborhood decomposition for the cases when Γ is a spherical arc [the case considered in Naiman (1986)] and when Γ is a spherical triangle. There is not much intuition lost if these sets are viewed as subsets of Euclidean space.

3.3. *Covering of facial neighborhoods.* The main result of this section is Theorem 3.1, which gives a set $D^{**}(F_I, \theta)$ that contains and approximates $D^*(F_I, \theta)$ and has a convenient coordinatization. The facial neighborhoods of the previous section have been defined by minimization of distance. There is another process by which one may describe points in these sets and which is used to construct the sets $D^{**}(F_I, \theta)$. For a fixed face F_I and a point $x \in F_I$, a point in $D^*(F_I, \theta)$ closest to $\Phi(x)$ is obtained by moving along a geodesic whose initial point is $\Phi(x)$ and whose tangent at $\Phi(x)$ lies in a certain polyhedral cone. Much of the work leading up to Lemma 3.2 involves finding a useful description for this cone.

For this section fix $I \subseteq \{1, \dots, q\}$ for which the corresponding face F_I is nonempty.

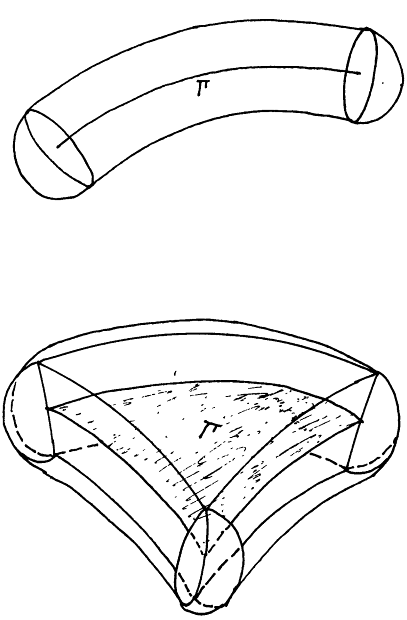

FIG. 1. *The tubular neighborhood decomposition for a spherical arc and a spherical triangle.*

DEFINITION 3.3. ω: $[0,1] \to S^m$ is referred to as a *great circular arc connecting u to v* if there exists an orthogonal transformation T on R^{m+1} such that $T(u) = (1, 0, \dots, 0)$, $T(v) = (\cos(\psi), \sin(\psi), 0, \dots, 0)$ for some $\psi \in [0, 2\pi]$ and $T(\omega(t)) = (\cos(t\psi), \sin(t\psi), 0, \dots, 0)$ for all $t \in [0,1]$.

The following facts are easy to verify. For any pair of points u and v in S^m with $\rho(u,v) \in (0,\pi)$, there are exactly two great circular arcs connecting u to v, one whose length is $\rho(u,v)$, which is the smallest length of any arc connecting u and v, and one whose length is $2\pi - \rho(u,v)$. If $\rho(u,v) = \pi$, then there are infinitely many great circular arcs connecting u to v, all of which have the same length. For any $u \in S^m$ and $\theta \in (0, 2\pi)$ and for any nonzero tangent vector $Y \in (S^m)_u$, there exists a unique great circular arc ω: $[0,1] \to S^m$, which fits Y, that is, $\omega(0) = u$ and $\omega'(0)$ is a positive multiple of Y and this arc has length θ.

LEMMA 3.1. *Let $v \in S^m$ and suppose $u = \Phi(x) \in \pi_{F_I}(v)$, where $x \in F_I$. Then $\Phi_*(C_{I,x})$ is a polyhedral cone in $(S^m)_{\Phi(x)}$ and if ω is a great circular arc connecting u to v, then $\omega'(0) \in \Phi_*(C_{I,x})^p$, where $\Phi_*(C_{I,x})^p$ denotes the polar cone to $\Phi_*(C_{I,x})$, that is,*

$$\Phi_*(C_{I,x})^p = \{W \in (S^m)_{\Phi(x)} \colon \langle Z, W \rangle \le 0, \text{ for all } Z \in \Phi_*(C_{I,x})\}.$$

PROOF. The first claim follows immediately from the fact that Φ_* is a linear transformation and $C_{I,x}$ is a polyhedral cone. Let $W \in C_{I,x}$ so that $\eta(t)$ defined by $x + tW$ is in Q for all $t \in [0,\delta]$ for some $\delta > 0$. Then $\Phi \circ \eta$: $[0,\delta] \to \Gamma$ is an arc with $\Phi \circ \eta(0) = u$. Since $u \in \pi_{F_I}(v)$, there exists $\tau \in (0,\delta)$ such that $\rho(\Phi \circ \eta(t), v) \ge \rho(u,v)$ for all $t \in [0,\tau]$. It follows easily that $\rho(\Phi \circ \eta(t), \omega(s)) \ge \rho(u, \omega(s))$ for all $t \in [0,\tau]$ and $s \in [0,1]$. Using Taylor expansions about $t = 0$ and $s = 0$, it follows that $\langle (\Phi \circ \eta)'(0), \omega'(0) \rangle \le 0$. The result then follows since $(\Phi \circ \eta)'(0) = \Phi_*(\eta'(0)) = \Phi_*(W)$. □

The following notation will be used below. Any $u \in S^m$ and any tangent vector $Y \in (S^m)_u$ may be viewed as elements of R^{m+1}, whose sum $u + Y$ (which, when normalized by dividing by its Euclidean length, defines an element of S^m) will be denoted by $[u + Y]$. Thus, $[u + Y]$ is obtained by moving in S^m the angular distance $\tan^{-1}(\|Y\|)$ from u along the great circular arc obtained by projecting the ray $\{u + tY \colon t \ge 0\}$ in R^{m+1} onto the unit sphere.

THEOREM 3.1 (Facial neighborhood covering theorem). *Define*

$$D^{**}(F_I, \theta) = \{[\Phi(x) + Y] \colon x \in F_I,\ Y \in \Phi_*(C_{I,x})^p \text{ and } \|Y\| \le \tan(\theta)\}$$

for $0 \le \theta \le \pi/2$. Then $D^*(F_I, \theta) \subseteq D^{**}(F_I, \theta)$.

PROOF. If $v \in D^*(F_I, \theta)$, fix $u = \Phi(x) \in \pi_{F_I}(v)$, where $x \in F_I$, such that $\rho(u, v) \leq \theta$. If $\rho(u, v) = 0$, then the result is immediate, so assume $\rho(u, v) > 0$. Let ω be the great circular arc which connects u to v and define $Y = (\tan(\rho(u, v))/\|\omega'(0)\|)\omega'(0)$. Clearly, $v = [\Phi(x) + Y]$ and $\|Y\| = \tan(\rho(u, v)) \leq \tan(\theta)$. It follows from Lemma 3.1 that $\omega'(0) \in \Phi_*(C_{I,x})^p$ and hence $Y \in \Phi_*(C_{I,x})^p$. $\square$

3.4. *Decomposition of the covering sets.* In this section, the cones $\Phi_*(C_{I,x})^p$ appearing in the definition of $D^{**}(F_I, \theta)$ are described in more detail. Each $\Phi_*(C_{I,x})^p$ can be expressed as a sum of a cone contained in the tangent space $\Phi(U)_{\Phi(x)}$ and a cone consisting of vectors normal to this tangent space. In fact, a stronger statement is true: There are smooth vector fields defined along F_I that span these two cones. Such vector fields are crucial for coordinatizing the sets $D^{**}(F_I, \theta)$.

The reader may find it helpful to examine Figure 2 which shows $D^{**}(F_I, \theta)$ for one of the edges of a spherical triangle. Note that the set is spanned by two vector fields along the edge of the triangle, where one of the vector fields Ψ is actually tangent to the surface defined by the triangle and the other N is normal to the surface. An arbitrary point in $D^{**}(F_I, \theta)$ can be represented as the sum of a point x in the edge, a positive multiple of Ψ and a multiple of N. Fix $I \subseteq \{1, \ldots, q\}$ corresponding to a nonempty face F_I for the remainder of this section.

Define $\Phi_\alpha = \Phi_*(\partial/\partial x_\alpha)$ for $\alpha = 1, \ldots, r$ and $x \in U$. Using the identification of $(S^m)_{\Phi(x)}$ with a subspace of R^{m+1}, Φ_α is the $(m+1)$ vector $\partial\Phi/\partial x_\alpha$. The vectors Φ_α span the tangent space $\Phi(U)_{\Phi(x)}$ of the submanifold $\Phi(U)$ at $\Phi(x)$ and they are linearly independent since Φ_* is one-to-one.

Define the coefficients of the induced Riemannian metric in terms of the above coordinate vectors $\bar{g}_{\alpha\beta} = \langle \Phi_\alpha, \Phi_\beta \rangle$, for $x \in U$. The *bar* over the g (and over any expression defined in terms of $\bar{g}$) is used to distinguish this expression for the metric from the expression for the induced metric in a submanifold and in different coordinates used in Section 3.5. Note that $(\bar{g}_{\alpha\beta})$ is a positive definite matrix for each $x \in U$. Let the components of the inverse of

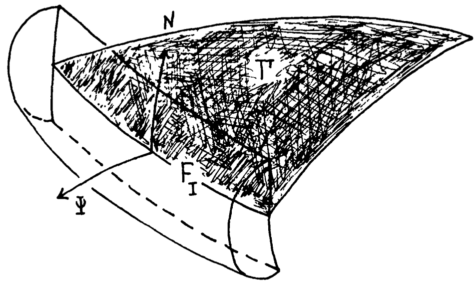

FIG. 2. *The set $D^{**}(F_I, \theta)$ when F_I is the edge of a spherical triangle, and the vector fields N and Ψ.*

this matrix be denoted by $\bar{g}^{\alpha\beta}$. Define a mapping Ψ: $U_x \to \Phi_*(U_x)$, for each $x \in U$ by

$$\Psi(W) = \sum_{\alpha,\beta=1}^{r} \bar{g}^{\alpha\beta} W^\alpha \Phi_\beta. \tag{3.3}$$

It is easy to verify that Ψ is the adjoint of $\Phi_*{}^{-1}$, the (nonsingular) linear transformation with the property

$$\langle W, Z\rangle = \langle \Psi(W), \Phi_*(Z)\rangle, \qquad \forall\, W, Z \in U_x.$$

It follows easily from (3.3) that

$$\langle \Psi(W), \Psi(Z)\rangle = \sum_{\alpha,\beta=1}^{r} \bar{g}^{\alpha\beta} W^\alpha Z^\beta$$

for all $W, Z \in U_x$.

REMARK 3.2. The projection of a given $Y \in (R^{m+1})_{\Phi(x)}$ onto the tangent space $\Phi(U)_{\Phi(x)}$ is given by the formula

$$\sum_{\alpha,\beta=1}^{r} \bar{g}^{\alpha\beta} \langle \Phi_\alpha, Y\rangle \Phi_\beta.$$

This fact, which is not difficult to verify, will be used repeatedly in Section 3.5.

LEMMA 3.2. *If $x \in F_I$, then $\Phi_*(C_{I,x})^p$ admits the direct sum decomposition*

$$\Phi_*(C_{I,x})^p = \Psi(C_{I,x}^p) \oplus \Phi_*(U_x)^\perp,$$

where $\Phi_*(U_x)^\perp$ denotes the orthogonal complement of the linear subspace $\Phi_*(U_x)$ in $(S^m)_{\Phi(x)}$.

PROOF. First note that

$$\begin{aligned}\Phi_*(C_{I,x}) &= \{\Phi_*(Y): Y \in U_x \text{ and } \langle v(i), Y\rangle \le 0, \forall\, i \in I\}\\ &= \{W \in \Phi_*(U_x): \langle v(i), \Phi_*{}^{-1}(W)\rangle \le 0, \forall\, i \in I\}\\ &= \{W \in \Phi_*(U_x): \langle \Psi(v(i)), W\rangle \le 0, \forall\, i \in I\}.\end{aligned}$$

By Lemma A.III(a), the polar cone of $\Phi_*(C_{I,x})$, as a subspace of $\Phi_*(U_x)$, is given by

$$\operatorname{con}\{\Psi(v(i)): i \in I\} = \Psi(\operatorname{con}\{v(i): i \in I\}), \tag{3.4}$$

where $\operatorname{con}(K)$ denotes the conical hull of K (see Appendix A). Since Lemma A.III(a) combined with the definition of $C_{I,x}$ gives

$$\operatorname{con}\{v(i): i \in I\} = C_{I,x}^p,$$

the set in (3.4) is $\Psi(C_{I,x}^p)$. Since $\Phi_*(C_{I,x}) \subset \Phi_*(U_x)$, the result follows. $\square$

The first term in the decomposition of Lemma 3.2 involves a linear mapping of the same cone since the cones $C_{I,x}^p$ are identified in a natural way. However, the linear map Ψ varies as x varies along F_I. Any fixed vector in the cone C_I^p maps in this manner to a vector field defined along F_I, and spanning vector fields of this form are used below to coordinatize the sets $D^{**}(F_I, \theta)$.

Every point in $D^{**}(F_I, \theta)$ can be expressed as a "sum" of a point $x \in F_I$ and a vector in the cone $\Phi_*(C_{I,x})^p$. To coordinatize $D^{**}(F_I, \theta)$ it suffices to coordinatize F_I and $\Phi_*(C_{I,x})^p$ separately. This process is carried out by the somewhat technical operation of decomposing F_I and $\Phi_*(C_{I,x})^p$ into *simplicial* pieces. The comments leading up to Theorem 3.2 show essentially that it suffices to consider the case when F_I is a *simplex* and C_I^p is a *simplicial cone*, the sum of whose dimensions is r.

To decompose $\Phi_*(C_{I,x})^p$, first partition C_I^p into simplicial cones and map this via Ψ to define a similar partition of $\Psi(C_I^p)$. Then apply Lemma 3.2. To be more precise, note that by Proposition A.3(e) and (f) (Appendix A) $C_I^p = \operatorname{con}\{v(i): i \in I\}$ is a pointed cone whose dimension is $r - \dim(F_I)$. Let $\mathscr{K}_I$ be a collection of subsets of C_I^p with each subset consisting of exactly $r - \dim(F_I)$ elements, such that:

P1. $\dim(\operatorname{con}(K)) = r - \dim(F_I)$ for each $K \in \mathscr{K}_I$.
P2. $\dim(\operatorname{con}(K) \cap \operatorname{con}(K')) < r - \dim(F_I)$ for all distinct $K, K' \in \mathscr{K}_I$.
P3. $C_I^p = \bigcup_{K \in \mathscr{K}_I} \operatorname{con}(K)$.

The existence of such a collection is guaranteed by Proposition A.2. In fact, $\mathscr{K}_I$ can be constructed from $\{v(i), i \in I\}$ by standard algorithms. Since Ψ is linear, $\operatorname{con}(\Psi(K)) = \Psi(\operatorname{con}(K))$ for each $K \in \mathscr{K}_I$. Since Ψ is nonsingular, $\Psi(K)$ consists of exactly $r - \dim(F_I)$ elements for each $K \in \mathscr{K}_I$. Furthermore:

P1′. $\dim(\operatorname{con}(\Psi(K))) = r - \dim(\mathrm{F}_I)$ for each $K \in \mathscr{K}_I$.
P2′. $\dim(\operatorname{con}(\Psi(K)) \cap \operatorname{con}(\Psi(K'))) < r - \dim(F_I)$ for all distinct $K, K' \in \mathscr{K}_I$.
P3′. $\Psi(C_{I,x}^p) = \cup_{K \in \mathscr{K}_I} \operatorname{con}(\Psi(K))$.

It will be necessary to make use of a partition of $\overline{F}_I$, the closure of F_I, into simplicial polytopes, which is guaranteed by Proposition A.1. Let $\mathscr{P}_I$ be a finite collection of subsets of $\overline{F}_I$, each consisting of $\dim(F_I) + 1$ elements, with the property that:

P1″. $\dim(\operatorname{conv}(P)) = \dim(F_I)$, for each $P \in \mathscr{P}_I$.
P2″. $\dim(\operatorname{conv}(P) \cap \operatorname{conv}(P')) < \dim(F_I)$ for all distinct $P, P' \in \mathscr{P}_I$.
P3″. $\overline{F}_I = \cup_{P \in \mathscr{K}_I} \operatorname{conv}(P)$.

$\operatorname{conv}(P)$ denotes the convex hull of P (see Appendix A).

THEOREM 3.2 (Covering set decomposition. *Let $\mathscr{K}_I$ and $\mathscr{P}_I$ be as given above. Define $D^{**}(P, K, \theta)$ to be the set*

$$\{[\Phi(x) + Y] : x \in \operatorname{conv}(P),\, Y \in \operatorname{con}(\Psi(K)) \oplus \Phi_*(R^r)^{\perp} \text{ and } \|Y\| \le \tan(\theta)\} \tag{3.5}$$

for all $0 \le \theta \le \pi/2$ *and* $(P, K) \in \mathscr{P}_I \times \mathscr{K}_I$. *Then*

$$D^{**}(F_I, \theta) \subset \bigcup_{(P, K) \in \mathscr{P}_I \times \mathscr{K}_I} D^{**}(P, K, \theta). \tag{3.6}$$

Furthermore, the two sets in (3.6) *differ by a set having zero volume and the terms in the union in* (3.6) *intersect in sets having zero volume.*

PROOF. See Appendix B. □

As a consequence of Theorem 3.2,

$$\mathrm{Vol}(D^{**}(F_I, \theta)) = \sum_{(P, K) \in \mathscr{P}_I \times \mathscr{K}_I} \mathrm{Vol}(D^{**}(P, K, \theta)). \tag{3.7}$$

Combining this with (3.2) and Theorem 3.1,

$$\mathrm{Vol}(D(\Gamma, \theta)) \le \sum_{F_I \in \mathscr{F}} \sum_{(P, K) \in \mathscr{P}_I \times \mathscr{K}_I} \mathrm{Vol}(D^{**}(P, K, \theta)). \tag{3.8}$$

Section 3.5 gives a method for bounding each of the terms in (3.8) provided θ is sufficiently small.

3.5. *Volume bounds.* For this section fix $I \subset \{1, \dots, q\}$ defining a nonempty face F_I, and fix $\mathscr{P}_I$, $\mathscr{K}_I$ and $(P, K) \in \mathscr{P}_I \times \mathscr{K}_I$ as in Section 3.4. An upper bound is developed below for the volume of $D^{**}(P, K, \theta)$ for sufficiently small $\theta \in [0, \pi/2]$, where $D^{**}(P, K, \theta)$ is defined in (3.5). The restriction on θ is analogous to the restriction appearing in Weyl (1939). For the case when F_I is the interior of Q, so that "locally" $\Phi(F_I)$ is a submanifold, the bound is the same as the volume expression in Weyl. The terms corresponding to the other (lower dimensional) faces are dealt with below by using a lemma due to Weyl and a determinantal identity.

The first step is to develop a coordinatization of $D^{**}(P, K, \theta)$. This is the analogue of the description in Figure 2; the result is given in Lemma 3.3. By assumption $P = \{p(j),\ j = 0, \dots, d\}$ for distinct $p(j) = (p(j)^1, \dots, p(j)^r)^t$ and conv(P) is a d-dimensional simplex whose relative interior will be denoted by $\tilde{P}$. Also

$$K = \left\{\kappa(i) = \left(\kappa(i)^1, \dots, \kappa(i)^r\right)^t, i = 1, \dots, r - d\right\}$$

and con(K) is $(r - d)$ dimensional. Since $K \subseteq C_I^p$ it follows easily that $p(j) - p(0) \perp \kappa(i)$ for $j = 1, \dots, d$ and $i = 1, \dots, r - d$. Define the standard open d-simplex

$$\mathscr{S}_d = \left\{\lambda = (\lambda^1, \dots, \lambda^d) \in R^d \colon \sum_{j=1}^{d} \lambda^j < 1, \text{and } \lambda^j > 0, \text{for } j = 1, \dots, d\right\}$$

and let $\Theta\colon \mathscr{S}_d \to \tilde{P}$, be the bijection defined by

$$\Theta(\lambda) = p(0) + \sum_{j=1}^{d} \lambda^j(p(j) - p(0)).$$

The usual tangent vectors associated with the coordinate mappings $\lambda_1, \dots, \lambda_d$ in R^d will be denoted by $\partial/\partial\lambda_j$, $j = 1, \dots, d$.

The set $\Phi(\tilde{P})$ is (locally) a d-dimensional submanifold of $\Phi(U)$ whose tangent space $(\Phi(\tilde{P}))_{\Phi(\Theta(\lambda))}$ corresponding to the base point $\Phi(\Theta(\lambda))$ is spanned by the tangent vectors defined by

$$\Lambda_\alpha = (\Phi \circ \Theta)_* \left(\frac{\partial}{\partial\lambda_\alpha} \right) = \frac{\partial}{\partial\lambda_\alpha}(\Phi \circ \Theta) = \sum_{j=1}^{r} \left(p(\alpha)^j - p(0)^j\right)\Phi_j \tag{3.9}$$

$$\text{for } \alpha = 1, \dots, d.$$

It will prove convenient to introduce the coefficients of the Riemannian metric induced on this submanifold in terms of these coordinates by defining

$$g_{\alpha\beta} = \langle \Lambda_\alpha, \Lambda_\beta \rangle = \sum_{\eta,\gamma=1}^{r} \left(p(\alpha)^\eta - p(0)^\eta\right)\left(p(\beta)^\gamma - p(0)^\gamma\right)\bar{g}_{\eta\gamma}$$

$$\text{for } \alpha, \beta = 1, \dots, d.$$

For each $x \in U$, the vectors $\Phi_\alpha \in (S^m)_{\Phi(x)}$ for $\alpha = 1, \dots, r$, which are viewed below as column $m + 1$ vectors, are linearly independent and the $(m + 1) \times r$ matrix-valued mapping $x \to (\Phi_1, \dots, \Phi_r)$ is C^∞. Since U is contractible, there exist orthonormal vectors $N(\alpha) \in (R^{m+1})_{\Phi(x)}$ for $\alpha = 0, \dots, m - r$ and $x \in U$, such that the following properties hold.

(N1) $N(0)$ is identified with $\Phi(x) \in S^m$.
(N2) $N(\alpha) \perp \Phi_\beta$ for all $\alpha = 0, \dots, m - r$ and $\beta = 1, \dots, r$.
(N3) $N(\alpha)$ is C^∞ function of the base point x for $\alpha = 0, \dots, m - r$.

For a proof see Atiyah (1967), Lemma 1.4.4(2) and the discussion in Steenrod (1951), Section 6.7. The $N(\alpha)$ are introduced in order to give a convenient description of $D^{**}(P, K, \theta)$. As in Weyl (1939), the volume turns out not to depend on these vector fields. Note that $N(\alpha)$, $\alpha = 1, \dots, m - r$, span the orthogonal complement of $\Phi(U)_{\Phi(x)}$ relative to $(S^m)_{\Phi(x)}$, while the $N(\alpha)$, $\alpha = 0, \dots, m - r$, span the orthogonal complement of $\Phi(U)_{\Phi(x)}$ relative to $(R^{m+1})_{\Phi(x)}$.

Define vector fields *along* $\Phi(\tilde{P})$:

$$\Psi(\kappa(i)) = \sum_{\beta=1}^{r} \sum_{\alpha=1}^{r} \bar{g}^{\alpha\beta}\kappa(i)^\alpha \Phi_\beta, \qquad i = 1, \dots, r - d.$$

Since the $\kappa(i)$ are normal to the tangent space of $\tilde{P}$, it follows from the adjoint property of Ψ that each $\Psi(\kappa(i))$ is normal to the tangent space of $\Phi(\tilde{P})$.

Define the matrix of inner products $\Sigma = (\sigma_{ij})$ by letting

$$\sigma_{ij} = \langle \Psi(\kappa(i)), \Psi(\kappa(j)) \rangle = \sum_{\alpha, \beta = 1}^{r} \kappa(i)^{\alpha} \kappa(j)^{\beta} \bar{g}^{\alpha\beta}, \qquad i, j = 1, \dots, r-d.$$

This matrix is a positive definite matrix since $\bar{g}^{\alpha\beta}$ is positive definite and the $\kappa(i)$ are linearly independent. Define the associated quadratic norm

$$\xi(t) = \left\{ \sum_{i,j=1}^{r-d} \sigma_{ij} t^i t^j \right\}^{1/2}, \tag{3.10}$$

for $t = (t^1, \dots, t^{r-d}) \in R^{r-d}$ which is the length of the linear combination $\sum_{i=1}^{r-d} t^i \Psi(\kappa(i))$. The form ξ can be used to define the ball of *radius* $a > 0$ in the positive orthant

$$\Xi(a) = \{ t \in [0, +\infty)^{r-d} : \xi(t) \le a \}.$$

Note the dependence of $\Psi(\kappa(i))$, σ_{ij}, ξ and $\Xi(a)$ on the base point $\lambda \in \tilde{P}$.

LEMMA 3.3. *Define*

$$E(\theta) = \{ (\lambda, t, u) \in \mathscr{S}_d \times \Xi(\tan(\theta)) \times R^{m-r} : \|u\|^2 \le \tan^2(\theta) - \xi^2(t) \}$$

and Ω: $E(\theta) \to R^{m+1}$ *by*

$$\Omega(\lambda, t, u) = \Phi(\Theta(\lambda)) + \sum_{i=1}^{r-d} t^i \Psi(\kappa(i)) + \sum_{i=1}^{m-r} u^i N(i),$$

where $\Psi(\kappa(i))$ *and* $N(i)$ *have* $\Phi(\Theta(\lambda))$ *as base point. Then*

$$D^{**}(P, K, \theta) = \{ [\Omega(\lambda, t, u)] : (\lambda, t, u) \in E(\theta) \}.$$

PROOF.

$$\operatorname{con}(\Psi(K)) \oplus \Phi_*(R^r)^{\perp} = \left\{ \sum_{i=1}^{r-d} t^i \Psi(\kappa(i)) + \sum_{i=1}^{m-r} u^i N(i) : t^i \ge 0,\ u^i \in R \right\},$$

so the result follows from (3.5) and the fact that

$$\left\| \sum_{i=1}^{r-d} t^i \Psi(\kappa(i)) + \sum_{i=1}^{m-r} u^i N(i) \right\|^2 = \xi^2(t) + \|u\|^2. \qquad \square$$

The $\Psi(\kappa(i))$ define linearly independent vector fields along $\Phi(\tilde{P})$, tangent to $\Phi(U)$ and normal to $\Phi(\tilde{P})$, while the $N(i)$ are orthonormal vector fields along $\Phi(\tilde{P})$ and normal to $\Phi(U)$. Each point in $D^{**}(P, K, \theta)$ is thus represented in Lemma 3.3 as a sum of a point in $\Phi(\tilde{P})$ and a vector normal to $\Phi(\tilde{P})$, while the normal vector itself decomposes as a sum of a vector tangent to $\Phi(U)$ and a vector normal to $\Phi(U)$.

The first step in finding the volume of $D^{**}(P, K, \theta)$ is to proceed as in Weyl (1939) and express the volume as an integral whose integrand is the determinant of a matrix of partial derivatives. Then express this determinant in terms of the second fundamental form for the submanifold $\Phi(\tilde{P})$. This leads to the expression in Theorem 3.3. Define the coefficients of the second fundamental form (with raised index):

$$(3.11)\qquad G_\alpha^\beta(i) = \sum_{j=1}^{d} g^{j\beta}\left\langle \Lambda_\alpha, \frac{\partial N(i)}{\partial \lambda^j} \right\rangle = -\sum_{j=1}^{d} g^{j\beta}\left\langle \frac{\partial \Lambda_\alpha}{\partial \lambda^j}, N(i) \right\rangle$$

for $\alpha, \beta = 1, \dots, d$ and $i = 0, \dots, m - r$. A key property of these coefficients used by Weyl (1939) to simplify the determinant in the integrand, and which will be utilized in a similar manner below, is the fact that

$$(3.12)\qquad \Pi\left(\frac{\partial}{\partial \lambda_\alpha} N(i) \right) = \sum_{\beta=1}^{d} G_\alpha^\beta(i) \Lambda_\beta,$$

where $\Pi(v)$ denotes the orthogonal projection of the tangent vector v onto $\operatorname{span}\{\Lambda_\beta, \beta = 1, \dots, d\}$, the tangent space of $\Phi(\tilde{P})$.

The following similar observation will be used.

LEMMA 3.4. *Let $H_\alpha^\beta(i)$ be the coefficients with the property that*

$$\Pi\left(\frac{\partial}{\partial \lambda_\alpha} \Psi(\kappa(i)) \right) = \sum_{\beta=1}^{d} H_\alpha^\beta(i) \Lambda_\beta, \qquad \alpha = 1, \dots, d, i = 1, \dots, r - d.$$

Then

$$H_\alpha^\beta(i) = -\sum_{u=1}^{r} \kappa(i)^u \sum_{w,\gamma=1}^{r} \bar{g}^{uw} \left\langle \Phi_w, \frac{\partial \Phi_\gamma}{\partial \lambda_\alpha} \right\rangle \sum_{\eta=1}^{d} g^{\eta\beta}(p(\eta)^\gamma - p(0)^\gamma).$$

PROOF. See Appendix C. □

Let $H(i)$ [resp. $G(i)$] denote the $d \times d$ matrix whose α, β entry is $H_\alpha^\beta(i)$ [resp. $G_\alpha^\beta(i)$] for $i = 1, \dots, r - d$ (resp. $i = 0, \dots, m - r$). Note the dependence of these matrices on λ. The proof of Theorem 3.3 is rather long and is given in Appendix C.

THEOREM 3.3. *The expression*

$$\left| \int_{\lambda \in \mathscr{A}_d} \sqrt{\det g_{\alpha\beta}} \int_{t \in \Xi(\tan(\theta))} \sqrt{\det \sigma_{ij}}\, h\left(\lambda, t, \sqrt{\tan^2(\theta) - \xi^2(t)}\right) dt\, d\lambda \right|,$$

where

$$h(\lambda, t, \Delta) \equiv \int_{\{u \in R^{m-r}: \|u\| \le \Delta\}} \frac{\det A(\lambda, t, u)}{\{1 + \xi^2(t) + \|u\|^2\}^{(m+1)/2}}\, du$$

and where

$$A(\lambda, t, u) \equiv I + \sum_{i=1}^{r-d} t^i H(i) + \sum_{i=1}^{m-r} u^i G(i)$$

is an upper bound for the m-dimensional volume of $D^{**}(P, K, \theta)$, *provided the matrix* $A(\lambda, t, u)$ *is nonsingular for* $(\lambda, t, u) \in E(\theta)$, *and gives the exact volume if the mapping* $(\lambda, t, u) \to [\Omega(\lambda, t, u)]$ *is one-to-one.*

In Section 4 it is shown that $A(\lambda, t, u)$ is nonsingular for all $(\lambda, t, u) \in E(\theta)$, if θ is sufficiently small. The remainder of this section focusses on evaluation of $h(\lambda, t, \Delta)$ for given $\lambda \in \mathscr{S}_d$, $t \in [0, +\infty)^{r-d}$ and $\Delta \geq 0$. At first glance, this calculation seems to be a cumbersome chore. However, this calculation simplifies quite nicely because $h(\lambda, t, \Delta)$ can be expressed, using a determinantal identity and a bit of calculus, as a sum of spherical integrals which are related by a formula in Weyl (1939) to the coefficients of the metric $\bar{g}_{\alpha\beta}$.

In order to describe the identity for the determinant of a sum of matrices used below, it is necessary to introduce more notation. The set of permutations of a given set X will be denoted by S_X and the sign of $\tau \in S_X$ will be denoted by $\operatorname{sgn}(\tau)$. Let $I(d, s)$ denote the collection of subsets of $\{1, \dots, d\}$ consisting of exactly s elements. For a given $d \times d$ matrix B, the following notation is used for its square submatrices. Let $B\binom{J}{L}$ denote the $s \times s$ submatrix of B obtained by using as row indices the elements of J and as column indices the elements of L for $J, L \in I(d, s)$. If the ordered elements of J are $j_1, \dots, j_s$, let $j_{s+1}, \dots, j_d$ denote the ordered elements of $J^c = \{1, \dots, d\} - J$ and define $\tau_J \in S_{\{1, \dots, d\}}$ by $i = j_{\tau_J(i)}$ for $i = 1, \dots, d$. Similarly, denote the ordered elements of L by $l_1, \dots, l_s$ and those of L^c by $l_{s+1}, \dots, l_d$. Also, define $\tau_L \in S_{\{1, \dots, d\}}$ by $i = l_{\tau_L(i)}$ for $i = 1, \dots, d$.

The following result is well known and is used implicitly in Weyl (1939) for the case when B is the identity matrix. The proof involves a fairly elementary argument and will be omitted.

LEMMA 3.5. *If B and C are $d \times d$ matrices, then*

$$\det(B + C) = \sum_{s=0}^{d} \sum_{J, L \in I(d, s)} \operatorname{sgn}(\tau_J)\operatorname{sgn}(\tau_L)\det\left(B\binom{J}{L}\right)\det\left(C\binom{J^c}{L^c}\right),$$

where the determinant of a 0×0 *matrix is defined to be* 1.

Applying Lemma 3.5 to the matrix $A(\lambda, t, u)$ yields

$$\det A(\lambda, t, u) = \sum_{s=0}^{d} \sum_{J, L \in I(d, s)} \phi(J, L)\det\left(\sum_{i=1}^{m-r} u^i G(i)\binom{J^c}{L^c}\right),$$

where

$$\phi(J, L) = \operatorname{sgn}(\tau_J)\operatorname{sgn}(\tau_L)\det\left(I\binom{J}{L} + \sum_{i=1}^{r-d} t^i H(i)\binom{J}{L}\right).$$

Since the coefficients $\phi(J, L)$ do not depend on the variables u, it follows that

$$h(\lambda, t, \Delta) = \sum_{s=0}^{d} \sum_{J, L \in I(d,s)} \phi(J, L) \int_{\{u \in R^{m-r}:\, \|u\| \le \Delta\}} \frac{\det\left(\sum_{i=1}^{m-r} u^i G(i)\binom{J^c}{L^c}\right)}{\{1 + \xi^2(t) + \|u\|^2\}^{(m+1)/2}}\, du.$$

The integrals in this last expression are evaluated in the following lemma. The proof involves changing to spherical coordinates and using elementary calculus. Fix $0 \le s \le d$ and $J, L \in I(d, s)$. Let μ denote the uniform measure on the unit sphere and let V_k denote the volume of the k-dimensional unit ball so that $V_0 = 1$, $V_1 = 2$ and the V_k satisfy the recursion formula $V_{k+2} = (2\pi/(k+2))V_k$. Let $q = d - s + m - r - 1$ and let $\mathscr{B}$ denote the incomplete beta function with parameters $(m - q)/2$ and $(q + 1)/2$, i.e.,

$$\mathscr{B}(y) = \int_{x=0}^{y} x^{(m-q)/2-1}(1 - x)^{(q+1)/2-1}\, dx.$$

LEMMA 3.6.

$$\int_{\{u \in R^{m-r}:\, \|u\| \le \Delta\}} \frac{\det\left(\sum_{i=1}^{m-r} u^i G(i)\binom{J^c}{L^c}\right)}{\{1 + \xi^2(t) + \|u\|^2\}^{(m+1)/2}}\, du$$

$$= \frac{1}{2}(m - r)V_{m-r}\{1 + \xi^2(t)\}^{(q-m)/2}\left[\mathscr{B}(1) - \mathscr{B}\left(\frac{1 + \xi^2(t)}{1 + \xi^2(t) + \Delta^2}\right)\right]$$

$$\times \int_{\{u \in R^{m-r}:\, \|u\| = 1\}} \det\left(\sum_{i=1}^{m-r} u^i G(i)\binom{J^c}{L^c}\right)\mu(du).$$

PROOF. See Appendix C. □

As the final step in the derivation of a tractable expression for $h(\lambda, t, \Delta)$ note that Weyl (1939) gives a formula for each of the spherical averages in Lemma 3.6 in terms of the metric defined by the embedding (the $g_{\alpha\beta}$'s).

4. Critical radii for tubular neighborhoods. In this section, conditions are given which guarantee that the matrix $A(\lambda, t, u)$ in Theorem 3.3 is nonsingular for $(\lambda, t, u) \in E(\theta)$, so that to use Hotelling's (1939) terminology, "local self-overlapping" of the tubular neighborhood of angular radius θ fails to occur and the volume formula of Section 3.5 is an upper bound. The much more difficult problem of giving conditions guaranteeing that "global self-overlapping" of the tubular neighborhoods fails to occur (so that the volume formula is exact) is not addressed here. In the one-dimensional case, when $\Phi(\tilde{P})$ is a curve, Hotelling (1939) gives a critical radius based on the minimum radius of curvature over all points of the curve. The situation for higher-dimensional sets Γ is considerably more complicated because a different radius of curvature is defined in each direction in the embedded manifold.

In the following discussion P and K are fixed as in Section 3.5 and the matrices $A(\lambda, t, u)$, $H(i)$ and $G(i)$ are as defined there. Recall that $G(i)$, $H(i)$ and σ_{ij} depend on λ. Let $\Sigma \equiv (\sigma_{ij})$. For an arbitrary $d \times d$ matrix M, let $\operatorname{eig}(M)$ denote the set of (possibly complex) eigenvalues of M. Let $\rho(M)$ denote the spectral radius of M, that is, $\rho(M) = \max\{|\lambda|: \lambda \text{ an eigenvalue of } M\}$. The l_2 norm of M is defined by $\|M\| = \sqrt{\sum_{i,j=1}^{d} M_{ij}^2}$.

Since $I + M$ is singular if and only if $-1 \in \operatorname{eig}(M)$, it follows easily that the matrix $A(\lambda, t, u)$ is nonsingular for $(\lambda, t, u) \in E(\theta)$ for $\theta < \zeta$ and is singular for some $(\lambda, t, u) \in E(\theta)$ for $\theta > \zeta$, where

$$(4.1)\quad \zeta = \inf\left\{\theta > 0: -1 \in \bigcup_{(\lambda, t, u) \in E(\theta)} \operatorname{eig}\left(\sum_{i=1}^{r-d} t^i H(i) + \sum_{j=1}^{m-r} u^j G(j)\right)\right\},$$

where the infimum is defined to be $+\infty$ if the set is empty.

It follows from Theorem 4.1 that $\zeta > 0$. While ζ gives the sharpest information one would want in order to determine when the volume expression may be suspect, it is quite difficult to calculate in general. Indeed, to find the matrices $G(i)$ is a straightforward but computationally intensive problem which requires determination of the $N(i)$. The size of this problem grows with m, the number of dimensions of the ambient space (hence with the number of observations in the statistical model). To complicate matters further, calculation of ζ requires determining for a given θ whether -1 is an eigenvalue of one of the matrices $\sum_{i=1}^{r-d} t^i H(i) + \sum_{j=1}^{m-r} u^j G(j)$ for some $(\lambda, t, u) \in E(\theta)$.

On the other hand, for practical purposes there is reason to suspect that in a sense ζ is too conservative because the volume formula ought to give reasonable approximations even when -1 *is* an eigenvalue for the matrix in a small set of points. It is an open problem to find a more appropriate measure of when the formula gives a good approximation or to find an estimate of the error for the volume formula when $\theta > \zeta$ fails.

Theorem 4.1 gives a conservative and easy to calculate bound for ζ. The following notation will prove useful. The unique $(r - d) \times (r - d)$ symmetric square root of Σ will be denoted by $\Upsilon = (v_{ij})$ so that $\|\Upsilon t\| = \xi(t)$ for $t \in R^{r-d}$. Define $d \times d$ matrices

$$\tilde{H}(i) = \sum_{j=1}^{r-d} v^{ij} H(j) \quad \text{for } i = 1, \dots, r - d,$$

where v^{ij} denotes the i, j element of Υ^{-1}.

THEOREM 4.1. *The inequalities $\zeta \geq \zeta^* > 0$ hold where ζ^* is defined by*

$$\tan^2(\zeta^*) = 1 / \sup_{\lambda \in \mathscr{S}_d} \max\left\{\sum_{i=1}^{r-d} \|\tilde{H}(i)\|^2, \sum_{i=1}^{m-r} \|G(i)\|^2\right\}.$$

PROOF. See Appendix D. $\square$

A simple computational formula for $\sum_{i=1}^{m-r} \|G(i)\|^2$ is given in the following. Thus, the dependence of ζ^* on the second fundamental form does not hinder one from calculating it. Denote the second derivative vectors of the embedding map $\Phi \circ \Theta$ by

$$(4.2) \qquad \Lambda_{\alpha\beta} = \frac{\partial^2}{\partial\lambda_\alpha \partial\lambda_\beta}\Phi \circ \Theta = \frac{\partial}{\partial\lambda_\alpha}\Lambda_\beta, \quad \text{for } \alpha, \beta = 1, \dots, d.$$

These of course depend on the choice of base point $\lambda \in \mathscr{S}_d$.

LEMMA 4.1. *The identity*

$$\sum_{i=1}^{m-r} \|G(i)\|^2 = \sum_{\alpha,\beta=1}^{d} \left\{ \left\| \sum_{\eta=1}^{d} g^{\beta\eta}\Lambda_{\alpha\eta} \right\|^2 - \left\| \sum_{u,w=1}^{r} \bar{g}^{uw} \left\langle \sum_{\eta=1}^{d} g^{\beta\eta}\Lambda_{\alpha\eta}, \Phi_u \right\rangle \Phi_w \right\|^2 \right\} - d$$

holds.

PROOF. See Appendix D. □

Acknowledgments. It would not have been possible to write this paper without having first read Weyl (1939) very carefully. Soren Johansen deserves thanks for giving me the reference. It was necessary to learn some differential geometry before penetrating the paper and I am indebted to Steven Zucker and Joseph Sampson for taking the time to explain various concepts. Alan Goldman's encyclopedic knowledge of convex geometry proved to be a great resource and Roger Horn deserves thanks for telling me about matrix norms. I also wish to thank David Siegmund for discussions, especially in explaining the applications described in Hotelling (1939). Henry Wynn, Iain Johnstone and the referees made comments which resulted in many improvements in the paper.

APPENDIX A

Some convex geometry. The following are standard definitions in convex geometry which are used throughout Section 3.1. Fix V, a finite-dimensional vector space over R endowed with a real-valued inner product denoted by $\langle\ ,\ \rangle$.

For $F \subseteq V$ the *affine hull* of F is the affine subspace given by

$$\operatorname{aff}(F) = \left\{ \sum_{i=1}^{q} \lambda_i v_i \colon \lambda_i \in R, v_i \in F, i = 1, \dots, q \text{ and } \sum_{i=1}^{q} \lambda_i = 1 \right\}.$$

The dimension of F is defined to be the dimension of $\operatorname{aff}(F)$. The *convex hull* of F is the convex set

$$\operatorname{conv}(F) = \left\{ \sum_{i=1}^{q} \lambda_i v_i : \lambda_i \geq 0,\ v_i \in F,\ i = 1, \dots, q \text{ and } \sum_{i=1}^{q} \lambda_i = 1 \right\}.$$

The *conical hull* of F is the set

$$\operatorname{con}(F) = \left\{ \sum_{i=1}^{q} \lambda_i v_i : \lambda_i \geq 0,\ v_i \in F, \text{ for } i = 1, \dots, q \right\}.$$

A *convex polyhedron* or *polyhedron* in V is a set of the form

$$\text{(A.1)} \qquad Q = \{u \in V : \langle v_i, u \rangle \leq c_i,\ i = 1, \dots, q\},$$

where $v_i \in V - \{0\}$ and $c_i \in R$ for $i = 1, \dots, q$. A point $u \in Q$ is said to be *inner* if $\langle v_i, u \rangle < c_i$ whenever the hyperplane $\{\langle v_i, u \rangle = c_i\}$ does not contain Q; otherwise the point is referred to as a boundary point. A convex polyhedron is said to be *pointed* if it contains no affine subspaces of V. A bounded convex polyhedron is referred to as a *polytope*.

If Q is a convex polyhedron given by (A.1), then a *face* of Q is a set of the form

$$G_I = \{u \in V : \langle v_i, u \rangle = c_i,\ \forall\, i \in I \text{ and } \langle v_i, u \rangle \leq c_i,\ \forall\, i \notin I\}$$

for some (possibly empty) set of indices $I \subseteq \{1, \dots, q\}$. The face is *proper* if it is not Q itself. A *facet* of a polyhedron Q is a proper face G of Q which is maximal, i.e., contained in no other proper face of Q. An *open face* of Q is a set of the form

$$\text{(A.2)} \qquad F_I = \{u \in V : \langle v_i, u \rangle = c_i,\ \forall\, i \in I \text{ and } \langle v_i, u \rangle < c_i,\ \forall\, i \notin I\},$$

for some (possibly empty) set of indices $I \subseteq \{1, \dots, q\}$.

Note that there are faces corresponding to each of the 2^p subsets of indices, though some of these faces may be empty, and faces corresponding to different pairs of sets may coincide. As an immediate consequence of the definitions we can express any convex polyhedron as a disjoint union of its open faces of various dimension. Thus,

$$\text{(A.3)} \qquad Q = \bigcup_{I \subseteq \{1, \dots, q\}} F_I,$$

for Q given in (A.1) and F_I defined in (A.2).

A *convex cone* in V is a subset of V which is closed under the formation of nonnegative linear combinations. A convex cone is called *polyhedral* if it is also a polyhedron. If C is convex cone in V, its *polar* (or *dual*) *convex cone* is defined by

$$C^p = \{u \in V : \langle v, u \rangle \leq 0,\ \forall\, v \in C\}.$$

If $x \in Q$, where a Q is a convex polyhedron, the *cone of feasible directions* at x relative to Q is the polyhedral convex cone defined by

$$C_{x,Q} = \{u \in V : x + \varepsilon u \in Q, \text{ for all sufficiently small } \varepsilon > 0\}.$$

Loosely speaking, $C_{x,Q}$ is the set of vectors which "point into" the polyhedron from x. In fact, the representation (A.1) leads to a simple description of $C_{x,Q}$ given in Proposition A.3(c).

The following basic results, whose proofs appear in Stoer and Witzgall (S–W) (1970), will be used below.

A.I (S–W Lemma 2.3.10). Every nonempty polyhedron in V has inner points.

A.II (S–W Theorem 2.4.7; see also 2.4.4). If G is a face of a polyhedron $Q \subseteq V$, then $\operatorname{aff}(G) \cap Q = G$.

A.III (S–W Section 2.8). If C is a polyhedral convex cone in V, then:

(a) There exists a finite set $S \subseteq V$ such that $C = \{u \in V: \langle v, u \rangle \leq 0, \forall\, v \in S\}$, in which case $C^p = \operatorname{con}(S)$.

(b) There exists a finite set S' such that $C = \operatorname{con}(S')$, in which case

$$C^p = \{u \in V: \langle v, u \rangle \leq 0, \forall\, v \in S'\}.$$

(c) $C^{pp} = C$.

A.IV (S–W Theorem 2.14.3). If G is a facet of a polyhedron $Q \subseteq V$, then $\dim(G) = \dim(Q) - 1$.

A.V (S–W Theorem 2.12.2). Every convex polytope in V is the convex hull of a finite set.

Proposition A.1. *Let $Q \subseteq V$ be a d-dimensional convex polytope. Then there exists T, a collection of subsets of Q with each subset containing exactly d points, such that for any inner point $x \in Q$:*

(a) $\operatorname{conv}(S \cup \{x\})$ *is d dimensional, for every $S \in T$.*

(b) $\dim(\operatorname{conv}(S \cup \{x\}) \cap \operatorname{conv}(S' \cup \{x\})) \leq d - 1$, *for all distinct* $S, S' \in T$.

(c) $\bigcup_{S \in T} \operatorname{conv}(S \cup \{x\}) = Q$

Proof. The proof is by induction on d. For $d = 1$ the result is obvious. Now assume the result holds for the $(d-1)$-dimensional case and let Q be a d-dimensional convex polytope. Let $G_1, \dots, G_m$ be the $(d-1)$-dimensional faces of Q and fix x_i an inner point of G_i, for $i = 1, \dots, m$, whose existence is guaranteed by A.I. By the inductive hypothesis, there exist $T_i \subseteq G_i$ for $i = 1, \dots, m$ such that (a′), (b′) and (c′) hold, where (a′), (b′) and (c′) are obtained by replacing d by $d - 1$, Q by G_i, x by x_i and T by T_i in (a), (b) and (c).

Let $T = \{S \cup \{x_i\}: S \in T_i$, for some $i = 1, \dots, m\}$. Since $x \notin \operatorname{aff}(S \cup \{x_i\})$ for all $S \in T_i$ and $i = 1, \dots, r$, (a) follows from (a′). Note that (b) is immediate from (b′) if S and S' are in the same T_i. If $S \in T_i$ and $S' \in T_j$ for some $i \neq j$, then since $S \cup \{x_i\} \subseteq G_i$ and $S \cup \{x_j\} \subseteq G_j$, it suffices to show

$\dim(G_i \cap G_j) \le d-2$. But if $\dim(G_i \cap G_j) \ge d-1 = \dim(G_i) = \dim(G_j)$, then $\mathrm{aff}(G_i \cap G_j) = \mathrm{aff}(G_i) = \mathrm{aff}(G_j)$, so by A.II, $G_i \cap G_j = G_i = G_j$, which is a contradiction. Finally, to prove (c), clearly $\bigcup_{S \in T} \mathrm{conv}(S \cup \{x\}) \subseteq Q$. If $u \in Q$, there exists a boundary point v of Q with $u = \eta x + (1-\eta)v$, for some $\eta \in [0,1]$. It follows that v lies in a proper face of Q, which in turn must be contained in some facet (maximal proper face) G. By A.IV, $G = G_i$ for some $i = 1, \dots, m$ and by (c′), $v \in \mathrm{conv}(S \cup \{x_i\})$ for some $S \in T_i$. It then follows that $u \in \mathrm{conv}(S \cup \{x_i, x\})$ and the proof is complete. □

PROPOSITION A.2. *Let C be a d-dimensional pointed polyhedral convex cone in V. Then there exists T, a collection of subsets of C, with each subset consisting of exactly d points, such that*:

(a) $\mathrm{con}(S)$ *is d-dimensional, for all $S \in T$.*
(b) $\dim(\mathrm{con}(S) \cap \mathrm{con}(S')) < d$, *for all distinct $S, S' \in T$.*
(c) $C = \bigcup_{S \in T} \mathrm{con}(S)$.

PROOF. Using AIII(c) $\mathrm{aff}(C^p)^{\perp} = \{0\}$, since C is pointed and $\mathrm{aff}(C^p)^{\perp} \subseteq C$. Thus, $\dim(\mathrm{aff}(C^p)) = \dim(V)$ and C^p has a nonempty interior. Fix $v \in (C^p)^{\mathrm{int}}$. Then clearly $\langle v, u \rangle < 0$, for all $u \in C$. If $Q = \{u \in C \colon \langle v, u \rangle = -1\}$, it follows easily that $C = \{\lambda u \colon u \in Q,\ \lambda \ge 0\}$. By A.III(b), C is the conical hull of some finite set $\{v_j,\ j = 1, \dots, m\}$, where $v_j \ne 0$, for $j = 1, \dots, m$. It follows that Q is the polytope given by the convex hull of $\{-\langle v, v_j \rangle^{-1} v_j,\ j = 1, \dots, m\}$, $\mathrm{con}(Q) = C$ and Q is $(d-1)$-dimensional. If T is a collection of subsets of Q given by Proposition A.1 and x is any inner point of Q, it follows easily that $\bigcup_{S \in T} S \cup \{x\}$ has the desired properties. □

PROPOSITION A.3. *Let $Q \subset V$ be a convex polytope of the form* (A.1). *If $x \in F_I$, where F_I is the open face given in* (A.2) *for some $I \subseteq \{1, \dots, q\}$, then*:

(a) $C_{x,Q} = \{u \in V \colon \langle v_i, u \rangle \le 0,\ \forall\, i \in I\}$.
(b) $C_{x,Q}$ *has the same dimension as* Q.
(c) $C^p_{x,Q} = \mathrm{con}\{v_i \colon i \in I\}$.
(d) *If* $W = \{u \in V \colon \langle v_i, u \rangle = 0,\ \forall\, i \in I\}$, *then* $\mathrm{aff}(F_I) = x + W$.
(e) $\dim(F_I) + \dim(C^p_{x,Q}) = \dim(V)$.
(f) *If* $\dim(Q) = \dim(V)$, *then* $C^p_{x,Q}$ *is pointed.*

PROOF. (a) follows immediately from the definitions. For (b) note that $Q \subset x + C_{x,Q}$, hence

$$\dim(Q) \le \dim(x + C_{x,Q}) = \dim(C_{x,Q}).$$

On the other hand, Q is the convex hull of a finite set by A.V, so there exists $\varepsilon > 0$ such that $x + \{u \in C_{x,Q} \colon \|u\| \le \varepsilon\} \subset Q$. Thus

$$\begin{aligned}\dim(Q) &\ge \dim\big(x + \{u \in C_{x,Q} \colon \|u\| \le \varepsilon\}\big)\\ &= \dim\{u \in C_{x,Q} \colon \|u\| \le \varepsilon\} = \dim(C_{x,Q}),\end{aligned}$$

where the equality uses the fact $C_{x,Q}$ is a cone, and the proof is complete. (c) follows immediately from A.III(a). To prove (d), it is easy to verify that $F_I \subseteq x + W$, which is an affine set, so $\operatorname{aff}(F_I) \subseteq x + W$. Define $\varepsilon = \min_{i \notin I}\{(c_i - \langle v_i, x\rangle)/\|v_i\|\} > 0$, if $I^c \neq \varnothing$ and $\varepsilon = +\infty$ if $I^c = \varnothing$. Then if $B = W \cap \{u \in V\colon \|u\| < \varepsilon\}$, a simple application of the Cauchy–Schwarz inequality shows $x + B \subseteq F_I$, so $\operatorname{aff}(x + W) = \operatorname{aff}(x + B) \subseteq \operatorname{aff}(F_I)$ and the proof is complete.

For (e), note that as an immediate consequence of (c),

$$\operatorname{aff}(C_{x,Q}^{p}) = \operatorname{aff}(\operatorname{con}\{v_i\colon i \in I\}),$$

which implies

$$\begin{aligned}\dim(C_{x,Q}^{p}) &= \dim(\operatorname{aff}(\operatorname{con}\{v_i\colon i \in I\})) = \dim(V) - \dim\big(\operatorname{aff}(\operatorname{con}\{v_i\colon i \in I\})^{\perp}\big)\\ &= \dim(V) - \dim(W),\end{aligned}$$

where W is defined in (d) and the last equality uses (d). For (f), suppose $C_{x,Q}^{p}$ contains some line, say $\{u + \lambda v\colon \lambda \in R\}$, where $u, v \in V$, with $v \neq 0$. Thus, $\langle u + \lambda v, w\rangle \leq 0$ for all $\lambda \in R$ and $w \in C_{x,Q}$. It follows easily that $\langle v, w\rangle = 0$ for all $w \in C_{x,Q}$, so $\dim(C_{x,Q}) \leq \dim(V) - 1$, which contradicts (b). □

APPENDIX B

Proof of Theorem 3.2. Using Lemma 3.2 and (P3′),

$$\Phi_*(C_{I,x})^{p} = \bigcup_{K \in \mathscr{K}_I} \operatorname{con}(\Psi(K)) \oplus \Phi_*(R^r)^{\perp}$$

and (3.6) follows from (P3″) and the fact that $F_I \subset \overline{F}_I$.

The second and third claims are proved using dimensionality arguments. Note that a subset of S^m has μ-measure zero if it is contained in a submanifold of dimension less than m. For the second claim,

$$\bigcup_{(P,K) \in \mathscr{P}_I \times \mathscr{K}_I} D^{**}(P, K, \theta) - D^{**}(F_I, \theta)$$

is contained in

$$\bigcup_{(P,K) \in \mathscr{P}_I \times \mathscr{K}_I} \{[\Phi(x) + Y] \colon x \in \operatorname{conv}(P) - F_I,\ Y \in \operatorname{con}(\Psi(K)) \oplus \Phi_*(R^r)^{\perp}\}. \tag{B.1}$$

$\operatorname{conv}(P) - F_I$ is contained in the boundary of $\overline{F}_I$ for every $P \in \mathscr{P}_I$. Therefore

$$\dim(\operatorname{conv}(P) - F_I) < \dim(F_I).$$

It follows that each set in the union in the right side of (B.1) is contained in a submanifold having dimension at most

$$\dim(F_I) - 1 + (r - \dim(F_I)) + (m - r) = m - 1.$$

To prove the last claim, fix distinct pairs (P', K'), $(P, K) \in \mathscr{P}_I \times \mathscr{K}_I$. If $P \neq P'$, then

$$D^{**}(P, K, \theta) \cap D^{**}(P', K', \theta) \subseteq \{[\Phi(x) + Y]: x \in \operatorname{conv}(P) \cap \operatorname{conv}(P'), Y \in \Psi(C_I^p) \oplus \Phi_*(R^r)^{\perp}\}$$

and since (P2″) gives $\dim(\operatorname{conv}(P') \cap \operatorname{conv}(P)) < \dim(F_I)$, the dimension of this set is less than

$$\dim(F_I) + (r - \dim(F_I)) + (m - r) = m.$$

If $P = P'$, then $K \neq K'$ and it follows that

$$D^{**}(P, K, \theta) \cap D^{**}(P', K', \theta)$$

is a subset of

$$\{[\Phi(x) + Y]: x \in P, Y \in (\operatorname{con}(\Psi(K)) \cap \operatorname{con}(\Psi(K'))) \oplus \Phi_*(R^r)^{\perp}\},$$

and since (P2′) gives $\dim(\operatorname{con}(\Psi(K)) \cap \operatorname{con}(\Psi(K'))) < r - \dim(F_I)$, this is a submanifold whose dimension is less than

$$\dim(P) + (r - \dim(F_I)) + (m - r) = m. \qquad \square$$

APPENDIX C

Proofs of results from Section 3.5.

PROOF OF LEMMA 3.4. Using (3.3) and the linearity of Π leads to

$$\begin{aligned}
\Pi\left[\frac{\partial}{\partial\lambda_\alpha}\Psi(\kappa(i))\right] &= \Pi\left[\frac{\partial}{\partial\lambda_\alpha}\left\{\sum_{u,v=1}^{r}\kappa(i)^u \bar{g}^{uv}\Phi_v\right\}\right]\\
&= \sum_{u=1}^{r}\kappa(i)^u\sum_{v=1}^{r}\Pi\left[\frac{\partial}{\partial\lambda_\alpha}\{\bar{g}^{uv}\Phi_v\}\right]\\
&= \sum_{u=1}^{r}\kappa(i)^u\sum_{v=1}^{r}\Pi\left[\frac{\partial\bar{g}^{uv}}{\partial\lambda_\alpha}\Phi_v + \bar{g}^{uv}\frac{\partial\Phi_v}{\partial\lambda_\alpha}\right]\\
&= \sum_{u=1}^{r}\kappa(i)^u\left(\left\{\sum_{v=1}^{r}\frac{\partial\bar{g}^{uv}}{\partial\lambda_\alpha}\Pi[\Phi_v]\right\} + \left\{\sum_{w=1}^{r}\bar{g}^{uw}\Pi\left[\frac{\partial\Phi_w}{\partial\lambda_\alpha}\right]\right\}\right).
\end{aligned}\tag{C.1}$$

From the analogue of Remark 3.2 for the submanifold under consideration and (3.9) one obtains

$$\begin{aligned}
\Pi(\Phi_v) &= \sum_{\eta,\beta=1}^{d} g^{\eta\beta}\langle\Phi_v, \Lambda_\eta\rangle\Lambda_\beta = \sum_{\eta,\beta=1}^{d} g^{\eta\beta}\left\langle\Phi_v, \sum_{j=1}^{r}\left(p(\eta)^j - p(0)^j\right)\Phi_j\right\rangle\Lambda_\beta\\
&= \sum_{\eta,\beta=1}^{d} g^{\eta\beta}\sum_{j=1}^{r}\left(p(\eta)^j - p(0)^j\right)\bar{g}_{vj}\Lambda_\beta.
\end{aligned}\tag{C.2}$$

Differentiating both sides of the equation $\Sigma_{w=1}^{r} \bar{g}^{uv}\bar{g}_{vw} = \delta_w^u$ and leads to the identity

$$\text{(C.3)}\qquad \begin{aligned} \frac{\partial \bar{g}^{uv}}{\partial \lambda_\alpha} &= -\sum_{w,\lambda=1}^{r} \bar{g}^{uw}\bar{g}^{v\gamma}\frac{\partial \bar{g}_{w\gamma}}{\partial \lambda_\alpha} \\ &= -\sum_{w,\gamma=1}^{r} \bar{g}^{uw}\bar{g}^{v\gamma}\left(\left\langle \frac{\partial \Phi_w}{\partial \lambda_\alpha}, \Phi_\gamma\right\rangle + \left\langle \Phi_w, \frac{\partial \Phi_\gamma}{\partial \lambda_\alpha}\right\rangle\right). \end{aligned}$$

Combining (C.2) and (C.3) and the identity $\Sigma_{v=1}^{r} \bar{g}^{v\gamma}\bar{g}_{vj} = \delta_j^\gamma$ gives the following expression for one of the terms in (C.1):

$$\text{(C.4)}\qquad \begin{aligned} \sum_{v=1}^{r} \frac{\partial \bar{g}^{uv}}{\partial \lambda_\alpha}\Pi[\Phi_v] &= -\sum_{v=1}^{r}\sum_{w,\gamma=1}^{r} \bar{g}^{uw}\bar{g}^{v\gamma}\left(\left\langle \frac{\partial \Phi_w}{\partial \lambda_\alpha}, \Phi_\gamma\right\rangle + \left\langle \Phi_w, \frac{\partial \Phi_\gamma}{\partial \lambda_\alpha}\right\rangle\right) \\ &\qquad \times \sum_{\eta,\beta=1}^{d} g^{\eta\beta}\sum_{j=1}^{r}\left(p(\eta)^j - p(0)^j\right)\bar{g}_{vj}\Lambda_\beta \\ &= -\sum_{w,\gamma=1}^{r} \bar{g}^{uw}\left(\left\langle \frac{\partial \Phi_w}{\partial \lambda_\alpha}, \Phi_\gamma\right\rangle + \left\langle \Phi_w, \frac{\partial \Phi_\gamma}{\partial \lambda_\alpha}\right\rangle\right) \\ &\qquad \times \sum_{\eta,\beta=1}^{d} g^{\eta\beta}\left(p(\eta)^\gamma - p(0)^\gamma\right)\Lambda_\beta. \end{aligned}$$

For the other term in (C.1), again using Remark 3.2 and (3.9):

$$\text{(C.5)}\qquad \begin{aligned} \sum_{w=1}^{r} \bar{g}^{uw}\Pi\left[\frac{\partial \Phi_w}{\partial \lambda_\alpha}\right] &= \sum_{w=1}^{r} \bar{g}^{uw}\sum_{\eta,\beta=1}^{d} g^{\eta\beta}\left\langle \frac{\partial \Phi_w}{\partial \lambda_\alpha}, \Lambda_\eta\right\rangle\Lambda_\beta \\ &= \sum_{w=1}^{r} \bar{g}^{uw}\sum_{\eta,\beta=1}^{d} g^{\eta\beta} \\ &\qquad \times \sum_{\gamma=1}^{r}\left(p(\eta)^\gamma - p(0)^\gamma\right)\left\langle \frac{\partial \Phi_w}{\partial \lambda_\alpha}, \Phi_\gamma\right\rangle\Lambda_\beta. \end{aligned}$$

Note that this is one of the terms appearing in (C.4). The proof is completed by summing (C.4) and (C.5), cancelling terms and substituting into (C.1). □

Proof of Theorem 3.3. Using Lemma 3.3 and expression (5) from Weyl (1939), the expression

$$\int_{\mathscr{S}_d}\int_{t\in\Xi(\tan(\theta))}\int_{\|u\|^2\le\tan^2(\theta)-\xi^2(t)} |\det M|/\|\Omega\|^{m+1}\,du\,dt\,d\lambda,$$

where

$$M = \left(\Omega, \frac{\partial\Omega}{\partial\lambda^1}, \dots, \frac{\partial\Omega}{\partial\lambda^d}, \frac{\partial\Omega}{\partial t^1}, \dots, \frac{\partial\Omega}{\partial t^{r-d}}, \frac{\partial\Omega}{\partial u^1}, \dots, \frac{\partial\Omega}{\partial u^{m-r}}\right),$$

is an upper bound for the volume of $D^{**}(P, K, \theta)$ if M is nonsingular in $E(\theta)$ and this expression gives the exact volume if the mapping from $E(\theta)$ into

$D^{**}(P, K, \theta)$ taking $(\lambda, t, u) \to [\Omega(\lambda, t, u)]$ is one-to-one. When M is nonsingular for (λ, t, u) in $E(\theta)$, the determinant does not change sign so the absolute value can be moved outside the integral.

Clearly $\partial\Omega/\partial t^i = \Psi(\kappa(i))$ and $\partial\Omega/\partial u^i = N(i)$. Since these vectors appear as columns in M, the determinant is unchanged if multiples of these vectors are subtracted from the first column. Thus, the first column can be replaced by $\Phi(\Theta(\lambda))$ [which is identified with $N(0)$ at the base point $\Phi(\Theta(\lambda))$] and

$$\det M = \det\left(N(0), \frac{\partial\Omega}{\partial\lambda^1}, \dots, \frac{\partial\Omega}{\partial\lambda^d}, \Psi(\kappa(1)), \dots, \Psi(\kappa(r-d)),\right. \tag{C.6}$$
$$\left. N(1), \dots, N(m-r)\right).$$

The vectors $\Psi(\kappa(i))$, $i = 1, \dots, r-d$, and $N(i)$, $i = 0, \dots, m-r$, are linearly independent and lie in $(\Phi(\tilde{P}))^{\perp}_{\Phi(\Theta(\lambda))}$, the [$(m+1-d)$-dimensional] orthogonal complement of the (d-dimensional) tangent space $(\Phi(\tilde{P}))_{\Phi(\Theta(\lambda))}$ relative to the $(m+1)$-dimensional tangent space $(R^{m+1}_{\Phi(\Theta(\lambda))})$. It follows that the vectors span the orthogonal complement of this space and det M is unchanged if any of the columns $\partial\Omega/\partial\lambda^\alpha$ is replaced by $\Pi(\partial\Omega/\partial\lambda^\alpha)$, its projection onto the tangent space $(\Phi(\tilde{P}))_{\Phi(\Theta(\lambda))}$.

Using the definition of Ω, (3.12) and Lemma 3.4 leads to

$$\begin{aligned}
\Pi\left(\frac{\partial\Omega}{\partial\lambda^\alpha}\right) &= \Pi\left(\Lambda_\alpha + \sum_{i=1}^{r-d} t^i \frac{\partial\Psi(\kappa(i))}{\partial\lambda_\alpha} + \sum_{i=1}^{m-r} u^i \frac{\partial N(i)}{\partial\lambda_\alpha}\right) \\
&= \Lambda_\alpha + \sum_{i=1}^{r-d} t^i \sum_{\beta=1}^{d} H^\beta_\alpha(i)\Lambda_\beta + \sum_{i=1}^{m-r} u^i \sum_{\beta=1}^{d} G^\beta_\alpha(i)\Lambda_\beta \\
&= \sum_{\beta=1}^{d}\left(\delta^\beta_\alpha + \sum_{i=1}^{r-d} t^i H^\beta_\alpha(i) + \sum_{i=1}^{m-r} u^i G^\beta_\alpha(i)\right)\Lambda_\beta = \sum_{\beta=1}^{d} a^\beta_\alpha \Lambda_\alpha,
\end{aligned}$$

where a^β_α denotes the α, β entry of $A(\lambda, t, u)$.

Substituting into (C.6) and permuting columns

$$\begin{aligned}
\det M &= (-1)^r \det\left(\sum_{\beta=1}^{d} a^\beta_1\Lambda_\beta, \dots, \sum_{\beta=1}^{d} a^\beta_d\Lambda_\beta, \Psi(\kappa(1)), \dots,\right. \\
&\qquad\qquad \left.\Psi(\kappa(r-d)), N(0), N(1), \dots, N(m-r)\right) \\
&= (-1)^r \det(\Lambda_1, \dots, \Lambda_d, \Psi(\kappa(1)), \dots, \\
&\qquad\qquad \Psi(\kappa(r-d)), N(0), N(1), \dots, N(m-r)) \\
&\quad \times \det\begin{pmatrix} A(\lambda, t, u) & 0 \\ 0 & I \end{pmatrix} \\
&= (-1)^r \det(\Lambda_1, \dots, \Lambda_d, \Psi(\kappa(1)), \dots, \\
&\qquad\qquad \Psi(\kappa(r-d)), N(0), N(1), \dots, N(m-r)) \\
&\quad \times \det A(\lambda, t, u),
\end{aligned}$$

where I denotes the $(m + 1 - d) \times (m + 1 - d)$ identity matrix. As in Weyl (1939), the square of the first determinant in this expression is the determinant of the matrix of inner products of its columns. Since the latter matrix is a block diagonal matrix, where the blocks are (g_{ij}), (σ_{ij}) and I, the result follows. □

PROOF OF LEMMA 3.6. Note that multiplication of each u^i by a scale factor $\rho \geq 0$ has the effect of multiplying each column of the $(d - s) \times (d - s)$ matrix $\sum_{i=1}^{m-r} u^i G(i)\binom{J^c}{L^c}$ by ρ, so by changing to spherical coordinates the integral becomes

$$\int_{\rho=0}^{\Delta} \frac{\rho^{d-s}}{\{1 + \xi^2(t) + \rho^2\}^{(m+1)/2}} \frac{d}{d\rho}(V_{m-r}\rho^{m-r})\, d\rho$$

$$\times \int_{\{u \in R^{m-r}:\, \|u\| = 1\}} \det\left(\sum_{i=1}^{m-r} u^i G(i) \binom{J^c}{L^c} \right) \mu(du),$$

and the first integral is given by

$$(m - r) V_{m-r} \int_{\rho=0}^{\Delta} \frac{\rho^q}{\{1 + \xi^2(t) + \rho^2\}^{(m+1)/2}}\, d\rho.$$

By making the substitution $x = (1 + \xi^2(t))/(1 + \xi^2(t) + \rho^2)$ this becomes

$$\tfrac{1}{2}(m - r) V_{m-r} \{1 + \xi^2(t)\}^{(q-m)/2}$$

$$\times \int_{x=(1+\xi^2(t))/(1+\xi^2(t)+\Delta^2)}^{1} x^{((m-q)/2)-1}(1 - x)^{((q+1)/2)-1}\, dx. \qquad \square$$

APPENDIX D

Proofs of results from Section 4.

PROOF OF THEOREM 4.1. The terms $\sum_{i=1}^{r-d} \|\tilde{H}(i)\|^2$ and $\sum_{i=1}^{m-r} \|G(i)\|^2$ are both continuous functions in the bounded domain $\mathscr{S}_d$, so the supremum in the definition of ζ^* is finite and ζ^* is strictly positive. For the other inequality, note that from the definition of ζ^*,

$$\max\left\{ \sum_{i=1}^{r-d} \|\tilde{H}(i)\|^2, \sum_{i=1}^{m-r} \|G(i)\|^2 \right\} \leq 1/\tan^2(\zeta^*) \tag{C.7}$$

for each $\lambda \in \mathscr{S}_d$. Fix $\theta < \zeta^*$ so that $\xi(t)^2 + \|u\|^2 \leq \tan^2(\theta) < \tan^2(\zeta^*)$ whenever (λ, t, u) in $E(\theta)$. Since the l_2 norm is a matrix norm [see Horn and Johnson (1985), page 297] it follows that $\rho(M) \leq \|M\|$ for any matrix M. In particular, for $(\lambda, t, u) \in E(\theta)$,

$$\begin{aligned}\rho\left(\sum_{i=1}^{r-d} t^i H(i) + \sum_{i=1}^{m-r} u^i G(i)\right)^2 &\le \left\| \sum_{i=1}^{r-d} t^i H(i) + \sum_{i=1}^{m-r} u^i G(i)\right\|^2 \\ &= \left\| \sum_{i=1}^{r-d} (\Upsilon t)^i \tilde{H}(i) + \sum_{i=1}^{m-r} u^i G(i)\right\|^2 \\ &\le \|\Upsilon t\|^2 \sum_{i=1}^{r-d} \|\tilde{H}(i)\|^2 + \|u\|^2 \sum_{i=1}^{m-r} \|G(i)\|^2 \\ &= \xi(t)^2 \sum_{i=1}^{r-d} \|\tilde{H}(i)\|^2 + \|u\|^2 \sum_{i=1}^{m-r} \|G(i)\|^2 \\ &\le \left(\xi(t)^2 + \|u\|^2\right)/\tan^2(\zeta^*) < 1,\end{aligned}$$

where the second inequality uses Cauchy–Schwarz inequality componentwise and the third inequality uses (C.7). Since (λ, t, u) is arbitrary,

$$-1 \notin \bigcup_{(\theta, \lambda, u) \in E(\theta)} \operatorname{eig}\left(\sum_{i=1}^{r-d} t^i H(i) + \sum_{i=1}^{m-r} u^i G(i)\right),$$

so $\zeta \ge \theta$. Since θ is arbitrary, the desired inequality follows. □

PROOF OF LEMMA 4.1. Using the definitions and the fact that $N(0)$ is identified with $\Phi \circ \Theta$,

$$\begin{aligned}\sum_{i=1}^{m-r} \|G(i)\|^2 &= \sum_{i=0}^{m-r} \|G(i)\|^2 - \|G(0)\|^2 \\ &= \sum_{\alpha,\beta=1}^{d} \sum_{i=0}^{m-r} G_\alpha^\beta(i)^2 - \sum_{\alpha,\beta=1}^{d} G_\alpha^\beta(0)^2 \\ &= \sum_{\alpha,\beta=1}^{d} \sum_{i=0}^{m-r} \left\{ \sum_{\eta=1}^{d} g^{\beta\eta} \langle N(i), \Lambda_{\alpha\eta} \rangle \right\}^2 \\ &\quad - \sum_{\alpha,\beta=1}^{d} \left\{ \sum_{\eta=1}^{d} g^{\beta\eta} \langle N(0), \Lambda_{\alpha\eta} \rangle \right\}^2 \\ &= \sum_{\alpha,\beta=1}^{d} \sum_{i=0}^{m-r} \left\langle N(i), \sum_{\eta=1}^{d} g^{\beta\eta} \Lambda_{\alpha\eta} \right\rangle^2 \\ &\quad - \sum_{\alpha,\beta=1}^{d} \left\{ \sum_{\eta=1}^{d} g^{\beta\eta} \left\langle \Phi \circ \Theta, \frac{\partial^2}{\partial\lambda_\alpha \partial\lambda_\eta} \Phi \circ \Theta \right\rangle \right\}^2.\end{aligned} \tag{C.8}$$

Differentiating the identity $\langle \Phi \circ \Theta, \Phi \circ \Theta \rangle = 1$ with respect to λ_α and λ_η leads to

$$\left\langle \Phi \circ \Theta, \frac{\partial^2}{\partial \lambda_\alpha \partial \lambda_\eta} \Phi \circ \Theta \right\rangle = -g_{\alpha\eta},$$

so the second term in (C.8) is just $\sum_{\alpha, \beta=1}^{d} \{\sum_{\eta=1}^{d} g^{\beta\eta} g_{\alpha\eta}\}^2 = \sum_{\alpha, \beta=1}^{d} \delta_\alpha^\beta = d$. To simplify the first term in (C.8), let $\Pi_{\mathscr{N}}$ denote the orthogonal projection onto $\mathscr{N} = \operatorname{span}\{N(0), \dots, N(m-r)\}$ and let $\Pi_{\mathscr{N}^\perp}(v)$ denote the orthogonal projection of v onto the orthogonal complement of $\mathscr{N}$ in R^{m+1}. Using the fact that $\mathscr{N}^\perp = \operatorname{span}\{\Phi_\alpha, \alpha = 1, \dots, r\}$ and Remark 3.2,

$$\begin{aligned} \sum_{i=0}^{m-r} \langle N(i), v \rangle^2 &= \|\Pi_{\mathscr{N}}(v)\|^2 = \|v\|^2 - \|\Pi_{\mathscr{N}^\perp}(v)\|^2 \\ &= \|v\|^2 - \left\| \sum_{u, w=1}^{r} g^{uw} \langle v, \Phi_u \rangle \Phi_w \right\|^2 \end{aligned}$$

and the result follows. □

REFERENCES

ATIYAH, M. F. (1967). *K-Theory*. Harvard Univ. Press, Cambridge, Mass.

BERTRAND, W. and DIGUET, W. (1848). *J. Math.* (*Liouville*) *Ser. 1* **13** 80–86.

BOHRER, R. (1967). On sharpening Scheffé bounds. *J. Roy. Statist. Soc. Ser. B* **29** 110–114.

BOHRER, R. and FRANCIS, G. K. (1972). Sharp one-sided confidence bounds over positive regions. *Ann. Math. Statist.* **43** 1541–1548.

BOOTHBY, W. B. (1986). *An Introduction to Differentiable Manifolds and Riemannian Geometry.* Academic, Orlando.

CASELLA, G. and STRAWDERMAN, W. E. (1980). Confidence bands for linear regression with restricted predictor variables. *J. Amer. Statist. Assoc.* **75** 862–868.

CORNELL, J. A. (1981). *Experiments With Mixtures*: *Designs*, *Models*, *and the Analysis of Mixture Data*. Wiley, New York.

DEMMEL, J. W. (1988). The probability that a numerical analysis problem is difficult. *Math. Comp.* **50** 449–480.

HALPERIN, M. and GURIAN, J. (1968). Confidence bands in linear regression with constraints on the predictor variables. *J. Amer. Statist. Assoc.* **63** 1020–1027.

HICKS, N. (1971). *Notes on Differential Geometry*. Van Nostrand Reinhold, London.

HORN, R. and JOHNSON, C. (1985). *Matrix Analysis*. Cambridge Univ. Press, New York.

HOTELLING, H. (1939). Tubes and spheres in n-spaces, and a class of statistical problems. *Amer. J. Math.* **61** 440–460.

JOHANSEN, S. and JOHNSTONE, I. M. (1990). Hotelling's theorem on the volume of tubes: Some illustrations in simultaneous inference and data analysis. *Ann. Statist.* **18** 652–684.

JOHNSTONE, I. and SIEGMUND, D. (1989). On Hotelling's formula for the volume of tubes and Naiman's inequality. *Ann. Statist.* **17** 184–194.

KNAFL, G., SACKS, J. and YLVISAKER, D. (1985). Confidence bands for regression functions. *J. Amer. Statist. Assoc.* **80** 683–691.

KNOWLES, M. (1987). Simultaneous confidence bands for random functions. Ph.D. thesis, Statist. Dept., Stanford Univ.

KNOWLES, M. and SIEGMUND, D. (1987). On Hotelling's approach to testing for a nonlinear parameter in regression. *Int. Statist. Inst. Rev.* **57** 205–220.

NAIMAN, D. Q. (1986). Conservative confidence bands in curvilinear regression. *Ann. Statist.* **14** 896–906.

NAIMAN, D. Q. (1987). Simultaneous confidence bounds in multiple regression using predictor variables constraints. *J. Amer. Statist. Assoc.* **82** 214–219.

SCHEFFÉ, H. (1953). A method for judging all contrasts in the analysis of variance. *Biometrika* **40** 87–104.

SCHEFFÉ, H. (1959). *The Analysis of Variance*. Wiley, New York.

SPIVAK, M. (1979). *A Comprehensive Introduction to Differential Geometry*. Publish or Perish, Berkeley.

STEENROD, N. E. (1951). *The Topology of Fibre Bundles*. Princeton Univ. Press, Princeton, New Jersey.

STEINER, J. (1840). Uber parallele flachen. *Monatsb. Deutsch*. Akad. Wiss. Berlin, 114–118.

STOER, J. and WITZGALL, C. (1970). *Convexity and Optimization in Finite Dimensions*. Springer, Berlin.

UUSIPAIKKA, E. (1984). Exact confidence bands on certain restricted sets in the general linear model. Unpublished.

WEYL, H. (1939). On the volume of tubes. *Amer. J. Math.* **61** 461–472.

WORKING, H. and HOTELLING, H. (1929). Application of the theory of error to the interpretation of trends. *J. Amer. Statist. Assoc. Suppl.* (*Proceedings*) **24** 73–85.

WYNN, H. P. (1975). Integrals for one-sided confidence bounds: A general result. *Biometrika* **62** 393–396.

WYNN, H. P. and BLOOMFIELD, P. (1971). Simultaneous confidence bands in regression analysis. *J. Roy. Statist. Soc. Ser. B* **33** 202–217.

DEPARTMENT OF MATHEMATICAL SCIENCES
THE JOHNS HOPKINS UNIVERSITY
BALTIMORE, MARYLAND 21218