Stochastic Modelling and Applications
Vol. 4, No. 1, 2001.

GEOMETRY OF THE VON MISES DISTRIBUTION AND OF TESTS FOR ITS MEAN DIRECTION

Ashis SenGupta and S. Rao Jammalamadaka

Applied Statistics Division
Indian Statistical Institute
Kolkata, WB 700 035, INDIA And
Department of Statistics & Applied Probability
University of California
Santa Barbara, CA 93106, USA

Abstract

The geometry of the circular normal distribution, $CN(\mu,\kappa)$ and unconditional tests for μ are described. For κ known, the $CN(\mu,\kappa)$ forms a curved exponential family. The unconditional locally most powerful test proposed on the basis of the statistical curvature is shown to possess desirable optimal properties and good power performance even for small sample sizes. Tests based on the maximum likelihood estimator and the likelihood ratio, are also considered. For κ unknown, it is shown that the unconditional asymptotically locally most powerful or Neyman's C_α-test can be applied and it reduces a very simple form.

Key words: Ancillary families, C_α-test, circular normal distribution, curved exponential family.

AMS Subject Classification: 62H10

1. Introduction and Summary

The aim of this paper is two-fold. First, it is in the spirit of the differential geometry in statistical inference as exposed by Amari (1985), Barndorff-Nielsen, Cox and Reid (1986), and others and elucidates the geometry for the circular normal distribution, $CN(\mu,\kappa)$, and related tests. Secondly, we are interested in unconditional tests for the mean direction μ as opposed to the existing conditional ones.

Locally most powerful (LMP) and asymptotically LMP, Neyman's C_α, tests are constructed for the cases when κ, the concentration parameter, is known or unknown respectively, their properties are investigated and numerical and theoretical comparisons are made with some standard tests. In the former case, we show that the resulting model is a member of the curved exponential family (CEF), a submanifold imbedded in the space S of the manifold of the CN family. We consider the rigging ancillary submanifold, obtain its metric tensor and thereby demonstrate that the ancillary family is orthogonal. We obtain the test based on the maximum likelihood estimator (MLE) and the LMP test for the one-sided case and the likelihood ratio test (LRT) for the two-sided case. We study the geometry of these tests and that of the associated ancillary families. Efron's (1975) statistical curvature is presented to justify the choice of the LMP test. Exact cut-off points are tabulated and the power is also computed. The power performance is quite favorable as compared to the conditional test given in Mardia (1972). The test is admissible, has a monotone power function and is consistent. The geometric quantities evaluated earlier can be used to study the asymptotic behavior of the test. In particular, we use statistical curvature to give the third order power loss of the LMP test, while that of the test based on the MLE and the LRT can be similarly obtained. Further, asymptotic comparison done by Amari (1985, Sec. 6.2) is evaluated with exact computations, to confirm the impression that large sample size ($n > 30$) would be necessary to make LMP possibly inferior to LRT (and hence conditional tests approximately).

We consider next the case of κ unknown. No similarity or useful invariance holds. As opposed to a conditional test, we proceed in the spirit of the unconditional test to generalize the LMP test to the asymptotically LMP or C_α- test of Neyman (1959). We demonstrate that the condition (3) of Moran (1970) holds and so a reduced simple form is available for the asymptotically LMP test. The test is also consistent.

2. Geometry of the CN distribution

2.1 κ unknown.

Let $\alpha_1, \ldots, \alpha_n$ be a random sample from $CN(\mu, \kappa)$ i.e. from the p.d.f.

$$\begin{aligned} f(\alpha) &= [2\pi I_0(\kappa)]^{-1} \exp[\kappa \cos(\alpha - \mu)], \\ & \quad 0 \leq \alpha, \mu < 2\pi, 0 \leq \kappa < \infty. \end{aligned} \tag{2.1}$$

It is clear that $(\sum \cos \alpha_i, \sum \sin \alpha_i) \equiv (C, S)$, say, is sufficient for (μ, κ) and (2.1) is a member of the regular exponential family. Now, we may rewrite $f(\alpha) \equiv \exp[\theta^i x_i - \phi(\theta)]$, where $x_1 = \cos\alpha, x_2 = \sin\alpha$. The natural parameters are $\theta^1 = \kappa\cos\mu, \theta^2 = \kappa\sin\mu$. The expectation parameters given by $\nu_i = E(X_i) \equiv \partial_i\phi(\theta)$ are $\nu_1 = \rho\cos\mu, \nu_2 = \rho\sin\mu$, where $\rho = I_1(\kappa)/I_0(\kappa) = A(\kappa)$, say. The mean direction μ and the concentration parameter κ are frequently used as the parameter $\gamma = (\gamma^1, \gamma^2), \gamma^1 = \mu, \gamma^2 = \kappa$ to specify the family $\mathcal{S} = \{CN(\mu, \kappa)\}$ of the circular normal distributions. The parameter space is then the infinite open-top rectangle in the first quadrant with its base on 0 to 2π. The natural basis $\{\partial_1\}$ is $\partial_i = \partial/\partial\mu, \partial_2 = \partial/\partial\kappa$. The tangent vector T_θ is spanned by these vectors. We can identify T_θ, the differentiation operator representation of the tangent space with $T_\theta^{(1)}$ as the random variable or 1-representation of the same tangent space. From the log-likelihood function, $l(\alpha, \gamma)$, the basis $\partial_i l$ of the 1-representation is $\partial_1 l = \kappa\sin(\alpha - \mu), \partial_2 l = A(\kappa) + \cos(\alpha - \mu)$. The space $T_\theta^{(1)}$ is spanned by these two random variables, so that it consists of all the linear trigonometric functions in α defined below whose expectation vanishes,

$$T_\theta^{(1)} = \{a\kappa\sin(\alpha - \mu) + b\cos(\alpha - \mu) + c\}, \text{ with } c = -bA(\kappa).$$

One can use any of the above three parameter representations to specify the distribution. In fact, it is easy to establish the relationship between two different coordinate systems. For example, the Jacobian matrix of the coordinate transformation from γ to θ is given by,

$$B_i^\nu = \frac{\partial\theta^\nu}{\partial\gamma^i} = \begin{bmatrix} -\kappa\sin\mu & \cos\mu \\ \kappa\cos\mu & \sin\mu \end{bmatrix} \tag{2.2}$$

Let us consider the metric in the manifold of circular normal distributions. The metric tensor $g_{ij}(\gamma)$ or Fisher information matrix in the coordinate system $\gamma = (\mu, \kappa)$ of the circular normal family $CN(\mu, \kappa)$ is easily calculated by using $g_{ij}(\gamma) = -E[\partial_i \partial_j l(\alpha, \gamma)]$. The various expectations involved are computed by using the results, (i) $E \sin p(\alpha - \mu) = 0, E \cos p(\alpha - \mu) = I_p(\kappa)/I_0(\kappa), p = 1, 2, ...,$ (ii) $\sin \alpha$ and $\cos \alpha$ are odd and even functions respectively, (iii) $f(\alpha)$ is symmetric in μ and (iv) $\int \partial^i_\kappa \partial^j_\mu f(\alpha) d\alpha \equiv 0, i, j = 1, 2, \ldots$. Note that $g_{12}(\gamma)$ and $g_{21}(\gamma)$ vanish identically, since $g_{12}(\gamma) = E[\sin(\alpha - \mu)] \equiv 0$. So the basis vectors ∂_1 and ∂_2 are always orthogonal. Thus we get,

Result 2.1 The coordinate system γ is an orthogonal system, composed of two families of mutually orthogonal coordinate curves, $\gamma^1 = \mu =$ const. and $\gamma^2 = \kappa =$ const. Note however that the length of ∂_i depends on the position γ, i.e.,

$$| \partial_1 |^2 = \text{Var } [\kappa \sin(\alpha - \mu)] = \kappa A(\kappa),$$

$$| \partial_2 |^2 = \text{Var } [A(\kappa) + \cos(\alpha - \mu)] = 1 - A(\kappa)/\kappa - A^2(\kappa).$$

One can now calculate the Riemannian distance between two points, i.e., two CN distributions and the Riemannian geodesic curve connecting two CN distributions.

2.2 κ known.

Let κ be known, say equal to 1. Note that the model M consisting of the $CN(\beta, 1)$ is a submanifold imbedded in the space $\mathcal{L}$ of the circular normal distributions $CN(\beta, \kappa)$ with the coordinate system $\theta = (\beta, \kappa)$. From (2.1), it is clear that the CN family is a particular member of the regular exponential family and our $CN(\beta, 1)$ can be rewritten as $q(x, u) = f\{x, \theta(u)\}$ where, $\theta(u) = (\theta^1(u), \theta^2(u)) \equiv (\cos u, \sin u)$. Since this gives a smooth imbedding in the space of the exponential family, the family $M = \{q(x, u)\}$ is a $(2, 1)$ CEF. For definitions and discussions on CEF see e.g., Efron (1975). The family M forms a one-dimensional submanifold, i.e. a curve, imbedded in the two dimensional manifold $\mathcal{L}$. It is a unit circle in the θ-plane of the natural coordinates θ, since $\theta(u)$ satisfies $(\theta^1)^2 + (\theta^2)^2 = 1$. If we use the expectation parameter $\eta = (\eta_1, \eta_2)$ the η-coordinates $\eta(u)$ of the distribution specified by u are, $\eta_1 = A(1) \cos u, \eta_2 = A(1) \sin u$. Hence, the family M is represented also by a circle, $\eta_1^2 + \eta_2^2 = A^2(1)$, centered at 0, in the η-plane of the expectation coordinates. In terms of the expectation parameter, the tangent vector B_{ai} of $M(a = 1)$ is

$$B_{ai} = d\eta_i/du = A(1)[-\sin u, \cos u]$$

and a metric in $\mathcal{L}$ is defined by the tensor $g^{ij}(u) = E\{\partial_i l \partial_j l\}$. Now,

$$l = -ln2\pi - lnI_0(\kappa) + \kappa \cos\alpha \cos\mu + \kappa \sin\alpha \sin\mu$$

$$= ln2\pi - lnI_0(A^{-1}(\rho)) + (A^{-1}(\rho)\{\eta_1 \cos\alpha) + (A^{-1}(+\eta_2 \sin\alpha)\}/\rho, \rho^2 = \eta_1^2 + \eta_2^2.$$

In terms of the natural parameters,

$$l = -ln2\pi - lnI_0(((\theta^1)^2 + (\theta^2)^2)^{1/2}) + \theta^1 \cos\alpha + \theta^2 \sin\alpha$$

$$\partial l/\partial\theta^1 = -A(\kappa)\cos\mu + \cos\alpha, \text{ and } \partial l/\partial\theta^2 = -A(\kappa)\sin\mu + \sin\alpha.$$

$$\text{So, } g_{11} = E(\partial l/\partial\theta^1)^2 = -A^2(\kappa)\cos^2\mu + [I_0(\kappa) - I_2\cos 2\mu]/2I_0(\kappa),$$

$$g_{22} = E(\partial l/\partial\theta^2)^2 = -A^2(\kappa) + [I_0(\kappa) - I_2\cos 2\mu]/2I_0(\kappa),$$

$$\text{and } g_{12} = E(\partial l/\partial\theta^1 \cdot \partial l/\partial\theta^2) = [-A^2(\kappa)\sin 2\mu + E(\sin 2\alpha)]/2$$

$$= -(1/2)A^2(\kappa)\sin 2\mu.$$

This gives the metric g_{ij} defined in $\mathcal{L}$, from which the metric g_ab of M is given by $g_{ab} = B_a^i B_b^j g_{ij}$ where $\theta^1 \equiv \theta^1(u) = \cos u$ and $\theta^2 \equiv \theta^2(u) = \sin u$. Then, $B_a^i(u) \equiv \partial_a\theta^i(u) = [-\sin u, \cos u]$, $(a = 1, i = 1, 2)$ and hence,

$$g_{ab} = g_{11}\sin^2 u + g_{22}\cos^2 u - 2g_{12}\sin u \cos u.$$

A vector orthogonal to M, say $n^i(u)$ is such that, $n^i B_b^j g_{ij} = 0$. Normalize n^i to $B_\kappa^i(u)$ such that $B_\kappa^i B_b^j g_{ij} = 0$ and $B_\kappa^i B_\lambda^j g_{ij} = 1$. Given n^i satisfying the orthogonality condition, this orthonormalization with respect to the metric g_{ij} is achieved by taking $B_\kappa^i(u) = [n^i n^j g_{ij}]^{-1/2} n^i$.

[Fig. 2.1 to be placed here]

Let us attach a one-dimensional submanifold $D(u)$ of $\mathcal{L}$ to each point u such that $D(u)$ transverses M at the point $\theta(u)$ or $\eta(u)$. Here we attach a straight line $D(u)$ at each $\eta(u)$ in the η-coordinate system. The equation of $D(u)$ is,

$$\eta_1 = A(1)\cos u + [A(1)\cos u]v, \quad \eta_2 = A(1)\sin u + [A(1)\sin u]v,$$

i.e. $D(u) = \{\eta \mid \eta_1, \eta_2\}$ where v is the parameter specifying points on $D(u)$. Here v can be regarded as a coordinate on the line $D(u)$, where the origin $v = 0$ is chosen at the intersection of M and $D(u)$.

The imbedding $\theta = \theta(u)$ is given by $\theta^1(u) = u, \theta^2(u) = 1$, so that the Jacobian matrix $B^i_a(u)$ of the above coordinate transformation is

$$B^i_a(u) = \partial\theta^i/\partial u^a = (1,0), a = 1, i = 1,2.$$

We attach to each point $u \in M$, a rigging ancillary submanifold $R(u)$ as shown in Fig. 2.1, i.e. $R(u)$ consists of the CN distributions with fixed mean β and varying κ.

[Fig. 2.2 to be placed here]

Let $v = \kappa = 1$ be the coordinate of a point $CN(\beta, \kappa)$ in $R(u)$. Then (u, v) forms a coordinate system of $\mathcal{L}$ with the coordinate transformation $\theta^1(u,v) = u, \theta^2(u,v) = v + 1$ and the Jacobian matrix,

$$B^i_\alpha = \frac{\partial\theta^i(u,v)}{\partial\xi^\alpha} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

where $\xi = (u, v)$ and $\partial_\alpha B^i_\beta = 0$ holds. The metric tensor $g_{\alpha\beta}$ is given by,

$$g_{\alpha\beta} = B^i_\alpha B^j_\beta g_{ij} = \begin{bmatrix} t_1(\kappa) & 0 \\ 0 & t_2(\kappa) \end{bmatrix},$$

where $t_i(\kappa), i = 1,2$, are given below and $g_{ij} \equiv < \partial_i, \partial_j > = E[\partial_i l \partial_j l]$

$$g_{11} = E[\kappa^2 \sin^2(\alpha - \beta)] = \kappa A(\kappa) \equiv t_1(\kappa),$$

$$g_{22} = E[\cos^2(\alpha - \beta) + A^2(\kappa) - 2A(\kappa)\cos(\alpha - \beta)]$$

$$= 1 - A(\kappa)/\kappa - A^2(\kappa) \equiv t_2(\kappa),$$

and $g_{12} = E[-\kappa A(\kappa)\sin(\alpha-\beta) + \kappa\sin(\alpha-\beta)\cos(\alpha-\beta)] = 0.$

Let a, b, c, standing only for 1, be indices for u and κ, α, μ, standing only for 2, be indices for v. So, $g_{a\kappa} = 0$ and this implies that the metric (Fisher information) of M is $g_{ab}(u) = t_1(1)$ where we put $v = 0$, i.e. $\kappa = 1$ and the metric of $R(u)$ on M is $g_\kappa \lambda(u) = t_2(1)$ and ∂_a and ∂_κ are mutually orthogonal, $g_{a\kappa}(u) = 0$. Thus we have,

Result 2.2 The ancillary family is orthogonal. The α-connection $\Gamma^{(\alpha)}_{\alpha\beta\gamma}(u)$ in the associated coordinate system $\xi = (u, v)$ is,

$$\Gamma^{(\alpha)}_{abc}(u) \equiv E[\{\partial_a\partial_b l(x,u) + \frac{1-\alpha}{2}\partial_a l \partial_b l\}\partial_c l]$$

$$= \Gamma^{(1)}_{abc} + \frac{1-\alpha}{2}T_{abc}, \text{ where } T_{abc} \text{ is a third order tensor.}$$

Now, $\partial_a \equiv \partial l/\partial\beta = \kappa\sin(\alpha-\beta), \partial_a^2 \equiv \partial^2 l/\partial\beta^2 = -\kappa\cos(\alpha-\beta)$. Hence,

$$\Gamma^{(\alpha)}_{abc}(u) = E[-\kappa^2\cos(\alpha-\beta)\sin(\alpha-\beta) + \frac{1-\alpha}{2}\kappa^3\sin^3(\alpha-\beta)] = 0.$$

Since $\partial_\alpha \equiv \partial l/\partial\kappa = \kappa = -A(\kappa) + \cos(\alpha-\beta)$, we get similarly,

$$\Gamma^{(\alpha)}_{\kappa\lambda\mu}(u) = E[-A\prime(\kappa)\{-A(\kappa) + \cos(\alpha-\beta)\} + \frac{1-\alpha}{2}\{\cos(\alpha-\beta) - A(\kappa)\}^3]$$

$$= [(1-\alpha)/4I_0(\kappa)][2E\cos^3(\alpha-\beta) - 3A(\kappa)(1 + I_2(\kappa) + 3A^2(\kappa)I_1(\kappa) - A^3(\kappa)]$$

$\neq 0$ for any α, identically in κ.

$$\Gamma^{(\alpha)}_{ab\kappa}(u) = E[-A(\kappa) + \cos(\alpha-\beta)\}\{-\kappa\cos(\alpha-\beta) + ((1-\alpha)/2)\kappa^2\sin^2(\alpha-\beta)\}]$$

$$= (1/2I_0(\kappa))[2\kappa A(\kappa)I_1(\kappa) - \kappa(1 + I_2(\kappa)) - (1-\alpha)\{\kappa^2 A(\kappa)(1 - I_2(\kappa))/2$$

$$+(I_1(\kappa) - \kappa(1 - I_2(\kappa)) + \kappa(1 - I_2(\kappa))/2)\}],$$

where $E \sin^2(\alpha - \beta)\cos(\alpha - \beta)$ may be obtained, for example, by using the identity, $\int \delta_\kappa \delta_\beta^2 f(\alpha) d\alpha \equiv 0$. Thus, $\Gamma_{ab\kappa}^{(\alpha)}(u) \neq 0$ for any α, identically in κ. Hence, $H_{ab\kappa}^{(\alpha)}(u) = \Gamma_{ab\kappa}^{(\alpha)}(u)$, which gives the imbedding curvature of $R(u)$ in L at $\theta = \theta(u)$ does not vanish for any α, identically in $v = \kappa - 1$. However, the Riemann-Cristoffel curvature vanishes identically, since every one-dimensional manifold must be curvature-free. The above observations give,

Result 2.3 The model M is not an α-flat submanifold in $\mathcal{L}$, is not an α-geodesic and, in particular, is not a 0-geodesic. However M is itself α-flat. The auxiliary submanifolds $R(u)$ are α-flat for any α. The coordinate u of M is α-affine for any α, while the coordinate v of $R(u)$ is not α-affine for any α.

3. CEF and LMP test

Consider the LMP test for $H_0 : \theta = \theta_0$ against one-sided alternatives. In the absence of a uniformly most powerful test in a curved exponential family (CEF), a locally most powerful test is an attractive choice. This test maximizes the power at "local" alternatives, i.e. for small departures of the parameter from the null value. However, Chernoff (1951) exhibits that such tests can sometimes have undesirable performance for non-local alternatives. Efron (1975) suggests that with a value of the statistical curvature $\gamma_{\theta_0}^2 < \frac{1}{8}$ one can expect linear methods to work "well". In repeated sampling situations, the curvature ${}_m\gamma_{\theta_0}^2$ based on m observations satisfies, ${}_m\gamma_{\theta_0}^2 = {}_1\gamma_{\theta_0}^2/m \equiv \gamma_{\theta_0}^2/m$ and hence one can determine the sample size which reduces the curvature below 1/8.

4. Tests for μ with κ known

Let $\alpha \sim CN(\beta, 1)$, i.e. κ is assumed to be known, equal to 1. We would like to test $H_0 : \beta = 0$ against $H_1 : \beta > 0$.

4.1 A Conditional test

Mardia (1972), noting that there is no uniformly most powerful (UMP) test against one-sided alternatives, has proposed the test based on the best critical region (BCR) for testing against a simple alternative $H_1 : \beta = \beta\prime > 0$. This test against a single specific alternative is of course quite restrictive. Also, the test is a conditional test (based on Fisher's principle of ancillarity). Let R and $\bar{\alpha}$ denote the length and the direction of the vector resultant i.e., $C = R\cos\bar{\alpha}, S = R\sin\bar{\alpha}$. Then given $R = r$ the most powerful test for $H_0 : \beta = 0$ against $H_1 : \beta = \beta\prime$ has the critical region ω, $\omega : \sin(\bar{\alpha} - \beta\prime/2) > D$, where $\bar{\alpha} \sim CN(\beta, r)$ and D is a constant chosen to satisfy the level condition. For further details, see Mardia, pp.138-139. Some power values are given in Table 4.1. Note that there is minimal change in power at δ as $\beta\prime$ changes and is almost negligible for lage r . For example, the blocks of power values are identical at $r = 7$ for various β and the maximum powers i.e., at $\delta = \beta\prime$, were seen to differ only in the eighth decimal place.

[Table 4.1 to be placed here]

4.2 Likelihood ratio test

The LRT reduces to $\omega : R - C > k$ when $\hat{\mu} > \mu_0 = 0$. Then,

$$\alpha = \int_0^n t_0(r)h_n\{I_0)\}^{-n}dr$$

where $t_0(r) = (1/2\pi)\int_a^b \exp(r\cos\alpha)d\alpha$, $a = 2\pi - \cos^{-1}(k/r+1)$ and $b = \cos^{-1}(k/r+1)\text{mod}2\pi$. Power at β is obtained by replacing a and b above by, $a' = 2\pi - \cos^{-1}(k/r+1) - \beta$ and $b = \cos^{-1}(k/r+1) - \beta \bmod 2\pi$. Comparison with LMP shows that LMP is superior in the more reasonable range of 0 to $\pi/2$, almost up to $\pi/2$. Further, Amari's (1985, p. 182) result with his $t = 1.2$ shows that for $\beta \cdot\cong t\sqrt{ng}$ where $g = A(1)$, in actual computation, requires quite large n (even more than 30) for the approximation to work, i.e. LRT to be superior to LMP which is discussed below. With $t > 2$ however, the LMP test still dominates in a wide range. For example, $t = 1.2, n = 7$ gives $\beta = .6788$ while $t = 1.2$ gives $\beta = 1.1314$ while for $n = 30$, the values are $\beta = .3279$ and $\beta = .5465$. Table 4.4 exhibits such power comparisons.

4.3 Locally most powerful test

In practice we will be usually interested in testing against alternatives close to the null, say $0 < \beta < \pi/2$. Here we consider such one-sided composite alternatives. Note that the given distribution is a member of (2,1) CEF. A UMP test does not exist and a LMP test is then a reasonable candidate. However, in view of the comments in Section 3, the curvature of CN(β, 1) can be illuminating. It follows that at $\beta = 0$, for any $\kappa, \gamma_0^2(k) = A(\kappa)/\kappa(A'(\kappa))^2$. In particular, let $\kappa = 1$. Then for n observations, ${}_n\gamma_0^2(1) < 1/8$, we need $n > 14.266$, i.e. $n = 15$. Thus, according to Efron's rule, a sample size of only 15 will suffice to reduce the curvature below the ciritical value and to expect the LMP test to work "well". Such a sample size, in practice, should be easily available. The LMP test, by definition, has the critical region ω,

$$\omega \ : \ \sum_i \partial lnf(\alpha_i)\partial\beta|_{\beta=0} > K, \text{ or } \omega : R\sin\bar{\alpha} > K, \text{ i.e. } \omega : S > K, \tag{4.1}$$

where K is a constant chosen to satisfy the level condition. The exact cut-off points are given in Table 4.2.

It is interesting and instructive to look at the geometry of the LMP test (Fig. 4.1). Recall from Section 2.2 that the family M forms a circle centered at $(0,0)$. Now, the ancillary family $R(u)$ is given by, $\lambda(\eta) = c$. i.e. $\eta_2 = c$. The straight line $R(u)$ intersects M at two points u_- and u_+ for $c < 1$ and at one point $u^* = 1$ for $c = 1$. A parametric representation independent of c seems infeasible.

Fig. 4.1 to be placed here

Result 4.1 The LMP test is an unconditional test, is unbiased, has monotone power function, is consistent and is admissible - all globally for $0 < \beta < \pi/2$.

Proof: The first property is obvious from (4.1). Since β is a location parameter, monotonicity and hence unbiasedness follow e.g., by stochastic ordering. Consistency is easy to establish and admissibility is a consequence of the uniqueness of the non-randomized LMP critical region.

We next obtain the exact power of the unconditional LMP test and compare its performance, pointwise, against the corresponding conditional (most powerful) tests given in Mardia (1972). An overall comparison may be obtained by using average power, averaged over the conditioning variable or looking at the power envelope. A general comparison may also be made by evaluating the deviations of the unconditional power from the maximum and minimum powers corresponding to the values of the conditioning variable. From this comparison in Table 4.3, the performance of the LMP test looks quite encouraging.

To obtain the cut-off points, note that,

$$P(R\sin\bar{\alpha} \geq K \mid H_0) = \int_0^n P(R\sin\bar{\alpha} \geq K \mid R = r, H_0)p_R(r)dr$$

$$= \int_0^n P(\bar{\alpha} \in \omega_r \mid H_0)p_R(r)dr, \tag{4.2}$$

where $\omega_r = \text{arc}[\{(\pi/2 - \delta_r), (\pi/2 + \delta r)\}(\text{mod}2\pi)], \delta_r = \cos^{-1}(K/r)$. But, $\bar{\alpha} \mid (R = r) \sim CN(\beta, \kappa = r)$, i.e. $CN(0, r)$ under H_0. Also, $p_R(r) = \{I_0(1)\}^{-n}I_o(r)h_n(r)$, where, $h_n(r) = r\int_0^\infty uJ_0^n(u)J_0(ru)du, 0 < r < n$. Some values of $h_n(r)$ for various r , depending on n, is given in Greenwood and Durand (1955) and some extensive tables are available from SenGupta and Sastri (1988). An iterative technique, e.g., the bisection method, was used to obtain K and the cut-off points are given in Table 4.2.

To compute the power, note that, β is a location parameter. So, for $\beta = \beta\prime > 0, (\alpha - \beta\prime) \mid (R = r) \sim CN(0, r)$. Then,

$$\text{Power } (\beta\prime) = P(R\sin\bar{\alpha} > K \mid \beta\prime) = \int_0^n P(\bar{\alpha} \in \omega\prime \mid \beta = 0)p_R(r)dr,$$

where $\omega\prime_r = \text{arc}[\{(\pi/2 - \delta_r - \beta\prime), (\pi/2 + \delta_r - \beta)\}(\text{mod) } 2\pi]$. Given K, power can be obtained through numerical integration. We briefly discuss below the computational procedures.

(i) Case 1. n large (> 10). For large n we have 20 tabulated values for $h_n(r)$ from Greenwood and Durand. We use the method of bisection. We compute for every $r, r = 0.5(0.5)10.0$,

$$t_i(r) = (1/2\pi)\int_{\pi/2-\delta(i)}^{\pi/2+\delta(i)} \exp(r\cos\alpha)d\alpha, \quad \delta(i) = \cos^{-1}(K(i)/r), i = 1, 2.$$

We started with $\delta(1) = \pi/30$ and $\delta(2) = \pi/2$ and used NAG subroutine D01 GAF to compute $t_i(r)$ and also G_i,

$$G_i = \int_0^n t_i(r)h_n(r)\{I_0(1)\}^{-n}dr, \quad i = 1, 2$$

until the convergence was obtained for some $\delta(i) = \delta$, say, or equivalently $K(i) = K$. To compute the power, using D01 GAF again, we calculated for $r = 0.5(0.5)10.0$,

$$t_i(r) = (1/2\pi)\int_{\pi/2-\delta_r-\beta\prime}^{\pi/2+\delta_r-\beta\prime} \exp(r\cos\alpha)d\alpha.$$

$$\text{Power } (\beta\prime) = (1/2\pi)\int_0^n t(r)h_n(r)\{I_0(1)\}^{-n}dr.$$

(ii) Case 2. n small. Here the tabulated values being small in number, we use Gaussian quadrature and IMSL subroutine DCADRE to compute $h_n(r) \cdot h_n(r)$ was computed through a standard iterative technique until convergence was achieved for the varying upper limit. The IMSL subroutine MMBSJ0 was used. Using these values of $h_n(r)$, computations were then done exactly as in Case 1.

Some cut-off points are given in Table 4.2 while certain power values are tabulated in Table 4.3. Table 4.1 gives the powers for the best test for H_0 against a single value for the alternative. Comparisons of the performances of the LMP test and this restrictive test can now be made as was indicated earlier. For example, for the best test, at $H_1 : \beta\prime = 10$, maximum power = .1130 (for $r = 7$) and minimum power = .0594 (for $r = 1$), while the LMP test has power .09.

[Tables 4.2 and 4.3 to be placed here]

We next consider two tests for the two-sided alternatives and explore the geometry of these tests. The small-sample optimality for these tests are not known. However, together with the two-sided LMP test which may be derived, we present in section 5 a large-sample higher-order power comparison of these three tests.

4.4 Test based on the MLE

It is easy to see that the MLE of β is given by, $\hat{\beta} = \tan^{-1}(S/C)$. For testing $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$, consider the test given by, $\omega : \hat{\beta} > K_1$ or $< K_2$. In terms of the coordinates $\eta(u) = (\cos u, \sin u)$, M forms a unit circle, $\eta_1^2 + \eta_2^2 = 1$, centered at (0, 0) in S.

From the defining equation for $\hat{\beta}$ it then follows that the ancillary family associated with the test T is given by the family of straight lines which pass through the center (actually origin) of the above circle. For a given significance level, the critical region is bounded by the pair of these lines. We introduce in each ancillary subspace (line in this case) A(u), a local coordinate system v, which is defined as the distance from the intersecting point of M and A(u). Then a point $\eta = (\eta_1, \eta_2)$ in a neighborhood of M can be expressed in terms of the local coordinate system (u,v) as $\eta_1 = (1-v)\cos u$, $\eta_2 = (1-v)\sin u$.

[Fig. 4.2 to be placed here]

4.5 Likelihood ratio test

We consider the problem of testing $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$. Let $\lambda(\alpha_1, \ldots, \alpha_n)$ be the LRT statistic, i.e.,

$$\lambda(\alpha_1, \ldots, \alpha_n) = \Pi_{i=1}^n [f(\alpha_i, 0)/f(\alpha_i, \hat{\beta})] = \exp[\sum_{i=1}^n \cos\alpha_i - \cos(\alpha_i - \hat{\beta})].$$

$$-2\ln\lambda = 2[(\cos\hat{\beta} - 1)C + (\sin\hat{\beta})S] \equiv U(C, S), \text{ say.}$$

Therefore, the ancillary family is given by, $\{U(\eta_1, \eta_2) = k\}$. Now, $\cos\hat{\beta} = C/(C^2 + S^2)^{1/2}$ and $\sin\hat{\beta} = S/(C^2 + S^2)^{1/2}$.

Therefore, $U(\eta_1, \eta_2) = k \Rightarrow \left(\frac{\eta_1}{\sqrt{\eta_1^2+\eta_2^2}} - 1\right)\eta_1 + \frac{\eta_2}{\sqrt{\eta_1^2+\eta_2^2}} \cdot \eta_2 = k/2$

i.e.,

$$\eta_2^2 = k(\eta_1 + k/4). \tag{4.3}$$

Thus, the ancillary subspaces are defined by the parabolas as given in (4.3). Now, let us represent these parabolas in a parametric form. Let us consider a parabola $A(u) : \eta_2^2 = k(\eta_1 + k/4)$, which intersects the unit circle M (our family of CN distribution) at the point Q. So, coordinate of Q is $(\cos u, \sin u)$. Let P be any point on A(u) . Let the coordinate of P be (η_1, η_2). Let $PN \perp OX$ and $OP = x$. Let ρ correspond to the value of MLE for $\eta = (C, S)$, i.e. $\angle PON = \rho$. We will express the coordinates of P in a parametric form where ρ is the parameter. Clearly, $\eta_1 = x\cos\rho, \eta_2 = x\sin\rho$. Now, by the property of a parabola,

$$x = k/2 + \eta_1 = k/2 + x\cos\rho, \text{ i.e., } x(1-\cos\rho) = k/2. \tag{4.4}$$

Now, from (4.3), as $(\cos u, \sin u)$ is a point on A(u), we have, $\sin^2 u = k\cos u + k^2/4$, i.e., taking positive sign, $k = 2(1-\cos u)$. Then from (4.4), $x(1-\cos\rho) = k/2 = (1-\cos u)$. So, $x = (1-\cos u)/(1-\cos\rho)$. Thus, the ancillary subspaces A(u) can be represented parametrically as:

$$\eta_1 = \cos\rho(1-\cos u)/(1-\cos\rho), \eta_2 = \sin\rho(1-\cos u)/(1-\cos\rho).$$

(Fig. 4.3 to be placed here)

Consider next the local coordinates associated with the test. In a neighborhood of M the parabola can be regarded as the pair of two ancillary subspaces or pieces of the curves A(-u) and A(u). We introduce a local coordinate system v in each of these ancillary subspaces which is defined as the arc length from the intersecting point of M and A(u). Then, from the formula for arc length,

$$v = \int_u^\rho [\{\eta_1\prime(\rho)\}^2 + \{\eta_2\prime(\rho)\}^2]^{1/2} d\rho,$$

where differentiation is w.r.t. ρ. Now,

$$\eta_1\prime(\rho) = \frac{-\sin\rho(1-\cos u)}{(1-\cos\rho)^2} \quad \text{and} \quad \eta_2\prime(\rho) = -\frac{1-\cos u}{1-\cos\rho}.$$

$$\text{So,} v = \int_u^\rho (1-\cos u)\left[\frac{\sin^2\rho + (1-\cos\rho)^2}{(1-\cos\rho)^4}\right]^{1/2} d\rho = \sqrt{2}\int_u^\rho \left[\frac{(1-\cos u)}{(1-\cos\rho)^{3/2}}\right] d\rho$$

5. Higher-order power comparison

For the case of the two-sided alternatives, the test based on the MLE and the LR test discussed above and also the LMP test are unconditional tests. Except for the LMP test, which is optimal for all sample sizes in the sense of maximum local power, no small-sample property of the other two tests is known. However, using standard results, e.g., following Amari (1985), we get the following results on the deficiencies of the tests.

Result 5.1 The third order power loss of the LMP, MLE and the LR tests are given by, respectively, at t as,

$$\begin{aligned} L_1(t) &= 1.78\tau^2\xi(t)[1 - 1/(2\tau^2) - J(t)]^2 \\ L_2(t) &= 1.78\tau^2\xi(t)J^2(t), \text{ and} \\ L_3(t) &= 1.78\tau^2\xi(t)[1/2 - J(t)]^2, \end{aligned}$$

where τ is the upper $100\alpha/2\%$ point of the standard normal distribution, $\phi(\cdot)$, and

$$\xi(t) = (t/2)[\phi(\tau - t) - \phi(\tau + t)] \text{ and } J(t) = 1 - t/[2\tau \tanh t\tau].$$

Proof: This result follows from Theorems 6.6, 6.7, and 6.8 of Amari using $\gamma^2 = 1.78$ from Section 4.3 above.

6. Tests for μ with κ unknown

When κ is unknown, the principle of similarity or meaningful invariance does not lead to any reduction and hence no unconditional useful test is available. One approach, a very restrictive one, may be to use a conditional test (Mardia, p.143). However, here, we show that an unconditional asymptotically optimal test, e.g. Neyman's C_α test can be derived. Following the notations as in Neyman (1959), let $\phi = lnf(\alpha, \kappa)$. Then at $\mu = 0, \phi_\mu = \kappa \sin\alpha, \phi_\kappa = \cos\alpha - A(\kappa)$. Assume $\kappa \leq K < \infty$. Then straightforward computations establish that all the conditions (i) - (v) for ϕ_μ and ϕ_κ to be Crámer functions, as stated in Definition 3 in Neyman, are satisfied. Thus for testing $H_0 : \mu = 0$ against $H_1 : \mu > 0$, the C_α- test is given from Theorem 3 of Neyman as

$$Z_n^* = \sum_{i=1}^{n}\{\phi_\mu(\alpha, \hat{\kappa}) - a_1^0\phi_\kappa(\alpha, \hat{\kappa})\}/\sqrt{n}\sigma_0(\hat{\kappa}) > \tau_\alpha \quad (6.1)$$

where $\hat{\kappa}$ is any locally root n consistent estimator of κ under $H_0, \sigma_0(\hat{\kappa})$ is the standard deviation of $\phi_\mu(\alpha,\kappa) - a_1^0\phi_\kappa(\alpha,\kappa)$ under H_0 and evaluated at $\kappa = \hat{\kappa}$ and a_1^0 is the partial regression coefficient of ϕ_μ on ϕ_κ. One may, e.g., take $\hat{\kappa}$ as the MLE of κ under H_0, i.e., $\hat{\kappa} = Max\{0, A^{-1}(C/n); C > 0\}$. Further, a_1^0 is seen to be 0 by direct computation. In fact, condition (3) of Moran (1970), i.e. $E(\partial^2 1/\partial\mu\partial\kappa) = 0$ under H_0, holds. Then, the numerator of Z_n^* reduces to $\kappa \sum \sin\alpha_i$ and thus $\sigma_0^2(\kappa)$ reduces to, $\sigma_0^2(\kappa) = Var_{\mu=0}(\kappa \sin\alpha) = \kappa A(\kappa)$. Thus (6.1) reduces to the simple form,

$$Z_n = \sqrt{\hat{\kappa}} \sum_{i=1}^{n} \sin\alpha_i/(nA(\hat{\kappa}))^{1/2} > \tau_\alpha. \tag{6.2}$$

(6.2) involves computation of $\hat{\kappa}$. This may be avoided to give an even simpler but nevertheless (asymptotically) equivalent test. Note that, $\sigma_0^2(\kappa) = \kappa^2 E_0(\sin^2\alpha)$ and $\sum_{i=1}^{n} \sin^2\alpha_i/n$ is a consistent estimator of $E_0(\sin^2\alpha)$. Then (6.2) reduces to,

$$T_n = \sum_{i=1}^{n} \sin\alpha_i/(\sum_{i=1}^{n} \sin^2\alpha_i)^{1/2} > \tau_\alpha$$

T_n is equivalent to Z_n in the sense that it has, by Slutsky's theorem, the same limiting distribution as that of Z_n.

For any sequence $\mu^* = \{\mu_n\}$ such that $\mu_n\sqrt{n} \to \tau$ the asymptotic value of the power of the test is given by

$$1 - (1/\sqrt{2\pi}) \int_{-\infty}^{\tau_\alpha} \exp\{-(t - \sigma_0(\kappa)\tau)^2/2\} dt.$$

Among all tests, T_n^*, for $H_0 : \mu = 0$ with asymptotic level of significance α, whatever be the sequence of alternatives $\mu_n > 0$ with $\mu_n \to \mu_0 = 0$, and whatever be the fixed $\kappa > 0$,

$$\underline{\lim}[\text{ Power } \{T_n(\mu_n,\kappa)\} - \text{ Power } \{T_n^*(\mu_n,\kappa)\}] \geq 0.$$

The test T_n is in this sense an asymptotically locally most powerful test.

7. Acknowledgements

The authors would like to thank Mr. C.H. Sastri and Mr. M. Kafai for their help with the computations of Tables 4.1 and 4.2.

References

[1] Amari, S. (1985). *Differential-geometrical methods im statistics.* Lecture notes in statistics, **28**, Berlin: Springer-Verlag.

[2] Barndorff-Nielsen, O.E., Cox, D.R. and Reid, N. (1986). The role of differential geometry in statistical theory. *Int. Statist. Review* **54**, 83-96.

[3] Chernoff, H. (1951). A property of some Type A regions. *Ann. Math. Statist.* **22**, 472-74.

[4] Efron, B. (1975). Defining the curvature of a statistical problem. *Ann. Statist.* **3**, 1189-1242.

[5] Greenwood, J.A. and Durand, D. (1955).The distribution of the length and components of the sum of n random unit vectors, *Ann. Math. Statist.* **26**, 233-246.

[6] Mardia, K.V. (1972). *Statistics of directional data.* London: Academic Press.

[7] Moran, P.A.P. (1970). On asymptotically optimal tests of composite hypotheses. *Biometrika* **57**, 47-55.

[8] Neyman, J. (1959). Optimal asymptotic tests of composite statistical hypotheses. In *Probability and Statistics*, ed. U. Grenander, pp. 213-234. New York: Wiley.

[9] SenGupta, A. and Sastri, C. H. (1988). Certain $h_n(r)$ values for the circular normal distribution. Unpublished report, C. S. U., Indian Statistical Institute.

TABLE 4.1 CONDITIONAL POWER OF MP TESTS FOR
$H_0 : \beta = 0$ VS $H_1 : \beta = \beta\prime(\alpha = .05, n = 7)$

$\beta\prime$	δ/r	1	2	3	4	5	6	7
	2	.051763	.053537	.055161	.056505	.057627	.058608	.059496
	4	.053587	.057313	.060801	.063730	.066209	.068408	.070421
	6	.055471	.061336	.066948	.071726	.075824	.079499	.082900
	8	.057415	.065616	.073634	.080546	.086542	.091977	.097051
2	10	.059416	.070163	.080888	.090240	.098435	.105930	.112980
	12	.061475	.074983	.088738	.010085	.111566	.121435	.130776
	14	.063589	.080086	.097213	.011243	.125989	.138555	.150507
	16	.065758	.085476	.106335	.125020	.141750	.157336	.172213
	18	.067978	.091158	.116128	.138640	.158882	.177804	.195904
	20	.070247	.097136	.126610	.153321	.177403	.199960	221556
	2	.051763	.053537	.055161	.056505	.057627	.058608	
	4	.053588	.057314	.060801	.063730	.066209	.068408	
	6	.055474	.061340	.066950	.071727	.075824	.079499	
	8	.057420	.065626	.073640	.080548	.086543	.091977	
4	10	.059427	.070181	.080899	.090243	.098436	.105930	
	12	.061492	.075014	.088757	.100861	.111567	.121435	
	14	.063614	.080133	.097241	.112445	.125991	.138556	
	16	.065793	.085543	.106378	.125033	.141753	.157337	
	18	.068025	.091252	.116188	.138659	.158886	.177805	
	20	.070308	.097261	.126693	.153346	.177409	.199961	
	2	.051762	.053536	.055161	.056505	.057627	.058608	
	4	.053587	.057313	.060801	.063730	.066209	.068408	
	6	.055475	.061342	.066951	.071727	.075824	.079499	
	8	.057424	.065632	.073643	.080549	.086543	.091977	
6	10	.059434	.070195	.080906	.090245	.098437	.105930	
	12	.061505	.075038	.088770	.100865	.111568	.121435	
	14	.063635	.080171	.097264	.112451	.125992	.138556	
	16	.065822	.085600	.106412	.125043	.141755	.157337	
	18	.068065	.091331	.116237	.138673	.158889	.177805	
	20	.070361	.097370	.126761	.153366	.177413	.199962	
	2	.051759	.053531	.055158	.056504	.057627	.058608	
	4	.053583	.057306	.060797	.063729	.066209	.068408	
	6	.055471	.061336	.066948	.071726	.075824	.079499	
	8	.057424	.065632	.073643	.080549	.086543	.091977	
10	10	.059440	.070205	.080912	.090247	.098437	.105930	
	12	.061520	.075065	.088785	.100869	.111569	.121436	
	14	.063662	.080220	.097291	.112459	.125994	.138556	
	16	.065865	.085680	.106457	.125055	.141758	.157338	
	18	.068127	.091451	.116307	.138692	.158893	.177806	
	20	.070447	.097539	.126861	.153395	.177420	.199963	
	2	.051742	.053507	.055148	.056502	.057626	.058608	
	4	.053553	.057626	.060779	.063725	.066208	.068407	
	6	.055433	.061279	.066923	.071720	.075822	.079499	
	8	.057383	.065569	.073615	.080542	.086542	.091977	
20	10	.059403	.070146	.080885	.090240	.098436	.105930	
	12	.061492	.075019	.088765	.100864	.111568	.121435	

TABLE 4.2 CUT-OFF POINTS, K OF LMP TEST ($\alpha = .05$)

n	K
5	2.4654
6	2.7013
7	2.9156
8	3.1159
9	3.3038
10	3.4816
11	3.6509
12	3.8129
13	3.9686
14	4.1189
15	4.2638

TABLE 4.3 COMPARISON OF POWERS, ζ, OF MP AND LMP TESTS.

$\beta\prime$	MP test at $H_1 : \beta = \beta\prime$ min ζ r	MP test at $H_1 : \beta = \beta\prime$ max ζ r	LMP test
2	.0518	.0595	.0566
4	.0536	.0704	.0639
6	.0555	.0829	.0718
8	.0574	.0971	.0805
10	.0594	.1130	.0900
12	.0615	.1308	.1002
14	.0637	.1505	.1112
16	.0652	.1722	.1223
18	.0682	.1959	.1354
20	.0705	.2216	.1487

TABLE 4.4 COMPARISON OF POWERS OF LMP AND LR TESTS

β \ n	LMP 7	LRT 7	LMP 20	LRT 20	LMP 30	LRT 30
0	0.050002	0.050001	0.050000	0.049999	0.049999	0.050001
2	0.056604	0.050383	0.061651	0.051171	0.064523	0.051796
4	0.063880	0.051531	0.075322	0.054703	0.082125	0.057213
6	0.071861	0.053454	0.091172	0.060639	0.103093	0.066339
8	0.080573	0.056163	0.109330	0.069046	0.127641	0.079302
10	0.090037	0.059676	0.129882	0.080012	0.155882	0.096251
12	0.100266	0.064010	0.152862	0.093626	0.187808	0.117324
14	0.111266	0.069191	0.178248	0.109974	0.223277	0.142613
16	0.123035	0.075242	0.205953	0.129125	0.262006	0.172139
18	0.135561	0.082190	0.235826	0.151114	0.303574	0.205812
20	0.148825	0.090058	0.267656	0.175932	0.347441	0.243418
22	0.162796	0.098872	0.301173	0.203518	0.392971	0.284600
24	0.177435	0.108651	0.336058	0.233747	0.439460	0.328861
26	0.192693	0.119413	0.371956	0.266427	0.486182	0.375575
28	0.208513	0.131167	0.408487	0.301301	0.532416	0.424012
30	0.224829	0.143919	0.445260	0.338048	0.577490	0.473372
32	0.241568	0.157665	0.481888	0.376293	0.620804	0.522822
34	0.258653	0.172392	0.518001	0.415617	0.661857	0.571545
36	0.275999	0.188078	0.553257	0.455573	0.700258	0.618778
38	0.293519	0.204692	0.587354	0.495703	0.735735	0.663848
40	0.311123	0.222190	0.620032	0.535550	0.768128	0.706197
50	0.397281	0.320760	0.756537	0.716789	0.884737	0.867142
60	0.472919	0.429871	0.845433	0.846710	0.942207	0.948019
70	0.530505	0.537022	0.895801	0.923116	0.967118	0.948019
80	0.566025	0.631476	0.920332	0.962209	0.976938	0.991258
90	0.577970	0.707318	0.927567	0.980699	0.979516	0.994935
120	0.472919	0.829434	0.845433	0.995149	0.942207	0.996855
150	0.224829	0.863363	0.445260	0.996906	0.577490	0.997005
180	0.050002	0.869581	0.050000	0.997142	0.049999	0.997022
210	0.006020	0.863363	0.000874	0.996906	0.000264	0.997005
240	0.000813	0.829434	0.000011	0.995149	0.000001	0.996855
270	0.000348	0.707318	0.000002	0.980699	0.00000	0.994935
300	0.000813	0.429871	0.000011	0.846710	0.000001	0.948019
330	0.006020	0.143919	0.000874	0.338048	0.000264	0.473372
360	0.050002	0.050001	0.050000	0.049999	0.049999	0.050001

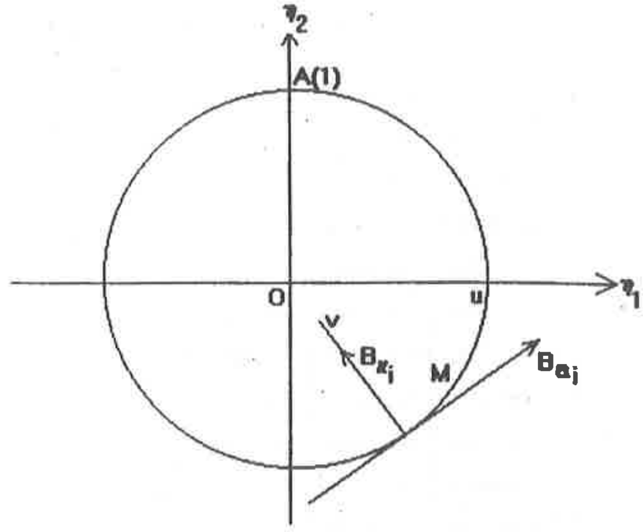

Figure 2.1 The curved exponential family CN(β,1)

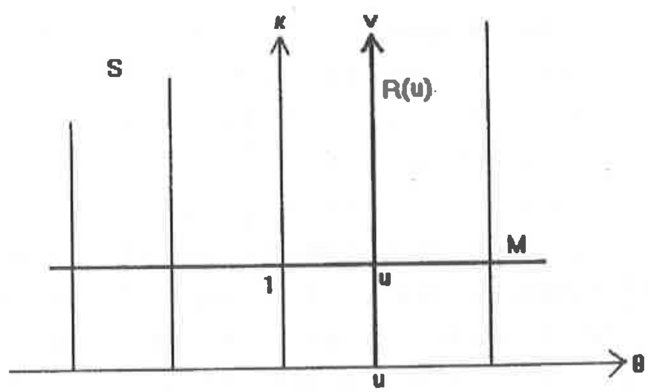

Figure 2.2 The rigging ancillary submanifold R(u)

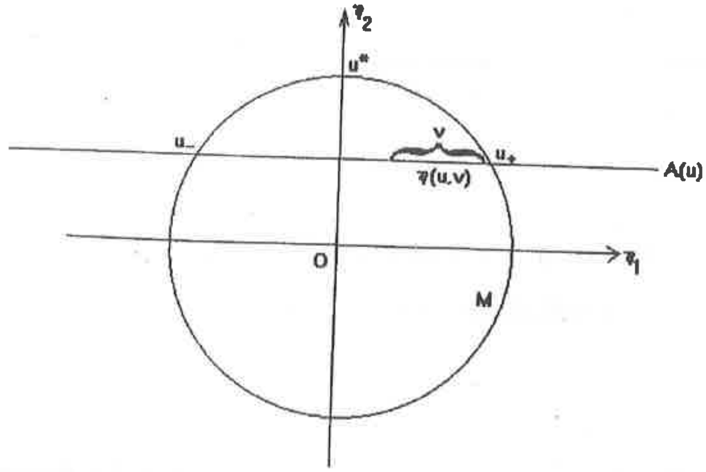

Figure 4.1 Ancillary family and (u,v) - coordinates of the locally most powerful test.

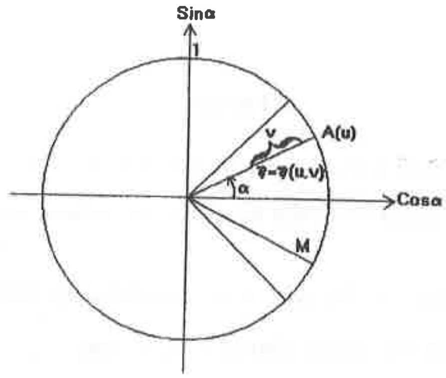

Figure 4.2 Ancillary family and (u,v) - coordinates of the test based on the maximum likelihood estimator.

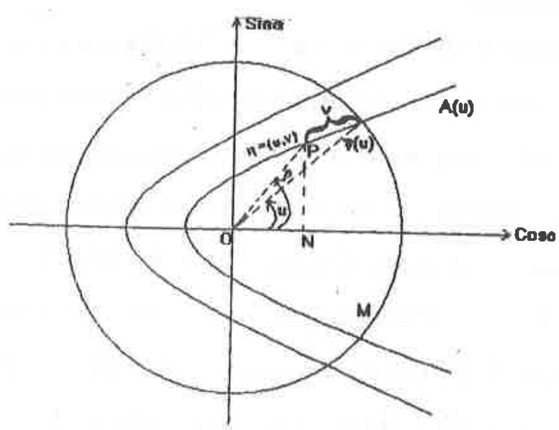

Figure 4.3 Ancillary family and (u,v) - coordinates of the likelihood ratio test.