

# A Family of Symmetric Distributions on the Circle

M.C. JONES and Arthur PEWSEY

We propose a new family of symmetric unimodal distributions on the circle which contains the uniform, von Mises, cardioid and wrapped Cauchy distributions, amongst others, as special cases. The basic form of the densities of this family is very simple, although its normalisation constant involves an associated Legendre function. The family of distributions can also be derived by conditioning and projecting certain bivariate spherically and elliptically symmetric distributions on to the circle. Trigonometric moments are available and circular variance and kurtosis are investigated in detail. The new family is compared with the wrapped normal and wrapped symmetric stable distributions. Theoretical aspects of maximum likelihood estimation are considered, and likelihood is used to fit the family of distributions to two example sets of data. The problem of the sample sizes necessary in practice to reliably discriminate between the von Mises, cardioid and wrapped Cauchy distributions is investigated. Finally, extension to a family of rotationally symmetric distributions on the sphere is briefly made.

KEY WORDS: Associated Legendre function; Cardioid distribution; Circular kurtosis; Circular uniform distribution; Distributions on the sphere; Von Mises distribution; Wrapped Cauchy distribution; Wrapped normal distribution; Wrapped symmetric stable distribution.

M.C. Jones is Professor of Statistical Science, Department of Statistics, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK (E-mail: *m.c.jones@open.ac.uk*). Arthur Pewsey is Titular de Escuela Universitaria, Departamento de Matemáticas, Escuela Politécnica, Universidad de Extremadura, Avenida de la Universidad s/n, 10071 Cáceres, Spain (E-mail: *apewsey@guadiana.unex.es*). The authors are grateful to Guy Nason for reminding them of the link between the Student's  $t$  and power of cosine distributions. Dr Pewsey gratefully acknowledges financial assistance from the Junta de Extremadura in the form of grant MOV02A044.

## 1. INTRODUCTION

The von Mises distribution has a central role as a symmetric distribution on the circle. Other popular symmetric distributions on the circle include the cardioid and wrapped Cauchy distributions. The circular uniform distribution is a special case of all of these. See Mardia (1972), Batschelet (1981), Fisher (1993), Mardia and Jupp (1999) and Jammalamadaka and SenGupta (2001) for book treatments of these distributions to which the reader is referred for standard details.

In this paper, we propose a simple general family of symmetric unimodal distributions on the circle that incorporates all of the above distributions. It has density

$$f_\psi(\theta) \propto \{1 + \tanh(\kappa\psi) \cos(\theta - \mu)\}^{1/\psi}, \quad \mu - \pi < \theta \leq \mu + \pi. \quad (1)$$

Here,  $\mu$  is a location parameter,  $\kappa \geq 0$  is a concentration parameter equivalent to the usual parameter of the von Mises distribution, and  $\psi \in \mathfrak{R}$ ; the cases  $\psi = 1, -1$  and, by continuity,  $0$ , yield the cardioid, wrapped Cauchy and von Mises distributions, respectively. See Section 2 for details including unimodality, normalising constant, special cases, graphs and genesis of the distribution. Further properties, centred on trigonometric moments and including detailed consideration of circular variance and kurtosis, are given in Section 3. A brief comparison of the new family of densities with the wrapped normal distribution and with the symmetric stable family of distributions is given in Section 4.

We consider the asymptotics of maximum likelihood estimation in Section 5 and the Appendix, with particular emphasis on the estimation of  $\kappa$  and  $\psi$ . The family of distributions is then fitted to two examples of animal orientation data from the literature (Section 6), in the second of these as a component of a mixture with the circular uniform distribution. In Section 7, we consider the sample sizes necessary in practice to be able to reliably discriminate between the von Mises, cardioid and wrapped Cauchy distributions. Finally, in Section 8, brief consideration is given to extending the family of symmetric circular distributions to a family of rotationally symmetric distributions on the sphere.

## 2. BASIC PROPERTIES

### 2.1 Unimodality and Identifiability

It can easily be shown that the density given by (1) is unimodal with mode at  $\theta = \mu$  and antimode at  $\theta = \mu + \pi$  for all  $\psi \in \Re$  and  $\kappa > 0$ . When  $\kappa = 0$ , (1) gives the uniform density for all  $\psi$ . Note that, in an obvious notation,  $f_\psi(\theta; -\kappa, \mu) = f_\psi(\theta; \kappa, \mu + \pi)$ . This distribution is the original distribution rotated to have mode at  $\mu + \pi$ ; as is common practice with e.g. the von Mises distribution, this duplication is the reason for restricting  $\kappa$  to be nonnegative throughout. In the remainder of Sections 2, 3 and 4, as well as Section 7, the location parameter  $\mu$  will be taken to be 0 without loss of generality.

### 2.2 Normalising Constant

Write  $z = \cosh(\kappa\psi) \in [1, \infty)$  so that  $f(\theta) \propto (z + \operatorname{sgn}(\psi)\sqrt{z^2 - 1} \cos \theta)^{1/\psi}$ . Then, when  $\psi > 0$ , use the following standard integral from Gradshteyn and Ryzhik (1994, (3.664.1)):

$$\int_0^\pi (z + \sqrt{z^2 - 1} \cos x)^{1/\psi} dx = \pi P_{1/\psi}(z)$$

where  $P_{1/\psi}(z)$  is the associated Legendre function of the first kind of degree  $1/\psi$  and order 0 (Gradshteyn and Ryzhik, 1994, Section 8.7–8.8). For negative  $\psi$ ,

$$\begin{aligned} \int_{-\pi}^\pi \frac{1}{(z - \sqrt{z^2 - 1} \cos x)^{-1/\psi}} dx &= \int_{-\pi}^\pi \frac{1}{(z + \sqrt{z^2 - 1} \cos(\pi - x))^{-1/\psi}} dx \\ &= \int_0^{2\pi} \frac{1}{(z + \sqrt{z^2 - 1} \cos x)^{-1/\psi}} dx \\ &= 2\pi P_{-(1/\psi)-1}(z) = 2\pi P_{1/\psi}(z), \end{aligned}$$

by (3.664.2) and (8.731.5) of Gradshteyn and Ryzhik. It follows that

$$f_\psi(\theta) = \frac{(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos \theta)^{1/\psi}}{2\pi P_{1/\psi}(\cosh(\kappa\psi))}. \quad (2)$$

### 2.3 Special Cases

*Case 1:*  $\psi = 0$ ; the von Mises distribution. The key is that, for small  $|\psi|$ ,

$$(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos \theta)^{1/\psi} \simeq \exp \left\{ \frac{1}{\psi} \log(1 + \kappa\psi \cos \theta) \right\} \simeq \exp(\kappa \cos \theta).$$

Also, setting  $\mu = 0$ ,  $z = \kappa$  and  $\nu = 1/\psi > 0$  in (9.6.49) of Abramowitz and Stegun (1965),  $\lim_{\psi \downarrow 0} P_{1/\psi}(\cosh(\kappa\psi)) = I_0(\kappa)$  where  $I_0$  is the modified Bessel function of the first kind of order 0. Essentially the same argument with  $\nu = -1/\psi$  gives the same limit for negative  $\psi$ . So density (2) tends to the von Mises density  $e^{\kappa \cos \theta} / (2\pi I_0(\kappa))$  as  $\psi$  tends to 0.

*Case 2:*  $\psi = 1$ ; the cardioid distribution.  $P_1(z) = z$  and so

$$f_1(\theta) = (2\pi)^{-1}(1 + 2\rho \cos \theta).$$

Here, we have employed the parametrisation  $0 \leq \rho = \tanh(\kappa)/2 \leq 1/2$ . Often, authors specify  $|\rho| \leq 1/2$  i.e.  $\kappa \in \mathfrak{R}$ ; we disallow negative  $\rho$  for the reason given in Section 2.1.

*Case 3:*  $\psi = -1$ ; the wrapped Cauchy distribution.  $P_{-1}(z) = P_0(z) = 1$  and so, in its usual parametrisation,

$$f_{-1}(\theta) = \frac{1}{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta)}, \quad 0 \leq \rho \leq 1,$$

where  $\rho = \tanh(\kappa/2)$ . As well as the book treatments mentioned in Section 1, see McCullagh (1996) for further insights into wrapped Cauchy models.

*Case 4:*  $0 < \psi$ ,  $\kappa \rightarrow \infty$ ; Cartwright's power-of-cosine distribution. This has density

$$\frac{2^{(1/\psi)-1} \Gamma^2((1/\psi) + 1) (1 + \cos(\theta))^{1/\psi}}{\pi \Gamma((2/\psi) + 1)} = \frac{2^{(2/\psi)-1} \Gamma^2((1/\psi) + 1) \cos(\theta/2)^{2/\psi}}{\pi \Gamma((2/\psi) + 1)} \quad (3)$$

(Cartwright, 1963). Note that Jammalamadaka and SenGupta (2001, Section 2.2.2) make an unnecessary restriction to  $\psi \leq 1$ . The normalizing constant — which has various alternative forms via properties of gamma and beta functions — follows because  $\lim_{z \rightarrow \infty} z^{-1/\psi} P_{1/\psi}(z) = \Gamma((2/\psi)+1) / \{2^{1/\psi} \Gamma^2((1/\psi)+1)\}$  for  $\psi > 0$  ((8.776.1) of Gradshteyn and Ryzhik, 1994). We also observe that Cartwright's distribution is that of  $\Theta = 2 \sin^{-1}(B)$  where  $B \sim \text{Beta}(\alpha, \beta)$  on  $[-1, 1]$  and  $\alpha = \beta = (1/\psi) + (1/2)$  for  $\psi > 0$ . Such circular beta distributions for general  $\alpha, \beta$  are discussed by Jammalamadaka and SenGupta (2001, Section 2.2.10), but the connection with Cartwright's distribution is not made there. Equivalently,  $\Theta = 2 \tan^{-1}(T/\sqrt{\nu})$  where  $T$  follows the Student  $t$  distribution with  $\nu$  degrees of freedom and  $\nu = (2/\psi) + 1 \geq 1$ , an observation also made in Minh and Farnum (2003).

*Case 5:*  $-2 \leq \psi \leq 0$ ,  $\kappa \rightarrow \infty$ ; Dirac delta function at 0. This is well known for the case  $\psi = 0$ . It follows for other  $-2 < \psi < 0$  using (8.776.1) of Gradshteyn and Ryzhik (1994) again. Write  $1/2 < \nu = -1/\psi < \infty$  and  $x = -\kappa\psi > 0$ . Then,  $(\cosh x - \sinh x \cos \theta)^\nu P_{\nu-1}(\cosh x) \sim (\cosh x)^{2\nu-1} (1 - \tanh x \cos \theta)^\nu \rightarrow \infty$  for  $\theta \neq 0$ , while for  $\theta = 0$ ,  $(\cosh x)^{2\nu-1} (1 - \tanh x)^\nu \sim e^{-x}$ . Finally, for  $\psi = -2$ ,  $P_{-1/2}(\cosh x) = \{(\pi/2) \cosh(x/2)\}^{-1} K(\tanh(x/2))$  and  $K(\tanh(x/2)) \simeq x/2$ ; here,  $K$  is the complete elliptic integral of the first kind and we have used (8.13.2) and (17.3.26) of Abramowitz and Stegun (1965). Now,  $(\cosh x - \sinh x \cos \theta)^{1/2} / \cosh(x/2) = O(1)$  for  $\theta \neq 0$ , while  $(\cosh x - \sinh x)^{1/2} / \cosh(x/2) \sim e^{-x}$  for  $\theta = 0$ , and the result follows.

*Case 6:*  $-\infty < \psi < -2$ ,  $\kappa \rightarrow \infty$ ; a new distribution with a pole at 0. Write  $0 < \nu = -1/\psi < 1/2$  and  $x = \kappa\psi > 0$ , and use (8.776.1) of Gradshteyn and Ryzhik (1994) once more. Then,  $2\pi(\cosh x - \sinh x \cos \theta)^\nu P_{\nu-1}(\cosh x) \simeq \sqrt{\pi}\Gamma(-\nu + (1/2))(1 - \tanh x \cos \theta)^\nu / \{2^{\nu-1}\Gamma(-\nu + 1)\}$ . A little further manipulation allows the limiting density to be written just as in (3) except with the sine function replacing the cosine. Now, this density is that of  $\Theta = 2 \cos^{-1}(B)$  where  $B \sim \text{Beta}(\alpha, \beta)$  on  $[-1, 1]$  and  $\alpha = \beta = (1/\psi) + (1/2)$  for  $0 > 1/\psi > -1/2$ ; and  $\Theta = 2 \cot^{-1}(T/\sqrt{\nu})$  where  $t \sim t_\nu$  and  $0 < \nu = (2/\psi) + 1 < 1$ . Notice that this completes the ‘missing’ range of trigonometric transformations of beta and  $t$  distributions from Case 4.

*Cases 7 and 8:*  $\kappa = 0$  and  $\psi \rightarrow \pm\infty$ ,  $\kappa$  finite; both yield the circular uniform distribution.

## 2.4 Graphs of Density

Density (2) is plotted in Fig. 1 for  $\psi = -1.5$  to  $1.5$  in steps of  $0.5$ , in Fig. 1(a) for  $\kappa = 1/2$  and in Fig. 1(b) for  $\kappa = 2$ . Note that these choices include standard distributions in the wrapped Cauchy when  $\psi = -1$ , von Mises when  $\psi = 0$  and cardioid when  $\psi = 1$ . Considerable similarities are apparent between members of the family for the smaller value of  $\kappa$ , these similarities increasing as  $\kappa$  continues to become smaller, when all members of the family tend to the uniform distribution. That  $\psi > 0$  produces a more spread distribution than  $\psi \leq 0$  for the same value of  $\kappa$  is clear from both frames of Fig. 1. For the larger value of  $\kappa$ , the  $\psi < 0$  densities are starting to reflect their delta function limit; positive  $\psi$  continues to yield well spread distributions. There is relatively little difference between distributions with different values of  $\psi$  for small  $\kappa$  and between distributions with different values of  $\kappa$  for  $\psi > 0$ .

\* \* \* Fig. 1(a)(b) about here \* \* \*

## 2.5 Conditional Derivations

Two different constructions involving conditioning bivariate symmetric distributions onto the unit circle lead to the same conditional distributions on the circle. The first concerns a spherically symmetric distribution with nonzero location, without loss of generality taken to be at  $(m, 0)$ ,  $m > 0$  i.e.  $(X_s, Y_s) \sim f_s(x, y) \propto g\{(x - m)^2 + y^2\}$  (Fang et al., 1990, Chapter 2). Then, the polar random variables  $(R_s, \Theta_s) \sim r_s g\{(r_s^2 + m^2) - 2mr_s \cos \theta_s\}$  and, thus, conditional on  $R_s = r_s \in (0, \infty)$ ,

$$\Theta_s | R_s = r_s \sim g\{(r_s^2 + m^2) - 2mr_s \cos \theta_s\}. \quad (4)$$

Alternatively, let  $(X_e, Y_e) \sim f_e(x, y) \propto g\{(x^2 - 2\delta xy + y^2)/(1 - \delta^2)\}$ ,  $-1 \leq \delta < 1$ , follow an elliptically symmetric distribution with mean at the origin (Fang et al., 1990, Chapter 2). Then, letting  $(R_e, \Phi)$  denote polar coordinates in this case,

$$\Phi | R_e = r_e \sim g\left\{\frac{r_e^2(1 - \delta \sin 2\phi)}{1 - \delta^2}\right\}.$$

This is, of course, an axial distribution, so double the angle to get  $\Theta_e \equiv 2\Phi | R_e = r_e \sim g\{r_e^2(1 - \delta \sin \theta_e)/(1 - \delta^2)\}$ . This has mode at  $\pi/2$  if  $\delta > 0$  and at  $-\pi/2$  if  $\delta < 0$ . Recentre the distribution at 0, so that  $\sin \theta_e$  transforms to  $\sin\{\theta_e + \text{sgn}(\delta)\pi/2\} = \text{sgn}(\delta) \cos \theta_e$ , and we have

$$\Theta_e | R_e = r_e \sim g\left\{\frac{r_e^2(1 - |\delta| \cos \theta_e)}{1 - \delta^2}\right\}. \quad (5)$$

Formulae (4) and (5) match provided we make the identifications  $|\delta| = 2mr_s/(r_s^2 + m^2)$ ,  $r_e = |r_s^2 - m^2|/\sqrt{r_s^2 + m^2}$  or equivalently  $m = (r_e/2)d_+(\delta)$ ,  $r_s = (r_e/2)d_-(\delta)$  where  $d_{\pm}(\delta) = (1 - |\delta|)^{-1/2} \pm (1 + |\delta|)^{-1/2}$ .

Three special choices for  $g$  cover virtually all of the family (1). As is well known, the von Mises distribution arises from conditioning on  $R$  in the bivariate normal distribution:  $g(z) = e^{-z/2}$  yields  $f_0(\theta) \propto e^{\kappa \cos \theta}$  where  $0 < \kappa = mr_s = r_e^2|\delta|/\{2(1 - \delta^2)\} < \infty$ . The bivariate  $t$  distribution (Fang et al., 1990, Section 3.3) has  $g(z) = (1 + z/\nu)^{-(\nu/2)-1}$ ,  $\nu > 0$ , which gives  $f_\psi(\theta)$  for  $0 \leq \tanh(\kappa\psi) = 2mr_s/(\nu + m^2 + r_s^2) = r_e^2|\delta|/\{\nu(1 - \delta^2) + r_e^2\} \leq 1$  and  $0 > \psi > -1$ . Conditioning a  $t$  distribution using the nonzero location spherically symmetric version, and hence this subset of our distributions,

has been considered briefly by Shimizu and Iida (2002). The bivariate symmetric beta distribution (Fang et al., 1990, Section 3.4) is based on  $g(z) = (1 - z)^{b-(3/2)}I(0 \leq z \leq 1)$  for  $b > 1/2$ . The spherical version of this distribution centred at  $(m, 0)$  completely envelops the circle of radius  $r_s$  centred at the origin only if  $m + r_s \leq 1$ . So we condition on  $R_s = c_s(1 - m)$  for  $0 \leq c_s \leq 1$ . With  $0 \leq \tanh(\kappa\psi) = 2mc_s/(1 + m - c_s^2(1 - m)) \leq 1$ , for  $b > 3/2$ , we obtain  $f_\psi(\theta)$  for  $\psi > 0$  while values of  $1/2 < b < 3/2$  provide members of family (1) with  $-1 < \psi < \infty$ , albeit with mode shifted from 0 to  $\pi$ . Likewise, the elliptical version of the beta distribution completely envelops only circles of radius  $\sqrt{1 - |\delta|}$  or less, so we need to condition on  $R_e^2 = c_e(1 - |\delta|)$  for  $0 \leq c_e \leq 1$ . The same values of  $\psi$  then ensue with  $0 \leq \tanh(\kappa\psi) = c_e|\delta|/(1 - c_e + |\delta|) \leq 1$ .

The only member of family (1) not available from this conditional derivation is the wrapped Cauchy distribution,  $\psi = -1$ . Interestingly, the wrapped Cauchy is the only member of the family that arises by projection on to, rather than conditioning on, the unit circle. It is the distribution obtained by doubling the angle in the axial distribution that is the angular central Gaussian distribution (Mardia and Jupp, 1999, Section 3.5.6). The latter is, in fact, obtained by projecting any bivariate elliptical distribution with mean at the origin, not just the Gaussian, on to the unit circle. As is also clear from this section of Mardia and Jupp, the wrapped Cauchy distribution does not arise from projecting a bivariate spherically symmetric distribution with nonzero mean on to the circle.

### 3. FURTHER PROPERTIES

#### 3.1 Trigonometric Moments and Characteristic Function

For  $p = 1, 2, \dots$ , use the further standard integrals from Gradshteyn and Ryzhik (1994, (8.711.2)):

$$\int_0^\pi \cos(px)(z + \sqrt{z^2 - 1} \cos x)^{1/\psi} dx = \frac{\pi\Gamma((1/\psi) + 1)}{\Gamma((1/\psi) + p + 1)} P_{1/\psi}^p(z), \quad \psi > 0,$$

and

$$\int_0^\pi \frac{\cos(px)}{(z - \sqrt{z^2 - 1} \cos x)^{-1/\psi}} dx = \frac{\pi\Gamma(-(1/\psi) - p)}{\Gamma(-(1/\psi))} P_{1/\psi}^p(z), \quad \psi < 0,$$

where  $P_{1/\psi}^p(z)$  is the associated Legendre function of the first kind of degree  $1/\psi$  and order  $p$  (Gradshteyn and Ryzhik, 1994, Section 8.7–8.8). Note that  $P_{1/\psi}^0(z) \equiv P_{1/\psi}(z)$ .

Now, for a symmetric circular density like  $f_\psi$ ,  $\phi_p = E(e^{ip\Theta}) = \alpha_p$ , where  $\{\phi_p : p = 0, \pm 1, \dots\}$  is the characteristic function of  $\Theta$  and  $\alpha_p = E(\cos(p\Theta))$  is the  $p$ th cosine moment. We immediately have

$$\phi_p = \begin{cases} \frac{\Gamma((1/\psi)+1)P_{1/\psi}^p(\cosh(\kappa\psi))}{\Gamma((1/\psi)+p+1)P_{1/\psi}^p(\cosh(\kappa\psi))}, & \psi > 0, \\ I_p(\kappa)/I_0(\kappa), & \psi = 0, \\ \frac{\Gamma((1/|\psi|)-p)P_{1/\psi}^p(\cosh(\kappa\psi))}{\Gamma(1/|\psi|)P_{1/\psi}^p(\cosh(\kappa\psi))}, & \psi < 0, \end{cases}$$

where  $I_p(\kappa)$  is the modified Bessel function of the first kind of order  $p$ . As  $P_1^1(z) = \sqrt{z^2 - 1}$  and  $P_1^p(z) = 0$ ,  $p = 2, 3, \dots$ ,  $\alpha_1$  correctly reduces to  $\tanh(\kappa\psi)/2$  and  $\alpha_p$  to 0,  $p = 2, 3, \dots$ , for the cardioid distribution. For  $\psi = -1$ ,  $\Gamma(1-p)P_0^p(z) = \{(z+1)/(z-1)\}^{p/2}$  (Gradshteyn and Ryzhik, 1994, (8.751.2)), so that  $\Gamma(1-p)P_0^p(\cosh(\kappa)) = \{\tanh(\kappa/2)\}^p = \alpha_p$  for the wrapped Cauchy distribution.

### 3.2 Circular Variance

A useful measure of variation on the circle is the circular variance  $0 \leq \nu \leq 1$  given by  $\nu = 1 - \rho$ , where  $\rho$  denotes the mean resultant length. For a distribution with  $\mu = 0$ , as here,  $\rho = \alpha_1$ . From Section 3.1,

$$\alpha_1 = \frac{|\psi|}{(1+\psi)} \frac{P_{1/\psi}^1(\cosh(\kappa\psi))}{P_{1/\psi}(\cosh(\kappa\psi))}.$$

The contour plot in Fig. 2 represents the numerically calculated value of  $\nu$  as a function of  $\kappa$  and  $\psi$ . The circular variance takes its maximal value of 1 for the uniform distribution which corresponds to the left-hand edge of Fig. 2 at which  $\kappa = 0$  and is also the limiting form to which the distributions are tending as  $|\psi| \rightarrow \infty$ . Note that the circular variance is not, as a first glance might suggest, symmetric in  $\psi$ .

\* \* \* Fig. 2 about here \* \* \*

The circular variance decreases with increasing  $\kappa$  for all  $\psi$ . Here is a demonstration of that fact. Recall that, for a distribution with  $\mu = 0$ ,  $\nu =$



$1 - E(\cos \theta)$ . Then, differentiation under the integral sign shows that, for  $\psi \neq 1$ ,

$$\frac{\partial \nu}{\partial \kappa} = -\operatorname{sech}^2(\kappa\psi) \operatorname{Cov}(\cos \theta, f_{\kappa\psi}(\cos \theta))$$

where  $f_{\kappa\psi}(y) = y/(1 + \tanh(\kappa\psi)y)$ . The covariance is positive because  $f_{\kappa\psi}(y)$  is a monotone increasing function of  $y$  for all  $\kappa$  and  $\psi$ , and hence  $\nu$  decreases. An even simpler argument holds for the case  $\psi = 1$ . Direct computation shows the circular variances of Cartwright's distribution for  $\psi > 0$ , the delta function limit for  $0 \geq \psi \geq -2$ , and the distribution of Case 6 in Section 2.3 for  $\psi < -2$ , to be  $\psi/(1 + \psi)$ , 0 and  $(2 + \psi)/(1 + \psi)$ , respectively. These limiting values of the circular variance as  $\kappa \rightarrow \infty$  are in accordance with Fig. 2.

A very similar contour plot arises if we instead consider another measure of variation on the circle, the circular dispersion of Fisher (1993, p.42). We used the formula  $(1 - \bar{\alpha}_2)/(2\rho^2)$  where  $\bar{\alpha}_2 = E[\cos\{2(\Theta - \mu_1)\}]$  and  $\mu_1$  denotes the mean direction, which we believe to be a correction to the formula given by Fisher. It should be noted however that for symmetric distributions, as here, Fisher's definition is equivalent to ours.

### 3.3 Circular Kurtosis

Circular kurtosis is normally considered to be a function of  $\alpha_2$  which is given by

$$\alpha_2 = \frac{\psi^2}{(1 + \psi)(1 + 2\psi)} \frac{P_{1/\psi}^2(\cosh(\kappa\psi))}{P_{1/\psi}(\cosh(\kappa\psi))}$$

for our family. Assuming  $\mu = 0$ , and hence  $\rho = \alpha_1$ , the usual standardisation of  $\alpha_2$ , designed to account for different scales, is Mardia's (1972, (3.7.11)) kurtosis coefficient  $(\alpha_2 - \alpha_1^4)/(1 - \alpha_1)^2$ . An alternative circular kurtosis is Mardia's (1972, (3.7.10))  $(\alpha_2 - 4\alpha_1 + 3)/(1 - \alpha_1)^2 = K$ , say. Both formulae arise from consideration of a small variability matching of circular distributions with linear distributions; the former arises from the latter through a desire to make the measure approximately take the value zero for the von Mises distribution.

We offer an alternative argument and formula which eschews approximation. The variance and coefficient of kurtosis of a random variable associated with a distribution on  $\mathfrak{R}$  symmetric about zero are precisely equal to the mean and squared coefficient of variation plus one of the square of that random variable. The latter two quantities can therefore reasonably be used as

measures of scale and tailweight for any distribution on  $\mathfrak{R}^+$ . Applying this argument to the random variable  $1 - \cos \Theta$  yields the variability measure  $1 - \alpha_1$  and kurtosis measure

$$\frac{\alpha_2 + 2\alpha_1(\alpha_1 - 4) + 5}{2(1 - \alpha_1)^2}. \quad (6)$$

In fact, (6) is nothing other than  $(K/2) + 1$ .

A plot of quantity (6), obtained numerically, as a function of  $\kappa$  and  $\psi$  is given in Fig. 3. Mardia's measure (3.7.11) affords a similar plot in the area of high kurtosis, but has rather different, and unattractive, behaviour for  $\psi > 0$ . The kurtosis of Fig. 3 shares a fairly similar shape with the variance of Fig. 2 except that the kurtosis takes its minimal value of  $1/2$  for the uniform distribution and appears to increase with increasing  $\kappa$  for all  $\psi$ . We have no proof of the latter. Direct computation shows the circular kurtosis of Cartwright's distribution for  $\psi > 0$ , and the distribution of Case 6 in Section 2.3 for  $\psi < -2$ , to be  $(\psi + 2)/(1 + 2\psi)$  and  $\psi^2/\{(1 + 2\psi)(2 + \psi)\}$ , respectively. That of the von Mises distribution,  $\psi = 0$ , is 2. For  $0 > \psi \geq -2$ , manipulations using (8.1.5) of Abramowitz and Stegun (1965) suggest that the limiting kurtosis is  $4(1 + \psi)/(2 + 5\psi)$  for  $0 > \psi > -2/5$  and is  $\infty$  otherwise. These limiting values of the circular kurtosis as  $\kappa \rightarrow \infty$  are not at odds with further numerical computation along the lines of Fig. 3.

\* \* \* Fig. 3 about here \* \* \*

## 4. COMPARISONS WITH DISTRIBUTIONS NOT IN THE FAMILY

### 4.1 Comparison with Wrapped Normal Distribution

The one major symmetric distribution on the circle which is not a member of family (1) is the wrapped normal distribution. It is well known that the von Mises and wrapped normal distributions match each other reasonably closely over the entire range of values for  $\kappa$  (Stephens, 1963, Collett and Lewis, 1981). Fig. 4 is a plot of the  $L_2$  distance between density (2) and the wrapped normal distribution fitted to (2) by matching first trigonometric moments. If  $\Theta_{\text{WN}}$  is a random variable from the wrapped normal distribution, then  $E(\cos p\Theta_{\text{WN}}) = \{E(\cos \Theta_{\text{WN}})\}^{p^2} = \alpha_1^{p^2}$  when  $E(\cos \Theta_{\text{WN}})$  is set equal to  $\alpha_1$  (which is always possible since there is a wrapped normal distribution

for any value of  $0 < \alpha_1 < 1$ ). A Parseval-type identity shows the  $L_2$  distance to be

$$\left\{ \frac{1}{\pi} \sum_{p=2}^{\infty} (\alpha_1^{p^2} - \alpha_p)^2 \right\}^{1/2}. \quad (7)$$

The quantity plotted in Fig. 4 is 1000 times (7) with the sum truncated after  $p = p_{\max} = 20$ . Taking the upper limit of the sum to be larger than this made no qualitative difference to any of the contours shown, but added small-scale artefacts which can probably be put down to numerical instability.

\* \* \* Fig. 4 about here \* \* \*

Broadly speaking, the good fit of the wrapped normal distribution to the von Mises distribution,  $\psi = 0$ , is apparent and persists, albeit to a slightly lesser degree, for  $\psi > 0$ , including the cardioid distribution,  $\psi = 1$ . The dissimilarity between (2) and the wrapped normal is considerably greater for  $\psi < 0$ , including the wrapped Cauchy distribution,  $\psi = -1$ , for  $\kappa$  not too large. Most non-wrapped-normal of all are distributions with  $-1 > \psi > -2$ , which are also amongst the distributions with greatest kurtosis according to Section 3.3.

#### 4.2 Comparison with Wrapped Symmetric Stable Distribution

We are aware of only one alternative three parameter family of symmetric distributions on the circle, namely the wrapped symmetric stable distributions of Mardia (1972, Section 3.4.8e), Jammalamadaka and SenGupta (2001, Section 2.2.8) and Gatto and Jammalamadaka (2004). As their name suggests, these distributions arise by wrapping real-line stable distributions onto the circle. In general, their densities can be written only in terms of an infinite series, specifically as

$$f_{\alpha}(\theta) = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{p=1}^{\infty} \rho^{p^{\alpha}} \cos(p(\theta - \mu)) \right\}, \quad \mu - \pi < \theta \leq \mu + \pi,$$

where  $\mu \in [0, 2\pi)$ ,  $\rho \in [0, 1]$  and  $\alpha \in (0, 2]$ . The wrapped symmetric stable family and our new family are somewhat complementary. Their common special case is the wrapped Cauchy distribution. The new family includes the von Mises distribution while the wrapped symmetric stable family includes the wrapped normal distribution.

Distributions ‘between and beyond’ the von Mises/wrapped normal and wrapped Cauchy distributions — those with  $\psi < 0$  in our case and with parameter  $\alpha < 2$  in the stable case — are quite similar. First, compare the dashed curves in Fig. 1(b) of this paper with the solid curves in Fig. 2.5 of Jammalamadaka and SenGupta (2001). Second, we repeat the comparative exercise of Section 4.1 for the wrapped symmetric stable distributions: Fig. 5 (numerical artefacts included) is the result of plotting 1000 times

$$\left\{ \frac{1}{\pi} \sum_{p=3}^{20} (\alpha_1^{p^\alpha} - \alpha_p)^2 \right\}^{1/2}, \quad (8)$$

the (approximate)  $L_2$  distance between density (2) and that of the wrapped symmetric stable distribution with first and second trigonometric moments equated. In particular, the value of  $\alpha$  has been chosen to be  $\alpha = \log_2(\log_e(\alpha_2)/\log_e(\alpha_1))$  so that  $\alpha_1^{2^\alpha} = \alpha_2$ . Empirically, we note that this formula appears to yield a value of  $\alpha \leq 2$  for almost all  $\psi \leq 0$ , but often fails to give sensible values for  $\alpha$  if we try to extend the exercise to distributions with  $\psi > 0$ . Indeed, it is fair to say that there are no direct analogues of the cardioid and other  $\psi > 0$  distributions in the wrapped symmetric stable family but, as we have already seen, the wrapped normal can be rescaled to be similar to those distributions.

\* \* \* Fig. 5 about here \* \* \*

## 5. MAXIMUM LIKELIHOOD ESTIMATION

In general, our family of distributions has three parameters, location  $\mu$ , concentration  $\kappa$  and power  $\psi$ . Recall the notation  $z = \cosh(\kappa\psi)$ . Write the log-likelihood as

$$\ell(\mu, \kappa, \psi) = \frac{1}{\psi} \sum_{i=1}^n \log(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu)) - n \log\{2Q_{1/\psi}(z)\}$$

where

$$Q_\nu(z) = \int_0^\pi (\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x)^\nu dx.$$

It will prove more convenient to work in terms of  $Q$ , its derivatives and related functions rather than in terms of associated Legendre functions and

their derivatives. All such functions are amenable to numerical integration, although care has to be taken with numerical accuracy, particularly for large absolute values of  $\kappa\psi$ , and hence values of  $\tanh(\kappa\psi)$  close to  $\pm 1$ .

The score equations and elements of the observed and expected information matrix for this model are given in the Appendix. Standard asymptotic likelihood theory applies. First, note that the maximum likelihood estimate of location,  $\mu$ , is asymptotically independent of the maximum likelihood estimates of the other two parameters, on which we now concentrate. Second, it may seem, from the Appendix, that elements of the expected information matrix  $\iota_{\kappa\psi} = O(1/\psi)$  and  $\iota_{\psi\psi} = O(1/\psi^2)$  as  $\psi \rightarrow 0$ . However, it can readily be shown that the leading terms in both expressions cancel out, as does the  $O(1/\psi)$  term of  $\iota_{\psi\psi}$  after considerably more manipulation.

The asymptotic correlation between  $\hat{\kappa}$  and  $\hat{\psi}$  is given in terms of elements of the expected information matrix by  $-\iota_{\kappa\psi}/(\iota_{\kappa\kappa}\iota_{\psi\psi} - \iota_{\kappa\psi}^2)$ . Taking their formulae from the Appendix, a contour plot of this asymptotic correlation is given in Fig. 6. Note that the vertical axis in Fig. 6 runs only from  $\psi = -2$  to 2 because of numerical difficulties for larger  $|\psi|$ . The general impression is of small to moderate correlations between  $\hat{\kappa}$  and  $\hat{\psi}$  for most  $\kappa$  and  $\psi$ , with the exception of higher correlations for small  $|\psi|$ , especially  $\psi < 0$ , and large  $\kappa$ . Notice that the asymptotic correlation is positive throughout. We have not found any simple reparametrisation of  $(\kappa, \psi)$  to have smaller asymptotic correlations.

\* \* \* Fig. 6 about here \* \* \*

## 6. EXAMPLES

### 6.1 Example 1

We first consider a grouped dataset of  $n = 714$  ‘vanishing angles’ of non-migratory British mallard ducks, taken from Table 1.1 of either Mardia (1972) or Mardia and Jupp (1999), but originally given by Matthews (1961). Over a period of a year, the ducks were taken under sunny conditions from Slimbridge in Gloucestershire to various sites at distances of between 30km and 250 km. Each duck’s vanishing angle is the bearing, measured clockwise from north, at which it disappeared from view when released. Mardia (1972, p.10) describes the distribution of the data as being “somewhat symmetrical”, and application of Pewsey’s (2002) test for reflective symmetry found no evidence that the underlying distribution is not symmetric.

Maximisation of the log-likelihood surface was carried out using a grid of starting values and the Nelder-Mead simplex algorithm (Nelder and Mead, 1965). Results for fitting the full family of distributions and for fitting each of the von Mises, cardioid and wrapped Cauchy submodels are given in Table 1. The von Mises distribution proves to be much the best of these three, but it is firmly rejected by a likelihood ratio test in favour of the member of the family that maximises the likelihood, which has  $\psi = -0.35$ . Indeed, approximate 95% confidence intervals for  $\psi$  are:  $(-0.54, -0.18)$  based on the profile log-likelihood and the  $\chi_1^2$  asymptotic distribution for the log-likelihood ratio;  $(-0.53, -0.17)$  based on  $\pm 1.96$  times the standard error obtained from the observed information and  $(-0.49, -0.21)$  based on  $\pm 1.96$  times the standard error obtained from the expected information. The profile log-likelihood for  $\psi$  (not shown) is well-behaved and reasonably symmetric about  $\psi = -0.35$ , which helps account for the similarity of these intervals.

\* \* \* Table 1 about here \* \* \*

The best fitting member of the family of distributions and the fitted von Mises density are plotted together with a histogram of the data in Fig. 7. The fit of the distribution with  $\psi = -0.35$  is noticeably better than that of the distribution with  $\psi = 0$ . Chi-squared tests of the goodness-of-fit of these distributions, based on the binning shown in Fig. 7 but combining class intervals to have expected values greater than 5, emphatically reject the von Mises model ( $p$ -value 0.0002) but are much more marginal about the  $\psi = -0.35$  model ( $p$ -value 0.0467).

\* \* \* Fig. 7 about here \* \* \*

In this and the following example, occasional ‘nonsense’ solutions to the general maximum likelihood problem were found. These were clearly ill-fitting versions of the distribution with unbelievable values of the log-likelihood, and were caused by the numerical difficulties alluded to in Section 4 when  $|\kappa\psi|$  is large.

## 6.2 Example 2

Our second example concerns an experiment in which ants were placed singly in the centre of an arena and the initial direction in which they moved in relation to a visual stimulus at an angle of 180 degrees from the zero direction was recorded (to the nearest  $10^\circ$ ). The data come from Jander

(1957) as plotted in Batschelet (1981, Fig. 3.2.2(A)); they are somewhat approximate in that frequencies in class intervals were recorded to the nearest five ants. In total,  $n = 730$ , yet both Batschelet (1981) and SenGupta and Pal (2001) forgot about the data being recorded in units of five and erroneously used  $n = 146$  in their analyses of these data.

An initial fitting of our family of distributions led to a maximum likelihood solution having  $\hat{\psi} = -1.30$  with maximised log-likelihood  $-926.55$ . However, inspection of a histogram of the data together with this fitted model (see Fig. 8 later) displays a lack-of-fit which is remedied by expanding our model to have density of the form  $pf_{\psi}(\theta) + (1-p)/(2\pi)$ , a  $p : 1-p$  mixture of the new family of distributions and the circular uniform distribution. Note that this model has a clear interpretation in animal orientation experiments: a (large) proportion  $p$  of animals follow some form of ‘typical’ oriented behaviour while the remainder are completely unaffected by the experimental conditions (here, presence of the stimulus).

Results for fitting each of the full family and von Mises, cardioid and wrapped Cauchy submodels, each mixed with the uniform distribution, are given in Table 2. Clearly, the difference between maximised log-likelihoods shows that the presence of the uniform mixing component improves matters greatly over its absence. Also, there is no loss in moving from the full family/uniform mixture, with  $\psi = -0.10$ , to the von Mises/uniform mixture model.

\* \* \* Table 2 about here \* \* \*

Also in Table 2 is the result of fitting the wrapped symmetric stable distribution mixed with the uniform to these data by maximum likelihood. This model has a maximised log-likelihood only very slightly smaller than ours. What is more, the general wrapped symmetric stable/uniform mixture can be specialised to a wrapped normal/uniform mixture without loss. Indeed, on likelihood grounds (and on chi-squared goodness-of-fit grounds not discussed in detail here), there is nothing to choose between the four best models and, in particular, between von Mises/uniform and wrapped normal/uniform mixtures. See also Fig. 8. (These results enhance those of SenGupta and Pal, 2001, who fitted wrapped symmetric stable/uniform mixture distributions to these data: we used more appropriate  $n$ , maximum likelihood parameter estimation and noted the opportunity to specialise to the wrapped normal/uniform mixture.) Notice that for all models supported

by the data, the proportion of ants reacting to the stimulus is estimated to be around  $2/3$ .

\* \* \* Fig. 8 about here \* \* \*

## 7. DISCRIMINATION BETWEEN VON MISES, CARDIOID AND WRAPPED CAUCHY DISTRIBUTIONS

The examples of Section 6 suggest that the principal members of the family of distributions, the von Mises, cardioid and wrapped Cauchy distributions, can be readily discriminated between, but in each example  $n$  was over 700. How readily can these distributions be discriminated between on the basis of smaller samples? Pewsey and Jones (2004) develop the work of Collett and Lewis (1981) on discriminating between von Mises and wrapped normal distributions, and we follow one of their approaches here. In terms of models  $M_1$  and  $M_2$ , say, it consists in: (i) defining  $T = \ell_{M_1} - \ell_{M_2}$  where  $\ell_M$  is the maximised log-likelihood value assuming model  $M$  to be appropriate, a positive value of  $T$  suggesting that  $M_1$  is to be preferred to  $M_2$ ; (ii) defining the misclassification probabilities  $p_1(\rho) = P(T \leq 0 \mid \text{sample actually comes from } M_1 \text{ with mean resultant length } \rho)$  and  $p_2(\rho) = P(T \geq 0 \mid \text{sample actually comes from } M_2 \text{ with mean resultant length } \rho)$ ; (iii) setting  $p(\rho) = \max(p_1(\rho), p_2(\rho))$ ; and (iv) estimating, using simulation, values of  $p(\rho)$  for a range of values of  $\rho$  and  $n$ . For each  $\rho$ , it is of particular interest to identify the lowest value of  $n$  for which  $p(\rho)$  is small e.g. 0.1, 0.05 or even 0.01.

Performing this algorithm for  $n = 25(25)200, 250, 300(100)600$ ,  $\rho = 0.05, 0.1(0.1)0.9, 0.95$  and taking  $M_1$  and  $M_2$  to be the von Mises and wrapped Cauchy distributions resulted in Fig. 9(a). Discrimination proves possible for relatively small sample sizes if the data are sampled from highly concentrated cases of the distributions. For instance, for  $\rho = 0.95$ , as few as  $n = 50$  observations are necessary in order to ensure that both misclassification probabilities are less than 0.01. For  $\rho = 0.6$ , on the other hand, the sample sizes required to ensure that  $p(0.6)$  is less than 0.1, 0.05 and 0.01 are around 100, 150 and 370, respectively. For small  $\rho$ , huge sample sizes would be required as the distributions become more and more alike and tend to the uniform.

\* \* \* Fig. 9 about here \* \* \*



Similar results for the comparison of the cardioid distribution with the von Mises distribution and with the wrapped Cauchy distribution are presented in Fig. 9(b),(c), respectively. Note that  $\rho = 1 - \nu \leq 1/2$  for the cardioid distribution. For  $\rho = 0.5$ , discrimination between the appropriate version of Cartwright's distribution and the von Mises distribution requires around 230 observations to ensure that  $p(0.5) \leq 0.05$ . For  $\rho = 0.4$ , the sample size required to ensure that  $p(0.4) \leq 0.1$  is around 400. For smaller  $\rho$  values, huge sample sizes will be required before discrimination can be conducted with small misclassification probabilities. On the other hand, clearly we can readily discriminate between the cardioid distribution and the von Mises distribution for data samples from concentrated cases of the latter. Similar patterns of results arise for the comparison between cardioid and wrapped Cauchy distributions (Fig. 9(c)), except that discrimination between the two is much easier. For example, the sample size required to ensure that  $p(0.4) \leq 0.05$  is now only around 180.

Following Royall (1997), it can be argued that the value '0' used in both  $p_1(\rho)$  and  $p_2(\rho)$  be replaced by a nonzero value  $t$ , a sensible value for which might be  $t = \log 8$ . This would correspond to deciding in favour of model  $M_1$ , say, only if the likelihood ratio in favour of  $M_1$  over  $M_2$  were  $e^t$  or bigger. Values of  $T$  between  $e^t$  and  $e^{-t}$  would result in indifference between  $M_1$  and  $M_2$ , and values smaller than  $e^{-t}$  would indicate model  $M_2$  rather than  $M_1$ . The probabilities  $p_1(\rho)$  and  $p_2(\rho)$  are no longer misclassification probabilities per se, but probabilities of weak and misleading evidence taken together. Pewsey and Jones (2004) give details of this second approach for discrimination between von Mises and wrapped normal distributions. Here, we simply record that the same approach applied to our pairwise distribution comparisons results, clearly, in larger sample sizes to achieve reliable discrimination. The required sample sizes for those values of  $\rho$  quoted above are, very roughly speaking, increased by between 50 and 100%.

## 8. ROTATIONALLY SYMMETRIC DISTRIBUTIONS ON THE SPHERE

The immediate extension of family (1) to distributions on the sphere in  $p \geq 2$  dimensions,  $S^{p-1}$ , has density  $f_\psi(x) \propto \{1 + \tanh(\kappa\psi)x^T\mu\}^{1/\psi}$  where  $x$  and  $\mu$  are unit vectors and superscript  $T$  denotes vector transpose. As such, the distributions are apparently novel special cases of the general family of spherical models with rotational symmetry (about  $\mu$ ), see e.g. Watson (1983, Chapter 4) and Mardia and Jupp (1999, Section 9.3.3). The pre-existing

special case is the von Mises-Fisher-Langevin distribution (Mardia and Jupp, 1999, Section 9.3.2) corresponding to  $\psi = 0$ .

Following Watson and Mardia and Jupp, rotationally symmetric spherical random variables can be written  $x = T\mu + \sqrt{1-T^2}\xi$  where  $\xi$  is the unit tangent to  $S^{p-1}$  at  $\mu$ , uniformly distributed on  $S^{p-2}$  and independent of the scalar random variable  $T = x^T\mu$ . When  $x$  is uniform on  $S^{p-1}$ ,  $T$  has density  $B((p-1)/2, 1/2)^{-1}(1-t^2)^{(p-3)/2}$  on  $t \in [-1, 1]$ , where  $B(\cdot, \cdot)$  denotes the beta function. It follows that

$$\begin{aligned} \int_{S^{p-1}} (1 + \tanh(\kappa\psi)x^T\mu)^{1/\psi} dx \\ = B((p-1)/2, 1/2)^{-1} \int_{-1}^1 (1 + \tanh(\kappa\psi)t)^{1/\psi} (1-t^2)^{(p-3)/2} dt. \end{aligned}$$

This can be evaluated for  $\psi$  both positive and negative by judicious use of (8.711.1) of Gradshteyn and Ryzhik (1994). For  $\psi \neq 0$ , the density on the sphere with respect to the uniform distribution turns out to be

$$f_\psi(x) = \frac{|\sinh(\kappa\psi)|^{(p/2)-1} (\cosh(\kappa\psi) + \sinh(\kappa\psi)x^T\mu)^{1/\psi}}{2^{(p/2)-1}\Gamma(p/2) P_{(1/\psi)+(p/2)-1}^{1-(p/2)}(\cosh(\kappa\psi))}.$$

This reduces to  $2\pi$  times (2) when  $p = 2$  and also reduces to a simpler form when  $p = 3$ , namely

$$f_\psi(x) = \frac{(1+\psi)\sinh(\kappa\psi)}{\psi\sinh(\kappa(1+\psi))} (\cosh(\kappa\psi) + \sinh(\kappa\psi)x^T\mu)^{1/\psi}$$

if  $\psi \neq -1$ , with another simple formula for  $\psi = -1$ . The nonzero mean spherically symmetric conditioning arguments of Section 2.5 can be extended, as is shown in special cases by Downs (1966) and Shimizu and Iida (2002).

## APPENDIX

Use primes to denote differentiation of  $Q_\nu(z)$  with respect to  $z$  and superscript circles to denote differentiation of  $Q_\nu(z)$  with respect to  $\nu$ . We will continue to write  $z$  for  $\cosh(\kappa\psi)$  in places.

### A.1 Score Equations

Differentiate the log-likelihood with respect to each of  $\mu$ ,  $\kappa$  and  $\psi$ , in that order. Then, the score equations are

$$0 = \frac{1}{\psi} \sinh(\kappa\psi) \sum_{i=1}^n \frac{\sin(\theta_i - \mu)}{\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu)},$$

$$0 = \sinh(\kappa\psi) \left\{ \sum_{i=1}^n \frac{(1 + \coth(\kappa\psi) \cos(\theta_i - \mu))}{\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu)} - n\psi \frac{Q'_{1/\psi}(z)}{Q_{1/\psi}(z)} \right\}$$

and

$$\begin{aligned} 0 &= \frac{\kappa}{\psi} \sinh(\kappa\psi) \left\{ \sum_{i=1}^n \frac{(1 + \coth(\kappa\psi) \cos(\theta_i - \mu))}{\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu)} - n\psi \frac{Q'_{1/\psi}(z)}{Q_{1/\psi}(z)} \right\} \\ &- \frac{1}{\psi^2} \left\{ \sum_{i=1}^n \log(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu)) - n \frac{Q_{1/\psi}^\circ(z)}{Q_{1/\psi}(z)} \right\}. \end{aligned}$$

Here,

$$Q'_\nu(z) = \nu \int_0^\pi (\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x)^{\nu-1} (1 + \coth(\kappa\psi) \cos x) dx$$

and

$$Q_\nu^\circ(z) = \int_0^\pi (\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x)^\nu \log(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x) dx.$$

## A.2 Observed Information

Elements of the observed information matrix are minus the second derivatives of  $\ell$ ; denote them by  $J$  with subscripts  $\mu\mu$ ,  $\mu\kappa$ , etc. We have:

$$J_{\mu\mu} = \frac{1}{\psi} \sinh(\kappa\psi) \sum_{i=1}^n \frac{\sinh(\kappa\psi) + \cosh(\kappa\psi) \cos(\theta_i - \mu)}{(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu))^2},$$

$$J_{\mu\kappa} = - \sum_{i=1}^n \frac{\sin(\theta_i - \mu)}{(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu))^2},$$

$$J_{\mu\psi} = \frac{\kappa}{\psi} J_{\mu\kappa} + \frac{1}{\psi^2} \sinh(\kappa\psi) \sum_{i=1}^n \frac{\sin(\theta_i - \mu)}{\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu)},$$

$$\begin{aligned} J_{\kappa\kappa} &= -\psi \sum_{i=1}^n \frac{1 - \cos^2(\theta_i - \mu)}{(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu))^2} \\ &+ n\psi^2 \left\{ \cosh(\kappa\psi) \frac{Q'_{1/\psi}(z)}{Q_{1/\psi}(z)} + \sinh^2(\kappa\psi) \left( \frac{Q''_{1/\psi}(z)}{Q_{1/\psi}(z)} - \frac{(Q'_{1/\psi}(z))^2}{Q_{1/\psi}^2(z)} \right) \right\}, \end{aligned}$$

$$\begin{aligned} J_{\kappa\psi} &= -\kappa \sum_{i=1}^n \frac{1 - \cos^2(\theta_i - \mu)}{(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu))^2} + n \{ \sinh(\kappa\psi) + \kappa\psi \cosh(\kappa\psi) \} \frac{Q'_{1/\psi}(z)}{Q_{1/\psi}(z)} \\ &+ n\kappa\psi \sinh^2(\kappa\psi) \left\{ \frac{Q''_{1/\psi}(z)}{Q_{1/\psi}(z)} - \frac{(Q'_{1/\psi}(z))^2}{Q_{1/\psi}^2(z)} \right\} - \frac{n}{\psi} \sinh(\kappa\psi) \left\{ \frac{Q_{1/\psi}^\circ(z)}{Q_{1/\psi}(z)} - \frac{Q'_{1/\psi}(z) Q_{1/\psi}^\circ(z)}{Q_{1/\psi}^2(z)} \right\}, \end{aligned}$$

$$\begin{aligned}
J_{\psi\psi} &= \frac{2\kappa}{\psi^2} \sum_{i=1}^n \frac{\sinh(\kappa\psi) + \cosh(\kappa\psi) \cos(\theta_i - \mu)}{\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu)} - \frac{\kappa^2}{\psi} \sum_{i=1}^n \frac{1 - \cos^2(\theta_i - \mu)}{(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu))^2} \\
&+ n\kappa^2 \sinh^2(\kappa\psi) \left\{ \frac{Q''_{1/\psi}(z)}{Q_{1/\psi}(z)} - \frac{(Q'_{1/\psi}(z))^2}{Q_{1/\psi}^2(z)} \right\} - \frac{2n\kappa}{\psi^2} \sinh(\kappa\psi) \left\{ \frac{Q'_{1/\psi}(z)}{Q_{1/\psi}(z)} - \frac{Q'_{1/\psi}(z)Q_{1/\psi}(z)}{Q_{1/\psi}^2(z)} \right\} \\
&+ \frac{n}{\psi^4} \left\{ \frac{Q_{1/\psi}^{\circ\circ}(z)}{Q_{1/\psi}(z)} - \frac{(Q_{1/\psi}^{\circ}(z))^2}{Q_{1/\psi}^2(z)} \right\} + n\kappa^2 \cosh(\kappa\psi) \frac{Q'_{1/\psi}(z)}{Q_{1/\psi}(z)} \\
&- \frac{2}{\psi^3} \left\{ \sum_{i=1}^n \log(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta_i - \mu)) - n \frac{Q_{1/\psi}^{\circ}(z)}{Q_{1/\psi}(z)} \right\}.
\end{aligned}$$

Here,

$$\begin{aligned}
Q''_{\nu}(z) &= \nu(\nu - 1) \int_0^{\pi} (\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x)^{\nu-2} (1 + \coth(\kappa\psi) \cos x)^2 dx \\
&- \nu \sinh^{-3}(\kappa\psi) \int_0^{\pi} (\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x)^{\nu-1} \cos x dx,
\end{aligned}$$

$$\begin{aligned}
Q'_{\nu}(z) &= \int_0^{\pi} (\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x)^{\nu-1} (1 + \coth(\kappa\psi) \cos x) dx \\
&+ \nu \int_0^{\pi} (\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x)^{\nu-1} (1 + \coth(\kappa\psi) \cos x) \\
&\quad \times \log(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x) dx,
\end{aligned}$$

$$Q_{\nu}^{\circ\circ}(z) = \int_0^{\pi} (\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x)^{\nu} \log^2(\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x) dx.$$

### A.3 Expected Information

Elements of the expected information matrix follow simply from those of the observed information matrix. Denote  $n^{-1}$  times the expected information matrix by  $\iota$ . Then:

$$\iota_{\mu\mu} = \frac{1}{\psi} \sinh(\kappa\psi) \frac{R_{1/\psi}(z)}{Q_{1/\psi}(z)}, \quad \iota_{\mu\kappa} = \iota_{\mu\psi} = 0,$$

$$\iota_{\kappa\kappa} = -\psi \frac{S_{1/\psi}(z)}{Q_{1/\psi}(z)} + \psi^2 \left\{ \cosh(\kappa\psi) \frac{Q'_{1/\psi}(z)}{Q_{1/\psi}(z)} + \sinh^2(\kappa\psi) \left( \frac{Q''_{1/\psi}(z)}{Q_{1/\psi}(z)} - \frac{(Q'_{1/\psi}(z))^2}{Q_{1/\psi}^2(z)} \right) \right\},$$

$$\begin{aligned}
\iota_{\kappa\psi} &= -\kappa \frac{S_{1/\psi}(z)}{Q_{1/\psi}(z)} + \{ \sinh(\kappa\psi) + \kappa\psi \cosh(\kappa\psi) \} \frac{Q'_{1/\psi}(z)}{Q_{1/\psi}(z)} \\
&+ \kappa\psi \sinh^2(\kappa\psi) \left\{ \frac{Q''_{1/\psi}(z)}{Q_{1/\psi}(z)} - \frac{(Q'_{1/\psi}(z))^2}{Q_{1/\psi}^2(z)} \right\} - \frac{1}{\psi} \sinh(\kappa\psi) \left\{ \frac{Q'_{1/\psi}(z)}{Q_{1/\psi}(z)} - \frac{Q'_{1/\psi}(z)Q_{1/\psi}(z)}{Q_{1/\psi}^2(z)} \right\},
\end{aligned}$$

$$\begin{aligned}
\nu_{\psi} &= \frac{2\kappa}{\psi^2} \frac{R_{1/\psi}(z)}{Q_{1/\psi}(z)} - \frac{\kappa^2}{\psi} \frac{S_{1/\psi}(z)}{Q_{1/\psi}(z)} + \kappa^2 \sinh^2(\kappa\psi) \left\{ \frac{Q''_{1/\psi}(z)}{Q_{1/\psi}(z)} - \frac{(Q'_{1/\psi}(z))^2}{Q_{1/\psi}^2(z)} \right\} \\
&- \frac{2\kappa}{\psi^2} \sinh(\kappa\psi) \left\{ \frac{Q'_{1/\psi}(z)}{Q_{1/\psi}(z)} - \frac{Q'_{1/\psi}(z)Q_{1/\psi}(z)}{Q_{1/\psi}^2(z)} \right\} \\
&+ \frac{1}{\psi^4} \left\{ \frac{Q_{1/\psi}^{\circ\circ}(z)}{Q_{1/\psi}(z)} - \frac{(Q_{1/\psi}^{\circ}(z))^2}{Q_{1/\psi}^2(z)} \right\} + \kappa^2 \cosh(\kappa\psi) \frac{Q'_{1/\psi}(z)}{Q_{1/\psi}(z)}.
\end{aligned}$$

Here,

$$R_{\nu}(z) = \int_0^{\pi} (\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x)^{\nu-1} (\sinh(\kappa\psi) + \cosh(\kappa\psi) \cos x) dx$$

and

$$S_{\nu}(z) = \int_0^{\pi} (\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos x)^{\nu-2} (1 - \cos^2 x) dx.$$

#### REFERENCES

- Abramowitz, M., and Stegun, I.A. (eds) (1965), *Handbook of Mathematical Functions*, New York: Dover.
- Batschelet, E. (1981), *Circular Statistics in Biology*, London: Academic Press.
- Cartwright, D.E. (1963), "The Use of Directional Spectra in Studying the Output of a Wave Recorder on a Moving Ship," in *Ocean Wave Spectra*, Englewood Cliffs, NJ: Prentice-Hall, pp. 203–218.
- Collett, D., and Lewis, T. (1981), "Discriminating Between the Von Mises and Wrapped Normal Distributions," *Australian Journal of Statistics*, 23, 73–79.
- Downs, T.D. (1966), "Some Relationships Among the Von Mises Distributions of Different Dimensions," *Biometrika*, 53, 269–272.
- Fang, K.T., Kotz, S., and Ng, K.W. (1990), *Symmetric Multivariate and Related Distributions*, London: Chapman and Hall.
- Fisher, N.I. (1993), *Statistical Analysis of Circular Data*, Cambridge: Cambridge University Press.
- Gatto, R., and Jammalamadaka, S.R. (2004), "Inference for Wrapped Symmetric Alpha-Stable Circular Models," *Sanhkya*, to appear.

- Gradshteyn, I.S., and Ryzhik, I.M. (1994), *Table of Integrals, Series, and Products*, fifth edition edited by A. Jeffrey, San Diego: Academic Press.
- Jammalamadaka, S.R., and SenGupta, A. (2001), *Topics in Circular Statistics*, Singapore: World Scientific.
- Jander, R. (1957), “Die Optische Richtungsorientierung der Roten Waldameise (*Formica Rufa* L.),” *Zeitschrift für Vergleichende Physiologie*, 40, 162–238.
- Mardia, K.V. (1972), *Statistics of Directional Data*, London: Academic Press.
- Mardia, K.V., and Jupp, P.E. (1999), *Directional Statistics*, Chichester: Wiley.
- Matthews, G.V.T. (1961), “‘Nonsense’ Orientation in Mallard *Ans Platyrhynchos* and Its Relation to Experiments on Bird Navigation,” *Ibis*, 103a, 211–230.
- McCullagh, P. (1996), “Möbius Transformation and Cauchy Parameter Estimation,” *Annals of Statistics*, 24, 787–808.
- Minh, D.L.P., and Farnum, N.R. (2003), “Using Bilinear Transformations to Induce Probability Distributions,” *Communications in Statistics – Theory and Methods*, 32, 1–9.
- Nelder, J.A., and Mead, R. (1965), “A Simplex Method for Function Minimization,” *Computer Journal*, 7, 308–313.
- Pewsey, A. (2002), “Testing Circular Symmetry,” *Canadian Journal of Statistics*, 30, 591–600.
- Pewsey, A., and Jones, M.C. (2004), “Discrimination Between the Von Mises and Wrapped Normal Distributions: Just How Big Does the Sample Size Have to Be?” under consideration.
- Royall, R.M. (1997), *Statistical Evidence; a Likelihood Paradigm*, London: Chapman & Hall.
- SenGupta, A., and Pal, C. (2001), “On Optimal Tests for Isotropy Against the Symmetric Wrapped Stable-Circular Uniform Mixture Family,” *Journal of Applied Statistics*, 28, 129–143.
- Shimizu, K., and Iida, K. (2002), “Pearson Type VII Distributions on Spheres,” *Communications in Statistics – Theory and Methods*, 31, 513–526.
- Stephens, M.A. (1963), “Random Walk on a Circle,” *Biometrika*, 50, 385–390.

Watson, G.S. (1983), *Statistics on Spheres*, New York: Wiley.

Table 1. Parameter Values Maximising the Log-Likelihood and Corresponding Maximised Log-Likelihood Values for the Full Distribution Family and Three Sub-Families for the Mallard Data.

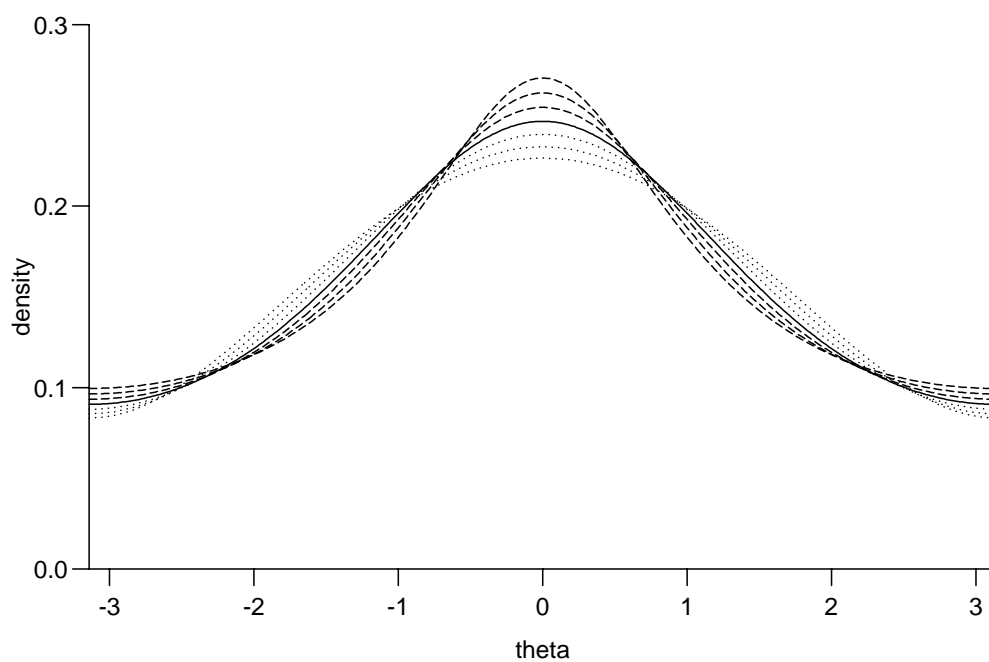
<i>Distribution</i>	$\psi$	$\mu$	$\kappa$	<i>Log-Likelihood</i>
full family	-0.35	-0.81	1.84	-869.36
von Mises	0	-0.80	2.12	-877.46
cardioid	1	-0.74	2.76	-981.55
wrapped Cauchy	-1	-0.82	1.50	-889.42

Table 2. Parameter Values Maximising the Log-Likelihood and Corresponding Maximised Log-Likelihood Values for Various Models Mixed With the Circular Uniform for the Ant Data. The columns headed  $\psi$  or  $\alpha$  and  $\kappa$  or  $\rho$  give  $\psi$  and  $\kappa$  for our family of distributions and  $\alpha$  and  $\rho$  for the wrapped symmetric stable distributions.

<i>Distribution</i>	$\psi$ or $\alpha$	$\mu$	$\kappa$ or $\rho$	$p$	<i>Log-Likelihood</i>
full family	-0.10	-3.10	5.01	0.67	-918.72
von Mises	0	-3.10	7.54	0.66	-918.80
cardioid	1	3.07	3.80	0.85	-1122.85
wrapped Cauchy	-1	-3.10	1.86	0.86	-923.77
wrapped symmetric stable	1.95	-3.10	0.93	0.67	-918.89
wrapped normal	2	-3.10	0.93	0.66	-918.90

Figure 1. Density (2) for  $\psi = -3/2, -1, -1/2$  (dashed),  $\psi = 0$  (solid), and  $\psi = 1/2, 1, 3/2$  (dotted), in Order of Decreasing Height at 0, for (a)  $\kappa = 1/2$ , (b)  $\kappa = 2$

[Figure 1(a)]





[Figure 1(b)]

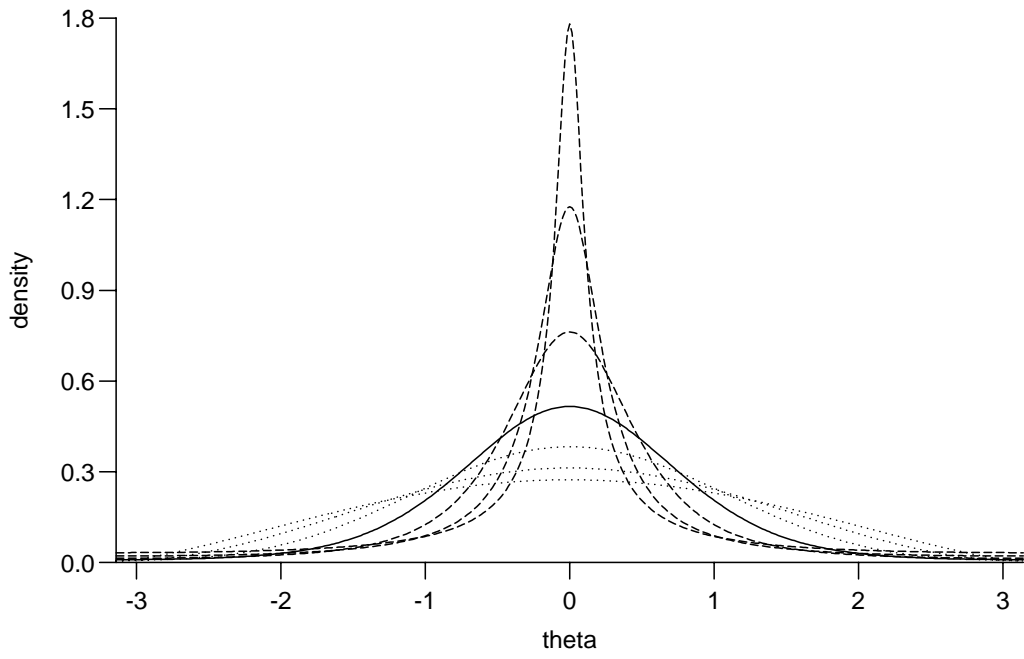


Figure 2. The Circular Variance as a Function of  $\kappa$  and  $\psi$

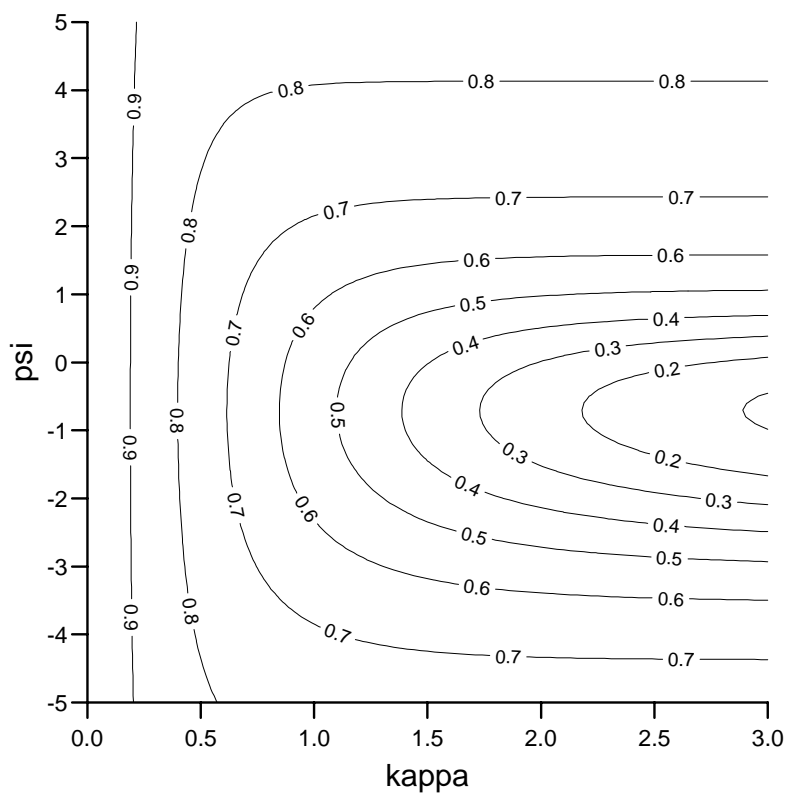


Figure 3. The Circular Kurtosis (6) as a Function of  $\kappa$  and  $\psi$

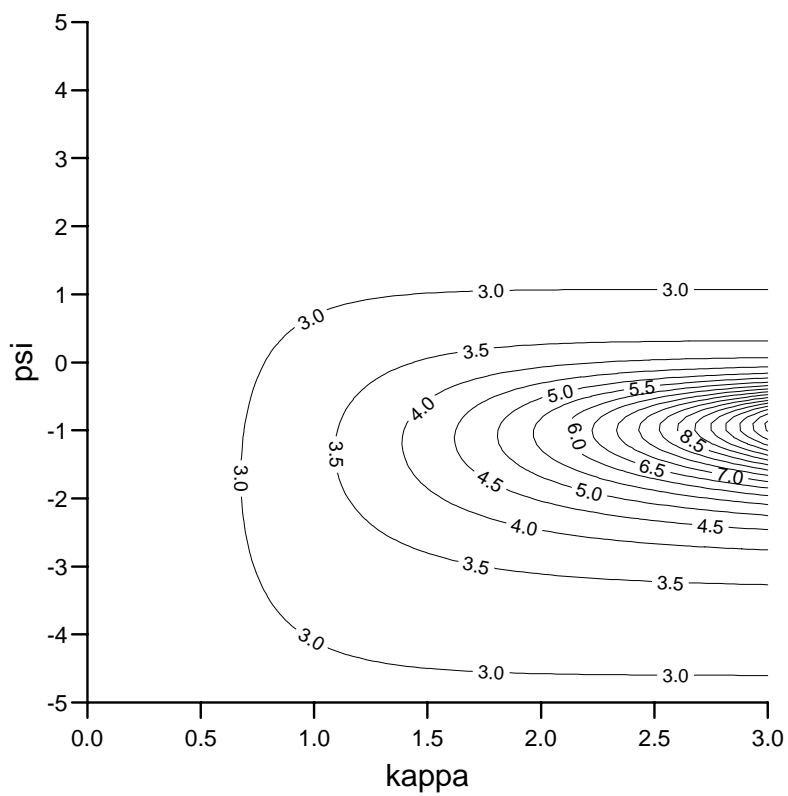


Figure 4. One Thousand Times the  $L_2$  Distance (7), Truncated After  $p_{\max} = 20$ , Between (2) and the Wrapped Normal Distribution as a Function of  $\kappa$  and  $\psi$

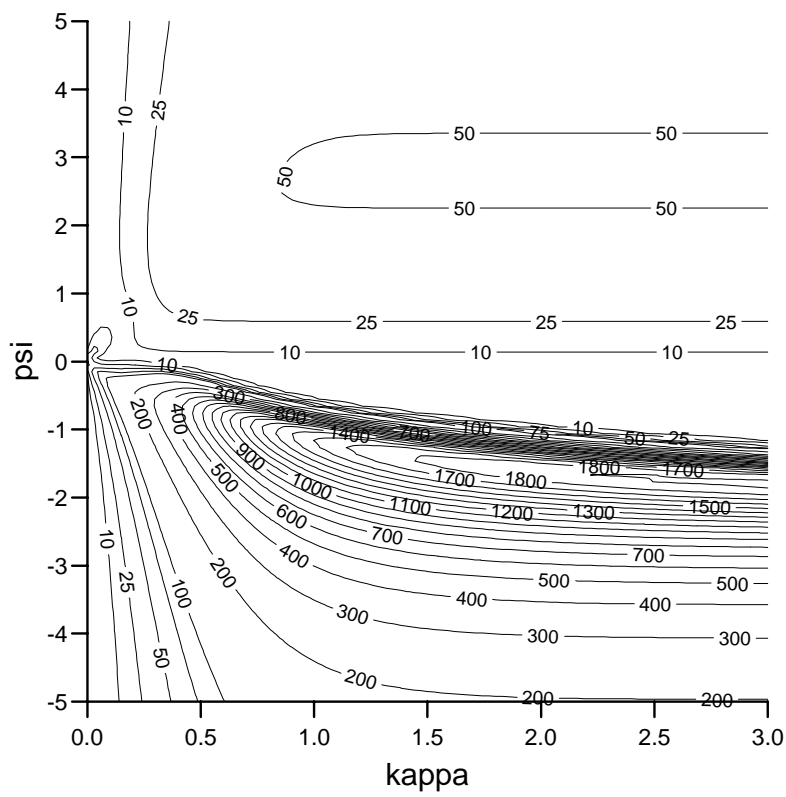


Figure 5. One Thousand Times the  $L_2$  Distance (8), Truncated After  $p_{\max} = 20$ , Between (2) and the Wrapped Symmetric Stable Distribution as a Function of  $\kappa$  and  $\psi \leq 0$

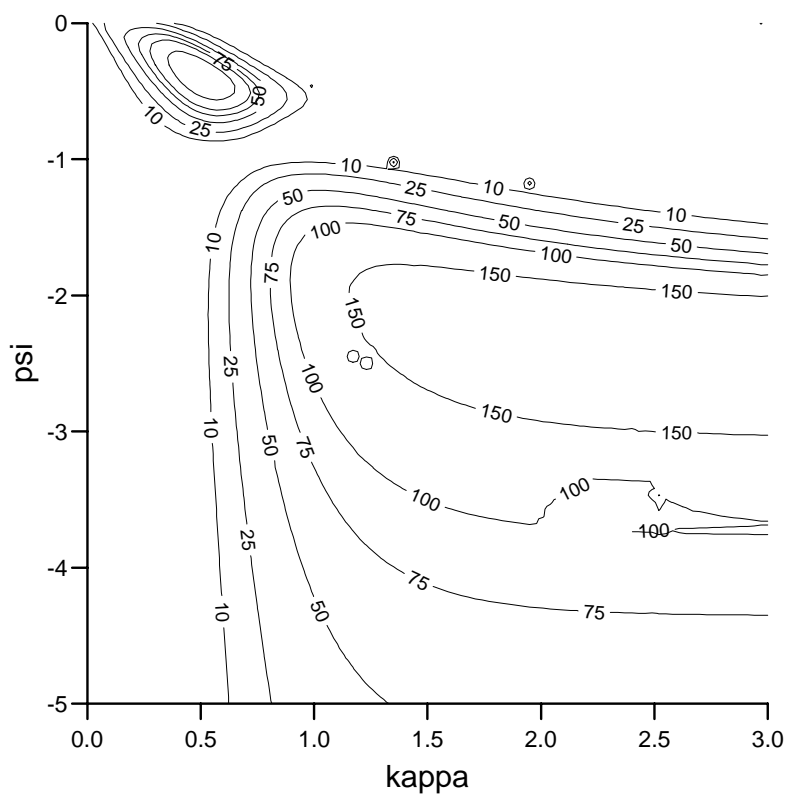


Figure 6. Asymptotic Correlation Between  $\hat{\kappa}$  and  $\hat{\psi}$  as a Function of  $\kappa$  and  $\psi$

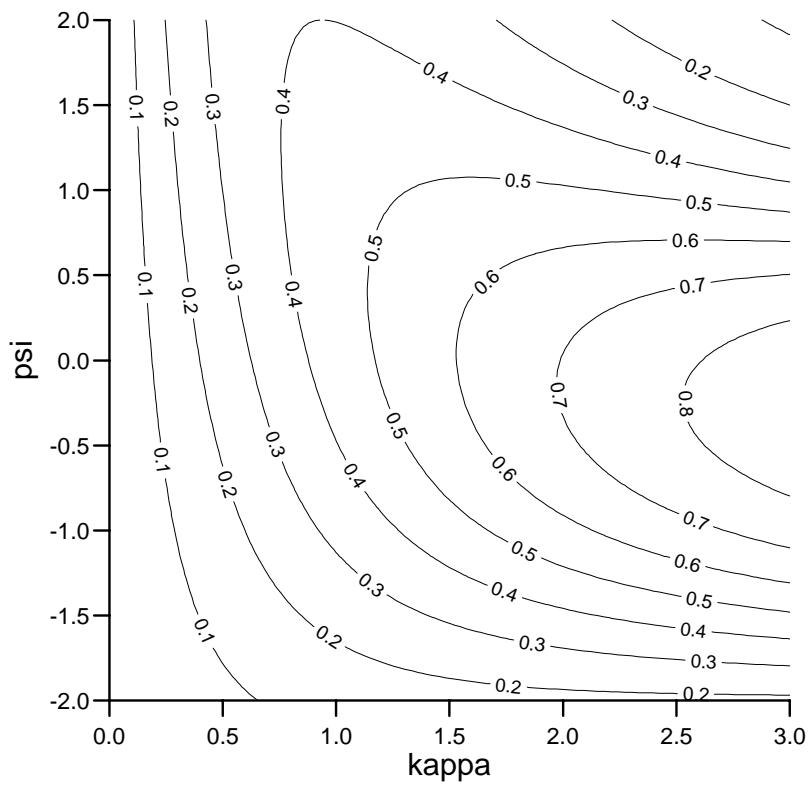


Figure 7. Histogram of the Mallard Data Together With Maximum Likelihood Fits of the Full Family of Distributions (Solid Curve) and the Von Mises Distribution (Dashed Curve). The data and densities are plotted on approximately  $(\hat{\mu} - \pi, \hat{\mu} + \pi)$

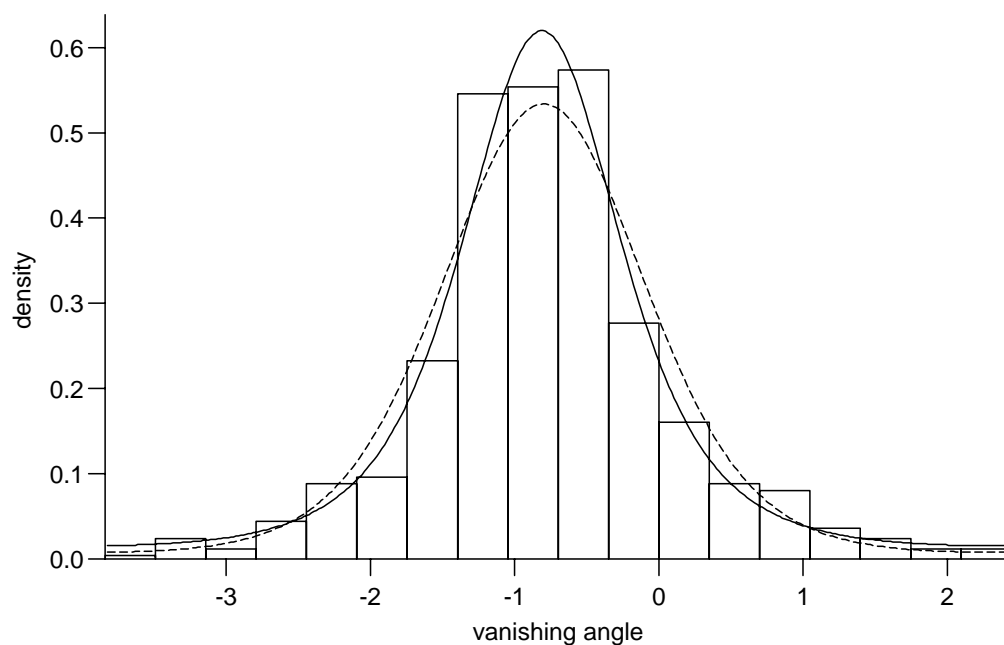


Figure 8. Histogram of the Ant Data Together With Maximum Likelihood Fits of the Full Family of Distributions (Solid Curve), the Von Mises/Uniform Mixture (Dashed Curve) and the Wrapped Normal/Uniform Mixture (Dotted Curve). The latter two curves are essentially coincident across most of the figure. The data and densities are plotted on approximately  $(\hat{\mu} - \pi, \hat{\mu} + \pi)$

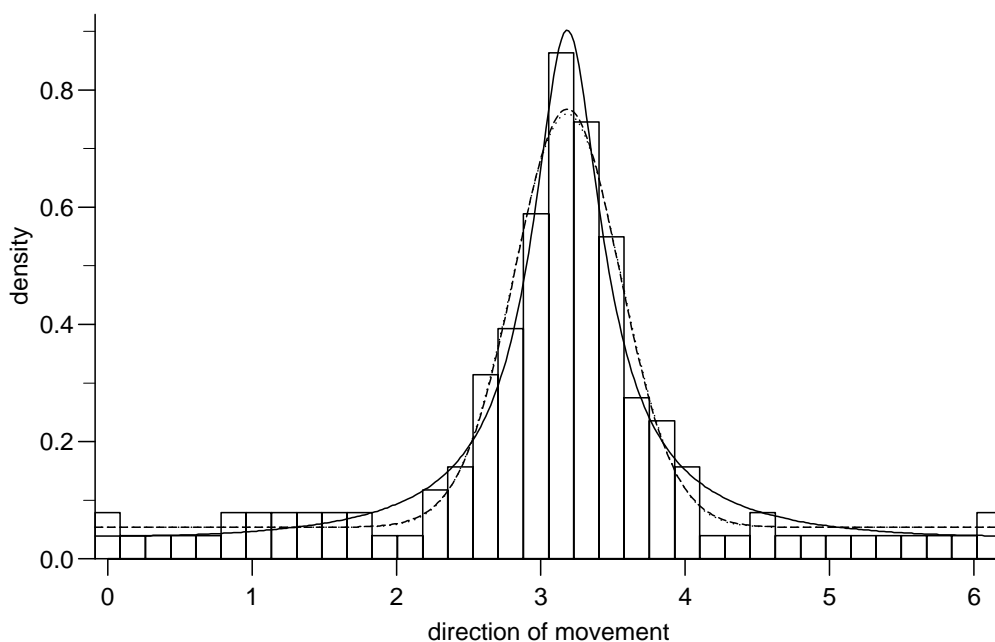
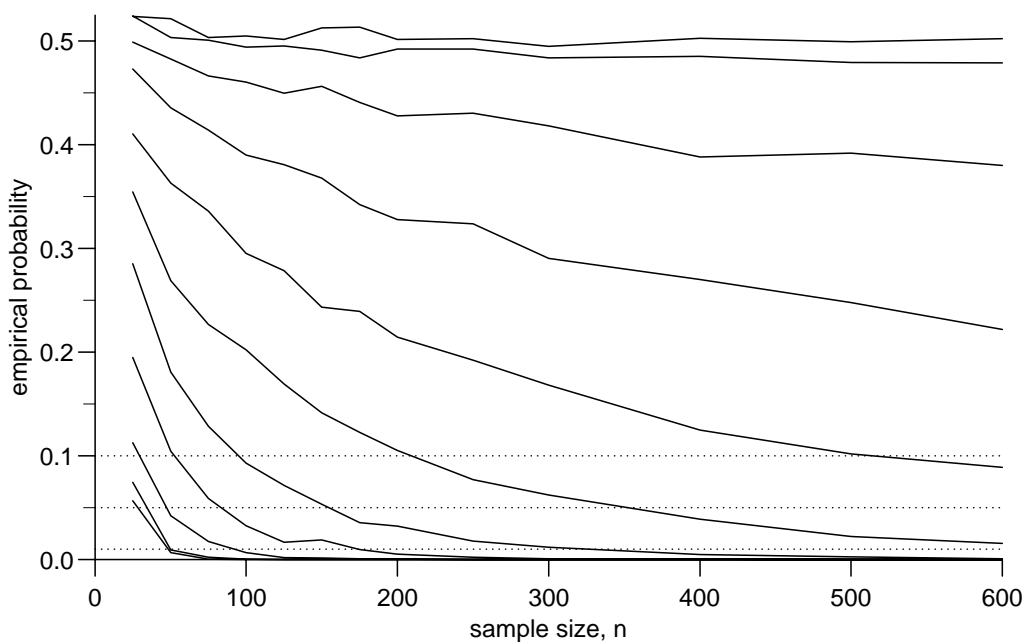


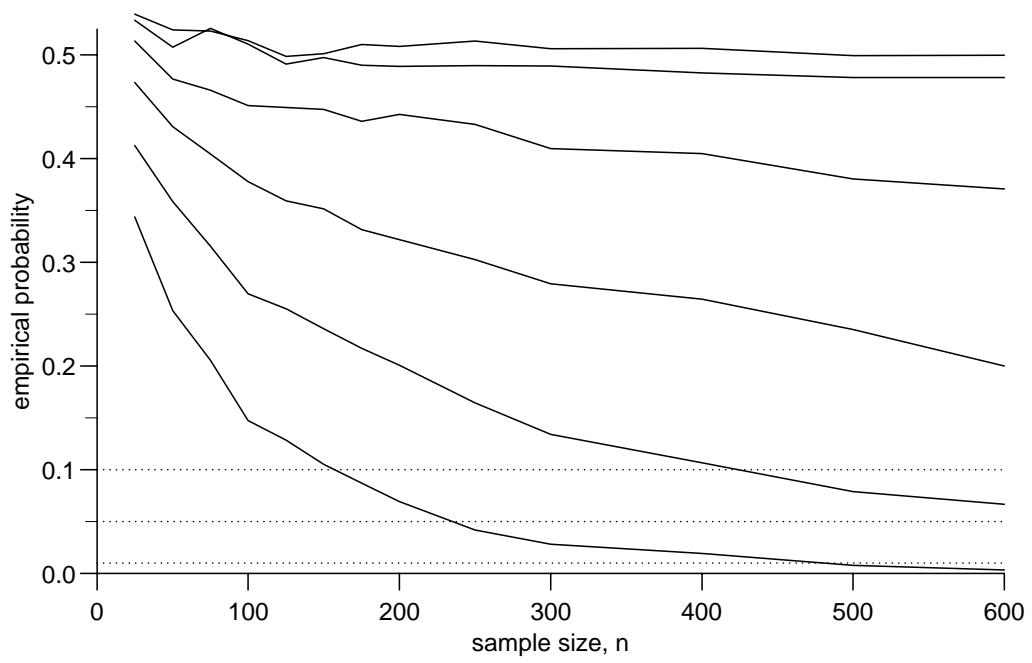


Figure 9. Estimated Values of  $p(\rho)$  as Functions of  $n$ . Plots correspond to: (a) von Mises versus wrapped Cauchy; (b) cardioid versus von Mises; (c) cardioid versus wrapped Cauchy. From the topmost downwards, the solid lines correspond to increasing values of  $\rho$ ;  $\rho$  takes values 0.05, 0.1(0.1)0.9, 0.95 in frame (a) and 0.05, 0.1(0.1)0.5 in frames (b) and (c). The dotted lines are at 0.1, 0.05 and 0.01

[Figure 9(a)]



[Figure 9(b)]



[Figure 9(c)]

