An extended family of circular distributions related to wrapped Cauchy distributions via Brownian motion

SHOGO KATO [1*] and M.C. JONES [2]

[1] *Institute of Statistical Mathematics, 10-3 Midori-Cho, Tachikawa, Tokyo 190-8562, Japan. E-mail: skato@ism.ac.jp*

[2] *Department of Mathematics & Statistics, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK E-mail: m.c.jones@open.ac.uk*

We introduce a four-parameter extended family of distributions related to the wrapped Cauchy distribution on the circle. The proposed family can be derived by altering the settings of a problem in Brownian motion which generates the wrapped Cauchy. The densities of this family have a closed form and can be symmetric or asymmetric depending on the choice of the parameters. Trigonometric moments are available, and they are shown to have a simple form. Further tractable properties of the model are obtained, many by utilising the trigonometric moments. Other topics related to the model, including alternative derivations, Möbius transformation and random variate generation, are considered. Discussion of the symmetric submodels is given. Finally, generalisation to a family of distributions on the sphere is briefly made.

Keywords: asymmetry; circular Cauchy distribution; directional statistics; four-parameter distribution; trigonometric moments.

Running Title: Circular distributions via Brownian motion

1 Introduction

As a unimodal and symmetric model on the circle, the wrapped Cauchy or circular Cauchy distribution has played an important role in directional statistics. It has density

$$f(\theta;\mu,\rho)=\frac{1}{2\pi}\frac{1-\rho^2}{1+\rho^2-2\rho\cos(\theta-\mu)},\quad -\pi\le\theta<\pi, \tag{1}$$

where $\mu\,(\in[-\pi,\pi))$ is a location parameter and $\rho\,(\in[0,1))$ controls the concentration of the model. This distribution has some desirable mathematical features as discussed in Kent and Tyler (1988) and McCullagh (1996). We write $\Theta\sim WC(\mu,\rho)$ if Θ has density (1).

In modelling symmetric circular data, the wrapped Cauchy distribution could be one choice as well as other familiar symmetric models such as the von Mises and wrapped normal distributions. (See also Jones and Pewsey, 2005.) In reality, however, it is not so common to find such data in fields of application. In many cases the data of interest are asymmetrically distributed and therefore probability distributions having asymmetric densities are desired.

Construction of a tractable circular model with an asymmetric shape has been a problem in statistics of circular data. To tackle this problem, some asymmetric extensions of well-known circular models have been proposed in the literature. Maksimov (1967), Yfantis and Borgman (1982) and Gatto and Jammalamadaka (2007) discussed an extension of the von Mises distribution generated through maximisation of Shannon's entropy with restrictions on certain trigonometric moments. Batschelet (1981) proposed a mathematical method of skewing circular distributions that has seen renewed interest very recently. Pewsey (2008) presented a four-parameter family of distributions on the circle by wrapping the stable distribution, which is asymmetric, onto the circle. Recent work by Kato and Jones (2010) proposed a family of distributions arising from the Möbius transformation which includes the von Mises and wrapped Cauchy distributions. Unlike familiar symmetric distributions, it is often difficult to deal with skew models in statistical analysis. This difficulty is partly due to the lack of some mathematical properties which many of the well-known symmetric models have. For example, existing asymmetric models often have complex normalising constants and trigonometric moments, which could cause trouble in analysis.

In this paper, we provide a four-parameter extended family of circular distributions based on the wrapped Cauchy distribution by applying Brownian motion which, to our knowledge, has not been used to propose a skew distribution on the circle. The four parameters enable the model to describe not only symmetric shapes but also asymmetric ones. An advantage of the proposed model is its mathematical tractability. For instance, it has a simple normalising constant and trigonometric moments. The property that the model can be derived from the Bayesian approach enables us to generate random samples using Markov Chain Monte Carlo. In addition to these, the distribution also has other mathematical features which bear advantages in statistical analysis. The current proposal is more tractable than the family discussed by Kato and Jones (2010), but is complementary to it in the sense that the latter has other advantages (particularly some associated with Möbius transformation).

The subsequent sections are organised as follows. In Section 2 we make the main proposal of this paper. The derivation of the proposed model is given, and the probability density function and probabilities of intervals under the density are discussed. Also, special cases of the model are briefly considered. Section 3 concerns the shapes of the density. Conditions for symmetry and for unimodality are explored, and the interpretation of the parameters is discussed through pictures of the density. In Section 4 we discuss the trigonometric moments and problems related to them. It is shown that the trigonometric moments can be expressed in a simple form. The mean direction, mean resultant length and skewness of the model are also considered. Applying the first trigonometric moment, the path of the expected exit points of a Brownian particle is obtained. Some other topics concerning our model are provided in Section 5. Apart from the derivation given in Section 2, there are other methods to generate the family, which are discussed in that section. In addition, we study the conformal invariance properties of the distribution and random variate generation using the Markov Chain Monte Carlo method. In Section 6 we investigate some properties of the symmetric cases of the proposed model. These submodels have some properties which the general family does not have. Finally, in Section 7, generalisation to a family of distributions on the sphere is made and its properties are briefly discussed. Appendices contain selected proofs.

2 A Family of Distributions on the Circle

2.1 Definition

It is a well known fact that the wrapped Cauchy distribution can be generated as the distribution of the position of a Brownian particle which exits the unit disc in the two-dimensional plane. (See, for example, Durrett (1984, Section 1.10).) The increments of Brownian motion are assumed to follow the multivariate normal distribution on $\mathbb{R}^n$. In addition to the wrapped Cauchy circular model, some distributions on certain manifolds are derived from, or have relationships with, Brownian motion. The Cauchy distribution on the real line is the distribution of the position where a Brownian particle exits the upper half plane. Considering the hitting time of the particle, one can obtain the inverse Gaussian distribution on the positive half-line. Kato (2009) proposed a distribution for a pair of unit vectors by recording points where a Brownian particle hits circles with different radii.

By applying Brownian motion, we provide a family of asymmetric distributions on the circle which includes the wrapped Cauchy distribution as a special case. The proposed model is defined as follows.

Definition 1. *Let $\{B_t\,;\, t \geq 0\}$ be $\mathbb{R}^2$-valued Brownian motion without drift starting at $B_0 = \rho_1(\cos\mu_1,\, \sin\mu_1)'$, where $0 \leq \rho_1 < 1$ and $-\pi \leq \mu_1 < \pi$. This Brownian particle will eventually hit the unit circle. Let τ_1 be the first time at which the particle exits the circle, i.e. $\tau_1 = \inf\{t\,;\, \|B_t\| = 1\}$. After leaving the unit circle, the particle will hit a circle with radius ρ_2^{-1} $(0 < \rho_2 < 1)$ first at the time τ_2, meaning $\tau_2 = \inf\{t\,;\, \|B_t\| = \rho_2^{-1}\}$. Then the proposed model is defined by the conditional distribution of B_{τ_1} given $B_{\tau_2} = \rho_2^{-1}(\cos\mu_2, \sin\mu_2)'$, where $-\pi \leq \mu_2 < \pi$.*

To put it another way, the proposed random vector, B_{τ_1} given B_{τ_2}, represents the position where a Brownian particle first hits the unit circle, given the future point at which the particle exits a circle with a larger radius. From the next subsection, we investigate some properties of the proposed model.

2.2 Probability density function

One feature of the proposed model is that it has a closed form of density with simple normalising constant. It is given in the following theorem. See Appendix A for the proof.

Theorem 1. *Let Θ be a $[-\pi, \pi)$-valued random variable defined as $B_{\tau_1} = (\cos\Theta, \sin\Theta)'$. Then the conditional density of Θ given $B_{\tau_2} = \rho_2^{-1}(\cos\mu_2, \sin\mu_2)'$ is given by*

$$f(\theta) = C\left[\left\{1+\rho_1^2 - 2\rho_1\cos(\theta-\mu_1)\right\}\left\{1+\rho_2^2 - 2\rho_2\cos(\theta-\mu_2)\right\}\right]^{-1}, \quad -\pi \le \theta < \pi, \tag{2}$$

where $-\pi \le \mu_1, \mu_2 < \pi,\ 0 \le \rho_1, \rho_2 < 1$, and the normalising constant C is

$$C = \frac{(1-\rho_1^2)(1-\rho_2^2)\{1+\rho_1^2\rho_2^2 - 2\rho_1\rho_2\cos(\mu_1-\mu_2)\}}{2\pi(1-\rho_1^2\rho_2^2)}.$$

For convenience, write $\Theta \sim EWC(\mu_1, \mu_2, \rho_1, \rho_2)$ if a random variable Θ has the density (2) ('E' for 'Extended'). Note that the density does not involve any infinite sums or special functions.

It is easy to see that distribution (2) reduces to the wrapped Cauchy (1) when either ρ_1 or ρ_2 is equal to zero.

Quite often, it is advantageous to express the random variable and parameters of the proposed model in terms of complex numbers rather than real numbers. Define a random vector by $Z = e^{i\Theta}$ where $\Theta \sim EWC(\mu_1, \mu_2, \rho_1, \rho_2)$. Then Z has density

$$f(z) = \frac{1}{2\pi}\frac{|1-\phi_1\overline{\phi_2}|^2}{1-|\phi_1\overline{\phi_2}|^2}\frac{1-|\phi_1|^2}{|z-\phi_1|^2}\frac{1-|\phi_2|^2}{|z-\phi_2|^2}, \quad z \in \partial D, \tag{3}$$

with respect to arc length on the circle, where $\phi_1 = \rho_1 e^{i\mu_1},\ \phi_2 = \rho_2 e^{i\mu_2}$ and $\partial D = \{z \in \mathbb{C}\,;\,|z| = 1\}$. It is clear that the parameters in this formulation, ϕ_1 and ϕ_2, take values on the unit disc in the complex plane denoted by $D = \{z \in \mathbb{C}\,;\,|z| < 1\}$. For brevity, we denote the distribution (3) by $EC^*(\phi_1, \phi_2)$. Also, as in McCullagh (1996), write $Z \sim C^*(\phi_1)$ if Z follows distribution (3) with $\phi_2 = 0$.

We remark that the above change-of-variable and reparametrisation do not actually change the distribution. In later sections, we utilise either of the representations (2) or (3), whichever is the more convenient.

2.3 Probabilities

As well as the density of the proposed model, the probabilities of intervals under the density can also be expressed without using infinite sums or special functions.

Theorem 2. *Let a random variable Θ follow the $EWC(\mu_1, \mu_2, \rho_1, \rho_2)$ distribution. If $\rho_1 \ne \rho_2$ or $\mu_1 \ne \mu_2$, then the probability of intervals under the density of Θ is given by*

$$
\begin{aligned}
&P(a < \Theta \le b) \\
&= \int_a^b C\left[\left\{1+\rho_1^2-2\rho_1\cos(\theta-\mu_1)\right\}\left\{1+\rho_2^2-2\rho_2\cos(\theta-\mu_2)\right\}\right]^{-1} d\theta \\
&= \frac{C}{D}\Bigg\{\rho_1\rho_2\sin(\mu_1-\mu_2)\left[\log\left\{\frac{1+\rho_2^2-2\rho_2\cos(\theta-\mu_2)}{1+\rho_1^2-2\rho_1\cos(\theta-\mu_1)}\right\}\right]_a^b \\
&\quad + \frac{2\rho_1\{\rho_1(1+\rho_2^2)-\rho_2(1+\rho_1^2)\cos(\mu_1-\mu_2)\}}{1-\rho_1^2}\left(\left[\arctan\left\{\frac{1+\rho_1}{1-\rho_1}\tan\left(\frac{\theta-\mu_1}{2}\right)\right\}\right]_a^b + A_{\mu_1}\right) \\
&\quad + \frac{2\rho_2\{\rho_2(1+\rho_1^2)-\rho_1(1+\rho_2^2)\cos(\mu_1-\mu_2)\}}{1-\rho_2^2}\left(\left[\arctan\left\{\frac{1+\rho_2}{1-\rho_2}\tan\left(\frac{\theta-\mu_2}{2}\right)\right\}\right]_a^b + A_{\mu_2}\right)\Bigg\},
\end{aligned}
$$

where C is defined as in Theorem 1, $-\pi \le a < b < \pi$,

$$D = (\rho_1^2+\rho_2^2)(1+\rho_1^2\rho_2^2) - 2\rho_1\rho_2(1+\rho_1^2)(1+\rho_2^2)\cos(\mu_1-\mu_2) + 4\rho_1^2\rho_2^2\cos^2(\mu_1-\mu_2),$$

and

$$A_\mu = \begin{cases} \pi, & \tan\left\{\frac{1}{2}(a-\mu)\right\} > \tan\left\{\frac{1}{2}(b-\mu)\right\}, \\ 0, & \text{otherwise.} \end{cases}$$

If $\rho_1 = \rho_2$ and $\mu_1 = \mu_2$, then

$$
\begin{aligned}
P(a < \Theta \le b) &= \int_a^b C\left\{1+\rho_1^2-2\rho_1\cos(\theta-\mu_1)\right\}^{-2} d\theta \\
&= \frac{2C}{(1-\rho_1^2)^2}\Bigg\{\left[\frac{\rho_1\sin(\theta-\mu_1)}{1+\rho_1^2-2\rho_1\cos(\theta-\mu_1)}\right]_a^b \\
&\quad + \frac{1+\rho_1^2}{1-\rho_1^2}\left(\left[\arctan\left\{\frac{1+\rho_1}{1-\rho_1}\tan\left(\frac{\theta-\mu_1}{2}\right)\right\}\right]_a^b + A_{\mu_1}\right)\Bigg\}.
\end{aligned}
$$

Proof. The result is straightforward from equations (2.553.3), (2.554.3) and (2.559.2) of Gradshteyn and Ryzhik (1994). □

2.4 Special cases

The proposed family (2) contains some known distributions as special cases.

Case 1: The wrapped Cauchy distribution. As mentioned in Section 2.2, model (2) becomes the wrapped Cauchy $WC(\mu_1, \rho_1)$ if $\rho_2 = 0$. Similarly, the model is $WC(\mu_2, \rho_2)$ when $\rho_1 = 0$.

Case 2: A special case of the Jones and Pewsey (2005) family. If $\rho_1 = \rho_2 \, (\equiv \rho)$ and $\mu_1 = \mu_2 \, (\equiv \mu)$, then density (2) reduces to

$$f(\theta) = \frac{1}{2\pi}\frac{1-\rho^2}{1+\rho^2}\left(\frac{1-\rho^2}{1+\rho^2-2\rho\cos(\theta-\mu)}\right)^2, \quad -\pi \le \theta < \pi. \tag{4}$$

This submodel corresponds to a special case of the family presented by Jones and Pewsey (2005). Their model has density

$$f_{JP}(\theta) = \frac{\{\cosh(\kappa\psi) + \sinh(\kappa\psi)\cos(\theta - \mu)\}^{1/\psi}}{2\pi P_{1/\psi}(\cosh(\kappa\psi))}, \quad -\pi \le \theta < \pi,$$

where $\kappa \ge 0$, $\psi \in \mathbb{R}$, $-\pi \le \mu < \pi$, and $P_{1/\psi}(z)$ is the associated Legendre function of the first kind of degree $1/\psi$ and order 0 (Gradshteyn and Ryzhik, 1994, Sections 8.7, 8.8). Our submodel is equivalent to their family with $\psi = -1/2$ and $\kappa = 2\log((1+\rho)/(1-\rho))$. In this case the associated Legendre function simplifies to $P_{-2}(z) = P_1(z) = z$. (See equations (8.2.1) and (8.4.3) of Abramowitz and Stegun (1970).) This fairly heavy-tailed circular distribution is shown by Jones and Pewsey to tend, suitably normalised, to the t distribution on $\sqrt{3}$ degrees of freedom as ($\rho \to 1$ and hence) $\kappa \to \infty$!

Case 3: One-point distribution. As ρ_1 (ρ_2) tends to one, the model converges to a point distribution with singularity at $\theta = \mu_1$ ($\theta = \mu_2$). Normalising by a scale factor of $(1-\rho)/\sqrt{\rho}$, the limiting version of the wrapped Cauchy distribution as $\rho \to 1$ is the ordinary Cauchy distribution. That argument can be extended to show that the Cauchy limiting distribution also arises in this case.

Case 4: Two-point distribution. Assume that $\rho_1 = \rho_2 (\equiv \rho)$ and $\mu_1 \ne \mu_2$. When ρ goes to one, the model (2) converges to the distribution of a random variable which takes values on μ_1 or μ_2 with probability 0.5. If $\mu_1 < \mu_2$ and ρ tends to one, $\{\sqrt{\rho}/(1-\rho)\}(\Theta - \mu_1)$ converges to $\frac{1}{2}C(0,1) + \frac{1}{2}I$, where $C(0,1)$ is the standard Cauchy and I has the distribution function $F(x) = \lim_{\lambda\to\infty}\int_{-\infty}^{x} 1/[\pi\{1+(t-\lambda)^2\}]dt$. Roughly speaking, this limiting distribution is a 50:50 mixture of the standard Cauchy and a point distribution with singularity at $X = \infty$. Similarly, one can show that if $\mu_1 > \mu_2$ and ρ tends to one, $\{\sqrt{\rho}/(1-\rho)\}(\Theta - \mu_1)$ converges to a 50:50 mixture of the standard Cauchy and a point distribution with singularity at $X = -\infty$. The limiting distribution of $\{\sqrt{\rho}/(1-\rho)\}(\Theta - \mu_2)$ can be discussed in a similar manner.

Case 5: The circular uniform distribution. When $\rho_1 = \rho_2 = 0$, the distribution reduces to the circular uniform distribution.

3 Shapes of Probability Density Function

3.1 Conditions for symmetry

As briefly stated in Section 1, the proposed model can be symmetric or asymmetric depending on the values of the parameters. The condition for symmetry is clearly written out in the following theorem.

Theorem 3. *The density (2) is symmetric if and only if* $\rho_1 = \rho_2$, $\rho_j = 0$ *or* $\mu_1 = \mu_2 + (j-1)\pi$ $(j = 1, 2)$.

See Appendix B for the proof. The three-parameter symmetric submodels arising under the above conditions will be discussed in Section 6.

3.2 Conditions for unimodality

Turning to conditions for unimodality, it is possible to express these in terms of an inequality. The process to obtain the inequality is similar to that in Kato and Jones (2010, Section 2.5). In the following discussion, take $\mu_2 = 0$ without loss of generality. First we calculate the first derivative of density (2) with respect to θ, which is given by

$$\frac{d}{d\theta} f(\theta) \propto \frac{[\rho_1 \sin(\theta - \mu_1)(1 + \rho_2^2 - 2\rho_2 \cos\theta) + \rho_2 \sin\theta\{1 + \rho_1^2 - 2\rho_1 \cos(\theta - \mu_1)\}]}{\{1 + \rho_1^2 - 2\rho_1 \cos(\theta - \mu_1)\}^2 (1 + \rho_2^2 - 2\rho_2 \cos\theta)^2}.$$

Then it follows that the extrema of the density are obtainable as solutions of the following equation:

$$a_0 + a_1 \cos\theta + a_2 \sin\theta + a_3 \cos\theta \sin\theta + a_4 \cos^2\theta = 0, \tag{5}$$

where

$$\begin{aligned} a_0 &= 2\rho_1\rho_2 \sin\mu_1, \quad a_1 = \rho_1(1+\rho_2^2)\sin\mu_1, \quad a_2 = -\rho_1(1+\rho_2^2)\cos\mu_1 - \rho_2(1+\rho_1^2), \\ a_3 &= 4\rho_1\rho_2 \cos\mu_1, \quad a_4 = -4\rho_1\rho_2 \sin\mu_1. \end{aligned}$$

Following Yfantis and Borgman (1982) and Kato and Jones (2010, Section 2.5), put $x = \tan(\theta/2)$ so that $\cos\theta = (1-x^2)/(1+x^2)$, $\sin\theta = 2x/(1+x^2)$. It follows that equation (5) can be expressed as a quartic equation in x whose discriminant D (Uspensky, 1948) can be written down in terms of $a_0, ..., a_4$ as in (6) of Kato and Jones (2010). The quartic equation has four real roots or four complex ones if $D > 0$, and two real roots and two complex ones if $D < 0$. Therefore the distribution is bimodal when $D > 0$ and unimodal when $D < 0$. Since a_j $(0 \le j \le 4)$ are functions of μ_1, ρ_1 and ρ_2, we can write the conditions for unimodality as a function of these three parameters. For general μ_2, it is easy to see that the discriminant is expressed in terms of $\mu_1 - \mu_2$, ρ_1 and ρ_2.

Figure 1 exhibits areas of positivity (bimodality of density) and negativity (unimodality of density) of the discriminant D for two pairs of parameters. Figure 1(a) suggests that density (2) with $\rho_1 = 0.5$ becomes unimodal for most cases when $-\pi/2 \le \mu_1 \le \pi/2$. It seems that bimodality is most likely to occur if $\pi/2 \le \mu_1 \le 3\pi/2$. The other frame of Figure 1 seems to show that distribution (2) with $\mu_2 = 2\pi/3$ is bimodal if ρ_1 or ρ_2 is sufficiently large and ρ_2/ρ_1 takes a value close to 1. In particular, when $\rho_1 = \rho_2$, the range of parameters corresponding to unimodality seems less wide than that when $\rho_1 = a\rho_2$ with $a \ne 1$. As will be shown in Section 6.2, for $\mu_1 = 2\pi/3$ and $\rho_1 = \rho_2$, the density becomes unimodal if $\rho_1 \le 2 - \sqrt{3} \simeq 0.268$ and bimodal otherwise.

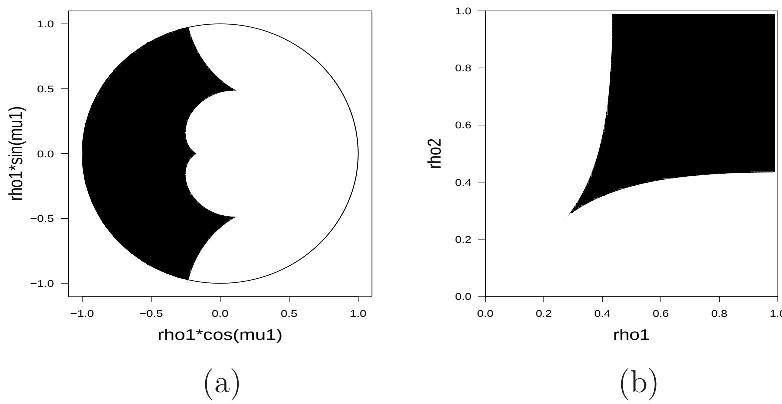


Figure 1: Discriminant D as functions of: (a) $(\rho_1 \cos \mu_1, \rho_1 \sin \mu_1)$ with $\rho_2 = 0.5$, and (b) (ρ_1, ρ_2) with $\mu_1 = 2\pi/3$. Positive D, corresponding to bimodality, is displayed in black, negative D, corresponding to unimodality, is exhibited in white.

3.3 Pictures of density

We plot density (2) for selected values of the parameters in Figure 2. Confirming the results in the previous two subsections, this figure shows that the density can be symmetric or asymmetric, unimodal or bimodal, depending on the choice of the parameters. Figure 2(a), displaying the density for fixed values of μ_2, ρ_1 and ρ_2, suggests that the density is symmetric if $\mu_1 = \mu_2$ or $\mu_1 = \mu_2 + \pi$ and asymmetric otherwise. Figure 2(b) exhibits how the shape of density (2) changes as ρ_1 increases. It is clear from this frame that the greater the value of ρ_1, the greater the concentration of the density. The frame also implies that the absolute value of the circular skewness of the density with fixed μ_1 and μ_2 takes a large value if ρ_1/ρ_2 is close to 1. Some symmetric cases of the density are shown in Figure 2(c). It appears to be that the density is unimodal when $\mu_1 = \mu_2$, whereas bimodality can occur when $\mu_2 = \mu_1 + \pi$. Figure 2(d) is a case in which two parameters, ρ_1 and ρ_2, are fixed to be equal (to 0.5), also implying that density (2) is symmetric. We will focus on this submodel in Section 6.2.

4 Trigonometric Moments and Related Problems

4.1 Trigonometric moments

It is often the case that asymmetric distributions on the circle have complicated trigonometric moments or Fourier coefficients. One feature of our asymmetric model, however, is the relative simplicity of its trigonometric moments. The expression for the moments is greatly simplified if the variables and parameters are represented in terms of complex numbers.

Theorem 4. *Let a random variable Z have the $EC^*(\phi_1, \phi_2)$ distribution. Then the nth trigonometric moment of Z is given by*

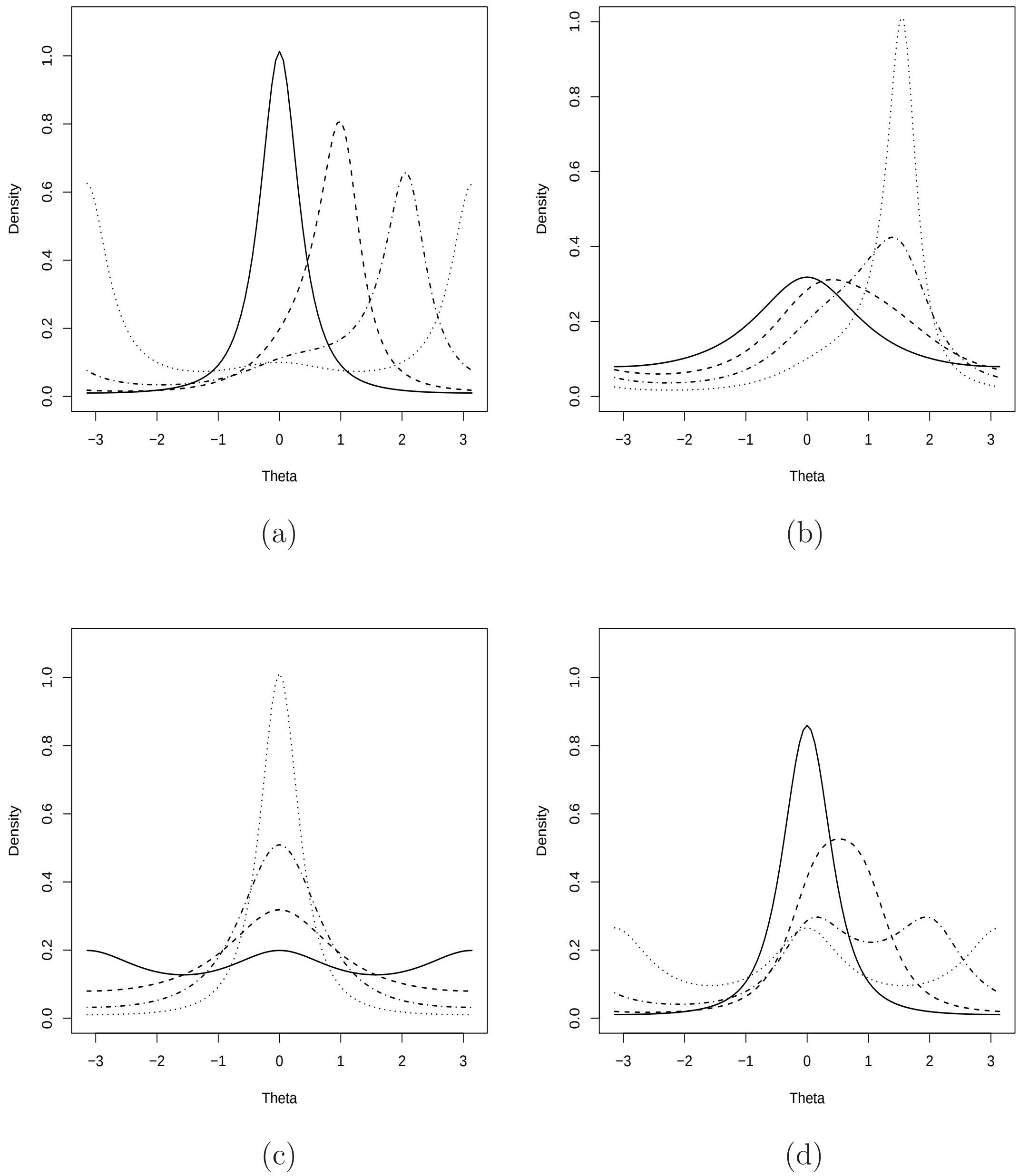
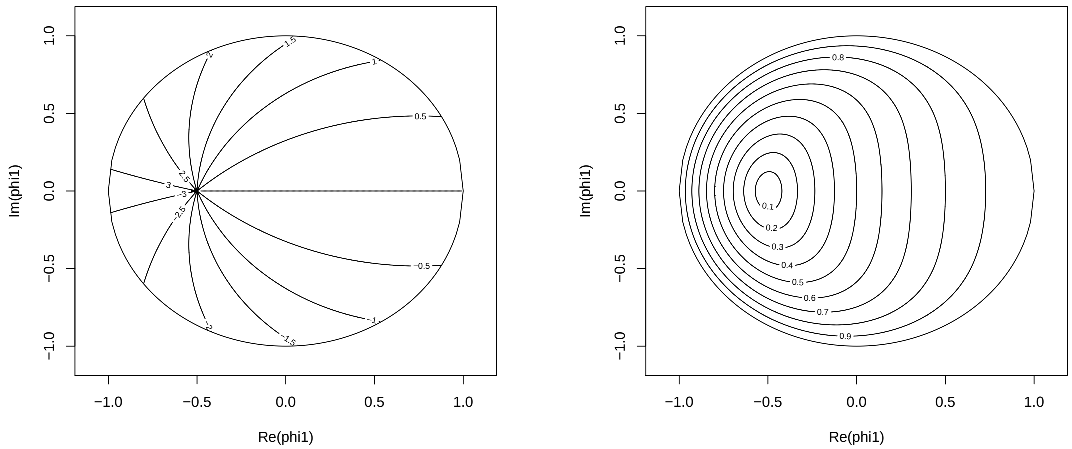


Figure 2: Density (2) for $\mu_2 = 0$ and (a) $\rho_1 = 2/3,\ \rho_2 = 1/3$ and: $\mu_1 = 0$ (solid), $\pi/3$ (dashed), $2\pi/3$ (dot-dashed) and π (dotted), (b) $\mu_1 = \pi/2,\ \rho_2 = 1/3$ and: $\rho_1 = 0$ (solid), $1/4$ (dashed), $1/2$ (dot-dashed) and $3/4$ (dotted), (c) $\rho_2 = 1/3$ and: $(\mu_1, \rho_1) = (\pi, 1/3)$ (solid), $(0, 0)$ (dashed), $(0, 1/3)$ (dot-dashed) and $(0, 2/3)$ (dotted), and (d) $\rho_1 = \rho_2 = 0.5$ and: $\mu_1 = 0$ (solid), $\pi/3$ (dashed), $2\pi/3$ (dot-dashed) and π (dotted).



(a) (b)

Figure 3: Plot of (a) mean direction and (b) mean resultant length of a random variable having the distribution $EC^*(\phi_1, 0.5)$, as a function of ϕ_1.

$$E(Z^n) = \begin{cases} \dfrac{(1-|\phi_2|^2)(1-\overline{\phi_1}\phi_2)\,\phi_1^{n+1} - (1-|\phi_1|^2)(1-\phi_1\overline{\phi_2})\,\phi_2^{n+1}}{(\phi_1-\phi_2)(1-|\phi_1\overline{\phi_2}|^2)}, & \phi_1 \neq \phi_2, \\ \dfrac{1+n+(1-n)|\phi_1|^2}{1+|\phi_1|^2}\phi_1^n, & \phi_1 = \phi_2. \end{cases}$$

Proof. The nth trigonometric moment can be expressed as

$$\begin{aligned} E(Z^n) &= \frac{1}{2\pi}\frac{|1-\overline{\phi_1}\phi_2|^2}{1-|\phi_1\overline{\phi_2}|^2}\int_{\partial D} z^n \left(\prod_{j=1}^{2}\frac{1-|\phi_j|^2}{|z-\phi_j|^2}\right)|dz| \\ &= \frac{1}{2\pi i}\frac{|1-\overline{\phi_1}\phi_2|^2}{1-|\phi_1\overline{\phi_2}|^2}\int_{\partial D} z^{n+1}\left\{\prod_{j=1}^{2}\frac{1-|\phi_j|^2}{(z-\phi_j)(1-\overline{\phi_j}z)}\right\}dz \qquad (6) \end{aligned}$$

If $\phi_1 \neq \phi_2$, the integrand in (6) is holomorphic on the unit disc except at two poles of order 1, i.e., $z = \phi_j,\ j = 1, 2$. When $\phi_1 = \phi_2$, the integrand in (6) has a single pole of order 2 at $z = \phi_1$. Hence, from the residue theorem (e.g., Rudin, 1987, Theorem 10.42), we obtain Theorem 4. □

4.2 Mean direction and mean resultant length

Let Z be distributed as $EC^*(\phi_1, \phi_2)$. From Theorem 4 it is easy to see that the first trigonometric moment for Z is simply expressed as

$$E(Z) = \frac{(1-|\phi_2|^2)\phi_1 + (1-|\phi_1|^2)\phi_2}{1-|\phi_1\phi_2|^2} \quad \text{for any } \phi_1 \text{ and } \phi_2. \qquad (7)$$

The following result provides the condition under which the first trigonometric moment is equal to zero.

Corollary 1. *The necessary and sufficient condition for* $E(Z) = 0$ *is given by* $\phi_1 = -\phi_2$.

Proof. It is straightforward to show the sufficient condition for $E(Z) = 0$. The necessary condition is proved as follows. Let $E(Z) = 0$ and $\phi_1 \neq -\phi_2$. Then, from Theorem 4, the following equation holds between the two parameters: $\phi_1\phi_2 = (\phi_1 + \phi_2)/\overline{(\phi_1 + \phi_2)}$. Taking the absolute values of both sides of the above equation, we have $|\phi_1\phi_2| = 1$. Since $|\phi_j| < 1$ $(j = 1, 2)$, there does not exist ϕ_1 and ϕ_2 such that $|\phi_1\phi_2| = 1$. Therefore, if $\phi_1 \neq -\phi_2$, then $E(Z) \neq 0$. □

The mean direction is a measure of location which is defined as $\arg\{E(Z)\}$ if $E(Z) \neq 0$ and is undefined if $E(Z) = 0$. The mean resultant length, which is a measure of concentration, is defined by $|E(Z)|$. Figure 3 shows the mean direction and mean resultant length of a random variable having density (3) with fixed $\phi_2 = 0.5$. Figure 3(a) suggests that the mean direction monotonically increases as $\arg(\phi_1)$ increases from $-\pi$ to π. As seen in Corollary 1, Figure 3(b) confirms that the mean resultant length is zero if $\phi_1 = -0.5$. It seems from this frame that unimodality holds for the mean resultant length as a function of ϕ_1.

4.3 Skewness

A measure of skewness for circular distributions (Mardia, 1972), s, is defined by

$$s = E[\mathrm{Im}\{(Ze^{-i\zeta})^2\}]/(1 - \delta)^{3/2}, \tag{8}$$

where ζ and δ are the mean direction and mean resultant length for a ∂D-valued random variable Z, respectively. For our model, it is possible to express the skewness for the proposed model in a fairly simple form as follows.

Corollary 2. *Let* Z *be distributed as* $EC^*(\phi_1, \phi_2)$. *Then the skewness for the distribution of* Z *is given by*

$$s = \frac{|1 - \phi_1\overline{\phi_2}|^2}{(1 - |\phi_1\overline{\phi_2}|^2)^3}\rho^{-2}(1-\rho)^{-3/2}\,\mathrm{Im}\left(\phi_1\overline{\phi_2}\right)(|\phi_1|^2 - |\phi_2|^2)(1 - |\phi_1|^2)(1 - |\phi_2|^2), \tag{9}$$

where $\rho = |(1 - |\phi_2|^2)\phi_1 + (1 - |\phi_1|^2)\phi_2| \,/(1 - |\phi_1\overline{\phi_2}|^2)$.

Proof. Note that the numerator of (8) can be expressed as

$$E[\mathrm{Im}\{(Ze^{-i\mu})^2\}] = \mathrm{Im}\left[E(Z^2)\left\{\overline{E(Z)}\Big/|E(Z)|\right\}^2\right].$$

Then, by using Theorem 4 and equation (7), the skewness (9) can be obtained after lengthy but straightforward calculations. □

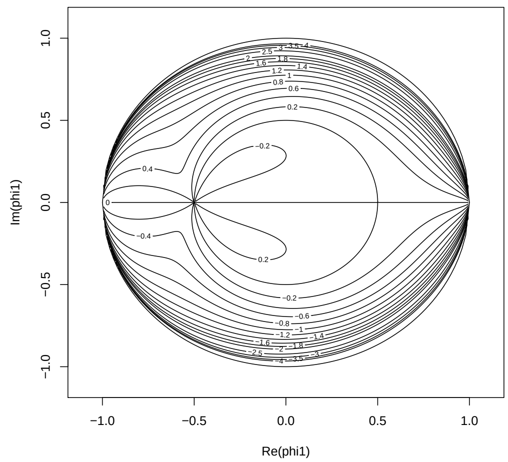


Figure 4: Skewness (9) of a random variable having the distribution $EC^*(\phi_1, 0.5)$, as a function of ϕ_1.

Now we write $s = s(\phi_1, \phi_2)$ in order to view the skewness (9) as a function of ϕ_1 and ϕ_2. It follows from Corollary 2 and Theorem 3 that the following properties hold for s.

Theorem 5. *The following properties hold for the skewness (9):*

1. $s(\phi_1, \phi_2) = 0 \iff$ *the density (3) is symmetric.*
2. $s(\phi_1, \phi_2) > 0\ (< 0) \iff Im(\phi_1\overline{\phi_2})(|\phi_1| - |\phi_2|) > 0\ (< 0)$.
3. $s(\phi_1, \phi_2) = -s(\phi_1, \overline{\phi_2}\phi_1^2/|\phi_1|^2) = -s(\overline{\phi_1}\phi_2^2/|\phi_2|^2, \phi_2), \quad \phi_1, \phi_2 \neq 0.$
4. $s(\phi_2, \phi_1) = s(\phi_1, \phi_2)$.
5. $s(\overline{\phi_1}, \overline{\phi_2}) = -s(\phi_1, \phi_2)$.
6. $s(\alpha\phi_1, \alpha\phi_2) = s(\phi_1, \phi_2), \quad \alpha \in \partial D.$
7. *Assume* $\rho_1, \rho_2 > 0$ *and* $\sin(\mu_1 - \mu_2) > 0\ (< 0)$. *Then* $\lim_{\rho_1 \to 1} s(\rho_1 e^{i\mu_1}, \rho_2 e^{i\mu_2}) = \infty\ (-\infty)$.

Figure 4 plots this skewness when $\phi_2 = 0.5$. As the first property of Theorem 5 shows, the skewness is equal to zero if and only if $\text{Im}(\phi_1\overline{\phi_2}) = 0$, $|\phi_1| = |\phi_2|$ or $|\phi_j| = 0$ $(j = 1, 2)$. This figure also confirms the second property of Theorem 5 that the model is positively (negatively) skewed if $|\phi_1| > |\phi_2|$ and $\text{Im}(\phi_1\overline{\phi_2}) > 0$ $(\text{Im}(\phi_1\overline{\phi_2}) < 0)$, or $|\phi_1| < |\phi_2|$ and $\text{Im}(\phi_1\overline{\phi_2}) < 0$ $(\text{Im}(\phi_1\overline{\phi_2}) > 0)$. It can also be seen that the fifth and seventh properties of Theorem 5 hold in this figure.

4.4 Path of the expected exit points

As described in Section 2.1, our model can be derived as the distribution of the position where a Brownian particle first hits the unit circle given a starting point, ϕ_1, and a future point, $\overline{\phi_2}^{-1}$, of the particle. Remembering this derivation and equation (7), one can calculate the location where the Brownian particle is expected to hit the unit circle for the first time. Then a natural problem arising from this might be to obtain the expected location where the particle hits a circle with radius between $\rho_1 = |\phi_1|$ and $1/\rho_2 = |\phi_2|^{-1}$ for the first time. The following theorem concerns the path consisting of such exit points of the particle.

Theorem 6. *Let $\{B_t\,;\,t \geq 0\}$ be $\mathbb{C}$-valued Brownian motion starting at $B_0 = \phi_1\,(\in D)$. Assume that $\tau_1(r) = \inf\{t\,;\,|B_t| = r\}$ and $\tau_2 = \inf\{t\,;\,|B_t| = |\phi_2|^{-1}\}$, where $|\phi_1| < r < |\phi_2|^{-1}$ and $\phi_2 \in D$. Let γ be the path of expected exit points defined by $\gamma = \{r\,\mathrm{ver}\{E(B_{\tau_1(r)}|B_{\tau_2} = \overline{\phi_2}^{-1})\}\,;\,r \in (|\phi_1|, |\phi_2|^{-1})\}$, where $\mathrm{ver}(z) = z/|z|$. Then γ can be expressed as*

$$\begin{aligned}\gamma &= \left\{r\,\mathrm{ver}\left\{(1-r^2|\phi_2|^2)r^{-1}\phi_1 + (1-r^{-2}|\phi_1|^2)r\phi_2\right\}\,;\,r \in (|\phi_1|, |\phi_2|^{-1})\right\} \quad (10)\\ &= \Bigg\{|\phi_1||\phi_1\phi_2|^{-t}\mathrm{ver}\Big[|\phi_1\phi_2|^{t}\Big\{1-|\phi_1\phi_2|^{2(1-t)}\Big\}\mathrm{ver}(\phi_1)\\ &\qquad + |\phi_1\phi_2|^{1-t}\left(1-|\phi_1\phi_2|^{2t}\right)\mathrm{ver}(\phi_2)\Big]\,;\,t \in (0,1)\Bigg\}. \quad (11)\end{aligned}$$

Proof. As in a similar manner to Theorem 1, the joint density of $(B_{\tau_1(r)}, B_{\tau_2})$ is obtainable. Similarly, the joint distribution of $(\mathrm{ver}(B_{\tau_1(r)}), B_{\tau_2})$ can be easily calculated using the fact that $\mathrm{ver}(B_{\tau_1(r)}) = r^{-1}B_{\tau_1(r)}$. Then it is straightforward to show that

$$\mathrm{ver}(B_{\tau_1(r)})\,|\,B_{\tau_2} = \overline{\phi_2}^{-1} \;\sim\; EC^*(r^{-1}\phi_1, r\phi_2).$$

It follows from equation (7) that

$$\begin{aligned}\mathrm{ver}\{E(B_{\tau_1(r)}\,|\,B_{\tau_2} = \overline{\phi_2}^{-1})\} &= \mathrm{ver}\left[E\left\{\mathrm{ver}(B_{\tau_1(r)})\,|\,B_{\tau_2} = \overline{\phi_2}^{-1}\right\}\right]\\ &= \mathrm{ver}\left\{(1-r^2|\phi_2|^2)r^{-1}\phi_1 + (1-r^{-2}|\phi_1|^2)\,r\phi_2\right\}.\end{aligned}$$

Thus we obtain (10). The other expression (11) immediately follows from the change-of-variable $r = |\phi_1|^{1-t}|\phi_2|^{-t}$. □

This theorem might be helpful to interpolate the exit points of a Brownian particle when the particle can be observed only at limited points or time.

Figure 5 shows paths of the expected exit points of the Brownian particle for a fixed starting point ϕ_1 and some selected values of the arrival point ϕ_2. The figure suggests that the path reduces to the straight line between ϕ_1 and $\overline{\phi_2}^{-1}$ if $\arg(\phi_1) = \arg(\phi_2)$, and this can be validated mathematically. However, in general, the path is a curve connecting ϕ_1 and $\overline{\phi_2}^{-1}$. Our simulation also suggests that, for all of the four combinations of the parameters ϕ_1 and ϕ_2 given in Figure 5,

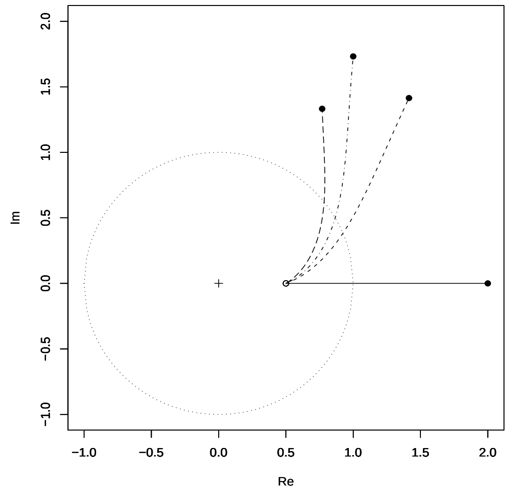


Figure 5: Plot of paths of the expected exit points, (10) or (11), for $\phi_1 = 0.5$ and $\phi_2 = 0.5$ (solid), $0.5\exp(\pi i/4)$ (dashed), $0.5\exp(\pi i/3)$ (dot-dashed) and $0.65\exp(\pi i/3)$ (long-dashed). The dotted line represents the unit circle.

the conditional distributions of $B_{\tau_1}(r)$ given $B_{\tau_2} = \overline{\phi_2}^{-1}$ are unimodal for any r. This unimodality makes it relatively easy to interpret the mean directions, which constitute the path of the expected exit points.

5 Some Other Topics

5.1 Alternative derivations

As discussed in Section 2.1, the proposed model can be derived by considering a problem in Brownian motion. Apart from the derivation given there, here is another method to generate our family via this stochastic process, kindly suggested by the referee.

Remember that, given a $\mathbb{C}$-valued Brownian path without drift starting from $\phi\,(\in D)$, the point of first exit from the unit disc D has a wrapped Cauchy distribution with density proportional to $1/|z-\phi|^2$. (Although we assume $\phi \in D$ by convention, the Brownian motion starting at $1/\phi$ has the same hitting distribution from the exterior because of the conformal invariance of Brownian motion.) From this it follows that another derivation of the proposed model (3) is given in the following theorem.

Theorem 7. *Let Z_1 and Z_2 be the points of first exit from the unit disc for two independent $\mathbb{C}$-valued Brownian motions without drift starting from ϕ_1 and ϕ_2, respectively, where $\phi_1, \phi_2 \in D$. Then the conditional distribution of $Z_1\,(\equiv Z)$ given $Z_1 = Z_2$ is given by the density (3).*

Remark: The referee embedded the above in a generalized formulation leading to densities of the form

$$f(z) \propto \prod_{j=1}^{d} \frac{1-|\phi_j|^2}{|z-\phi_j|^2}, \quad z \in \partial D; \quad \phi_1, \ldots, \phi_d \in D, \tag{12}$$

many of whose properties are also readily obtainable using the residue theorem.

In addition to the derivations using Brownian motion, there are other methods to generate our family. The following result shows that the proposed model appears in a Bayesian context.

Theorem 8. *Let $\Theta|\nu$ have the wrapped Cauchy distribution $WC(\nu, \rho_1)$ with known ρ_1. Assume that the prior distribution of ν is distributed as the wrapped Cauchy $WC(\mu_2, \rho_2)$. Then the posterior distribution of ν given $\Theta = \mu_1$ has density (2).*

The above derivation enables us to generate random samples from our model using the Markov Chain Monte Carlo method; see Section 5.3 for details.

Remark: In Theorem 8, the marginal distribution of Θ is given by the wrapped Cauchy distribution, $WC(\mu_1 + \mu_2, \rho_1\rho_2)$.

Before we discuss a third method to derive our model, here we briefly recall a known property of the wrapped Cauchy distribution or the Poisson kernel as follows.

Theorem 9. *[Rudin, 1987, Theorem 11.9] Assume $\overline{D}$ is the closed unit disc in the complex plane. Suppose a function u is continuous on $\overline{D}$ and harmonic in D, and suppose that $|u(z)| < \infty$ for any $z \in \overline{D}$. Then*

$$\int_{\partial D} u(z) \frac{1}{2\pi} \frac{1-|\phi_1|^2}{|z-\phi_1|^2} dz = u(\phi_1), \quad \phi_1 \in D.$$

Note that the integrand in the left-hand side of the above equation is actually the product of $u(z)$ and the density of the wrapped Cauchy $C^*(\phi_1)$. From the fact that

$$g(z) = \frac{|\phi_2^{-1}|^2 - |z|^2}{|z - \overline{\phi_2}^{-1}|^2}, \quad \phi_2 \in D,$$

is a continuous real function on $\overline{D}$ and harmonic in D, it follows that

$$f(z) = \frac{g(z)}{2\pi g(\phi_1)} \frac{1-|\phi_1|^2}{|z-\phi_1|^2}, \quad z \in \partial D,$$

satisfies the definition of a density. Clearly, $f(z)$ is equal to density (3).

The following theorem implies that our model appears as a special case of a further general family of distributions. The proof is straightforward and omitted.

Theorem 10. *Let f and g be probability density functions on the circle. Assume that h is the convolution of f and g, namely, $h(\tau) = (f * g)(\tau) = \int_{-\pi}^{\pi} f(\theta)g(\tau-\theta)d\theta$. Then a function $k(\theta) = f(\theta)g(\tau-\theta)/h(\tau)$ is a probability density function on the circle.*

Density (2) can be derived on setting $\tau = \mu_2$ and f and g as the densities of $WC(\mu_1, \rho_1)$ and $WC(0, \rho_2)$, respectively. In this case $h(\tau)$ also has the form of the wrapped Cauchy density $WC(\mu_1, \rho_1\rho_2)$ since the wrapped Cauchy distribution has the additive property.

5.2 Möbius transformation

The Möbius transformation is well known in complex analysis as a conformal mapping which projects the unit disc onto itself. It is defined as

$$\mathcal{M}(w) = \alpha \frac{w + \beta}{\overline{\beta} w + 1}, \quad w \in \overline{D}; \quad \alpha \in \partial D, \ \beta \in D.$$

Although this projection is usually defined on the interior of the unit disc, here we extend the domain of the mapping so that the boundary of the unit disc, i.e. the unit circle, is included. It is easy to see that the unit circle is mapped onto itself via the transformation, namely, $\mathcal{M}(\partial D) = \partial D$. In directional statistics the transformation appears in some papers such as McCullagh (1996), Jones (2004) and Kato and Jones (2010).

Here we study the conformal invariance properties of our family. The distribution (3) has a density that is relative invariant of weight 2 in the sense that

$$f(z; \phi_1, \phi_2) = f\{\mathcal{M}(z); \mathcal{M}(\phi_1), \mathcal{M}(\phi_2)\} \, |\mathcal{M}'(z)|^2.$$

Similarly, the extended family (12) has a density that is relatively invariant with weight k in the sense that $f(z; \phi_1, \ldots, \phi_k) = f\{\mathcal{M}(z); \mathcal{M}(\phi_1), \ldots, \mathcal{M}(\phi_k)\} \, |\mathcal{M}'(z)|^k$. The relative invariance with weight 1 of the wrapped Cauchy distribution can be obtained by putting $k = 1$.

The proposed family (3) is not closed under the Möbius transformation except for the wrapped Cauchy special cases. To say the same thing in a different way, if Z follows the distribution (3), then the density for $\mathcal{M}(Z)$ is not of the form (3) except for $\phi_1 = 0$ or $\phi_2 = 0$. This fact can be understood clearly in the following context; as seen in Theorem 7, our model (3) can be derived as the conditional distribution of Z_1 given $Z_1 = Z_2$, where Z_1 and Z_2 are independent wrapped Cauchy variables. Invertibility of the transformations $\mathcal{M}$ implies that $Z_1 = Z_2$ if and only if $\mathcal{M}(Z_1) = \mathcal{M}(Z_2)$. However, as is known as the Borel paradox (e.g. Pollard, 2002, Section 5.5), the conditional distribution of Z_1 given $Z_1 = Z_2$ is generally not the same as the conditional distribution of Z_1 given $\mathcal{M}(Z_1) = \mathcal{M}(Z_2)$. Consequently, the Möbius transformation of the conditional distribution of Z_1 given $Z_1 = Z_2$, which is the Möbius transformation of model (3), is not the same as the conditional distribution of $\mathcal{M}(Z_1)$ given $\mathcal{M}(Z_1) = \mathcal{M}(Z_2)$, which is the model (3). In a similar manner one can see that the extended model (12) with $k \geq 2$ is not closed under the Möbius transformation.

5.3 Random variate generation

To generate random variates from model (2), one can adopt some general algorithms such as the acceptance-rejection and inversion methods. However these methods are not efficient in some situations. For instance, when ρ_1 or ρ_2 is large, the acceptance ratio for the former method is not high since the peak of the density goes to infinity as ρ_1 or ρ_2 tends to one.

Applying the fact that the proposed model can be derived in the Bayesian context (see Theorem 8), here we present a method for random variate generation based on Markov Chain Monte Carlo. To generate many random samples from our model with large ρ_1 or ρ_2, this algorithm is more efficient than the acceptance-rejection method. The following steps generate random variates from model $EWC(\mu_1, \mu_2, \rho_1, \rho_2)$.

1. Take an initial value θ_0.
2. For $j = 1, 2, \ldots$, repeat the following steps:
   (a) Generate uniform $(0, 1)$ random variates $\tilde{\theta}$ and u.
   (b) Put $\theta = \mu_1 + 2 \arctan \left[\{(1-\rho_1)/(1+\rho_1)\} \tan \left(\frac{1}{2}(\tilde{\theta} - \mu_1) \right) \right]$.
   (c) Calculate
   $$\alpha(\theta_{j-1}, \theta) = \min \left\{ 1 \, , \, \frac{1+\rho_2^2 - 2\rho_2 \cos(\theta_{j-1} - \mu_2)}{1+\rho_2^2 - 2\rho_2 \cos(\theta - \mu_2)} \right\}.$$
   (d) Define θ_j as
   $$\theta_j = \begin{cases} \theta, & \text{if } u \leq \alpha(\theta_{j-1}, \theta), \\ \theta_{j-1}, & \text{if } u > \alpha(\theta_{j-1}, \theta). \end{cases}$$
3. Record θ_j $(j = N, N+1, \ldots)$, where N is a number which is sufficiently large enough for the algorithm to have converged.

The Metropolis-Hastings algorithm is useful for this model because (i) we can generate from the proposal distribution, i.e. the wrapped Cauchy $WC(\mu_1, \rho_1)$, without rejecting any uniform variates, and (ii) the ratio α is expressed in a simple form. The reason (i) holds is that the uniform distribution is transformed into the wrapped Cauchy distribution by the Möbius transformation (McCullagh, 1996). Therefore the algorithm enables us to generate variates from our model without rejection after N steps.

6 Symmetric Cases

So far, we have mostly considered the full family of distributions with densities (2) and (3). In this section we focus on the symmetric special cases of the proposed model. Some properties which the general family do not have hold for the symmetric cases. In Section 6.1, model (2) with $\mu_1 = \mu_2$ or $\mu_2 + \pi$ is discussed. In Section 6.2, we briefly consider another symmetric model, namely, model (2) with $\rho_1 = \rho_2$.

6.1 Symmetric case I: $\mu_1 = \mu_2$ or $\mu_1 = \mu_2 + \pi$

Model (2) with $\mu_1 = \mu_2$ or $\mu_1 = \mu_2 + \pi$ is essentially the same as the distribution with density

$$f(\theta) = \frac{1-\rho_1\rho_2}{2\pi(1+\rho_1\rho_2)} \frac{1-\rho_1^2}{1+\rho_1^2-2\rho_1\cos(\theta-\mu)} \frac{1-\rho_2^2}{1+\rho_2^2-2\rho_2\cos(\theta-\mu)}, \quad -\pi \le \theta < \pi, \tag{13}$$

where $-1 < \rho_1 < 1$, $0 \le \rho_2 < 1$, and $-\pi \le \mu < \pi$. Note the extension in the range of ρ: this density corresponds to (2) with $\mu_1 = \mu_2$ when $\rho_1 \ge 0$ in (13), while the density is equivalent to (2) with $\mu_1 = \mu_2 + \pi$ when $\rho_1 < 0$ in (13). Clearly, the density is symmetric about $\theta = \mu$ and $\theta = \mu + \pi$.

One might, however, wish to restrict interest to the case where $0 \le \rho_1 < 1$ because then density (13) is unimodal. To see this, note that in this case $a_0 = a_1 = a_4 = 0$ in (5) so that stationary points of the density occur at $\sin\theta = 0$ and, potentially, also at $\cos\theta = -a_2/a_3 = (\rho_1+\rho_2)(1+\rho_1\rho_2)/4\rho_1\rho_2$. However, the latter expression can easily be proved not be less than unity, showing that (13) is unimodal. On the other hand, the symmetric model (13) with negative ρ_1 has a relationship with mixtures of two wrapped Cauchy distributions. The density for this submodel can be expressed as

$$f(\theta) \;=\; p\frac{1}{2\pi}\frac{1-\rho_1^2}{1+\rho_1^2-2\rho_1\cos(\theta-\mu)} + (1-p)\frac{1}{2\pi}\frac{1-\rho_2^2}{1+\rho_2^2-2\rho_2\cos(\theta-\mu)}$$

where $p = -\rho_1(1-\rho_2^2)/\{(1+\rho_1\rho_2)(\rho_2-\rho_1)\}$. Inter alia, this representation allows random numbers from density (13) with $\rho_1 \le 0$ to be more easily generated than they were for the general case in Section 5.3.

6.2 Symmetric case II: $\rho_1 = \rho_2$

In this section we discuss the submodel of the density (2) with $\rho_1 = \rho_2 = \rho\,(\in [0,1))$. The density for this submodel reduces to

$$\begin{aligned} f(\theta) \;&=\; \frac{(1-\rho^2)\{1+\rho^4-2\rho^2\cos(\mu_1-\mu_2)\}}{2\pi(1+\rho^2)} \\ &\quad \times \left[\left\{1+\rho^2-2\rho\cos(\theta-\mu_1)\right\}\left\{1+\rho^2-2\rho\cos(\theta-\mu_2)\right\}\right]^{-1}, \quad -\pi \le \theta < \pi. \end{aligned}$$

The model is symmetric about $\theta = \overline{\mu}$ and $\overline{\mu} + \pi$, where $\overline{\mu} = \arg(\sum_{j=1}^2 \cos\mu_j + i\sum_{j=1}^2 \sin\mu_j)$. As displayed in Figure 2(d), this submodel can be unimodal or bimodal depending on the choice of the parameters. The condition for unimodality can be simplified for this submodel as follows:

$$-2\arccos\left(\frac{2\rho}{1+\rho^2}\right) \le \mu_1 - \mu_2 \le 2\arccos\left(\frac{2\rho}{1+\rho^2}\right).$$

This result was used in Section 3.2 to find the region of Figure 1(b) in which the discriminant takes negative values, corresponding to unimodality, when $\rho_1 = \rho_2$ and $\mu_1 - \mu_2 = 2\pi/3$.

7 A Generalisation on the Sphere

As described in Section 2.1, the proposed model (2) can be derived using Brownian motion. By adopting a multi-dimensional Brownian motion instead of the two-dimensional one, we can extend the model (2) to a distribution on the unit sphere. The generalised model is defined as follows.

Definition 2. *Let $\{B_t\,;\, t \geq 0\}$ be $\mathbb{R}^d$-valued Brownian motion starting at $B_0 = \rho_1\eta_1$, where $d \geq 2$, $0 \leq \rho_1 < 1$, $\eta_1 \in S^d$, and $S^d = \{x \in \mathbb{R}^d\,;\, \|x\| = 1\}$. Assume $\tau_1 = \inf\{t\,;\, \|B_t\| = 1\}$ and $\tau_2 = \inf\{t\,;\, \|B_t\| = \rho_2^{-1}\}$, where $0 < \rho_2 < 1$. Then the proposed model is defined by the conditional distribution of B_{τ_1} given $B_{\tau_2} = \rho_2^{-1}\eta_2$, where $\eta_2 \in S^d$.*

For simplicity, write $X = B_{\tau_1}$. The probability density function of this extended model is available, and it is given in the following theorem.

Theorem 11. *The conditional distribution of X given $B_{\tau_2} = \rho_2^{-1}\eta_2$ is of the form*

$$f(x) = \frac{1}{A_{d-1}} \frac{(1 + \rho_1^2\rho_2^2 - 2\rho_1\rho_2\eta_1'\eta_2)^{d/2}}{1 - \rho_1^2\rho_2^2} \frac{1 - \rho_1^2}{\|x - \rho_1\eta_1\|^d} \frac{1 - \rho_2^2}{\|x - \rho_2\eta_2\|^d}, \quad x \in S^d, \qquad (14)$$

where A_{d-1} is the surface area of S^d, i.e., $A_{d-1} = 2\pi^{d/2}/\Gamma(d/2)$.

See Appendix C for the proof. Note that the density (14) reduces to the circular case (2) if $d = 2$, $x = (\cos\theta, \sin\theta)'$, $\eta_1 = (\cos\mu_1, \sin\mu_1)'$ and $\eta_2 = (\cos\mu_2, \sin\mu_2)'$. It might be appealing that the density (14), which is not rotationally symmetric in general, can be expressed in a relatively simple form.

When $\rho_2 = 0$, the distribution becomes the so-called 'exit' distribution on the sphere, whose density is given by

$$f(x) = \frac{1}{A_{d-1}} \frac{1 - \rho_1^2}{\|x - \rho_1\eta_1\|^d}, \quad x \in S^d. \qquad (15)$$

We write $X \sim \mathrm{Exit}_d(\rho_1\eta_1)$ if a random variable X has the density (15). This model is rotationally symmetric about $x = \eta_1$. See, for example, Durrett (1984, Section 1.10) for details about the distribution. In particular, when $d = 2$, the model (15) becomes the wrapped Cauchy distribution. It is noted that this distribution is a submodel of Jones and Pewsey's (2005) family of distributions on the sphere with density

$$f(x) = \frac{1}{A_{d-1}} \frac{|\sinh(\kappa\psi)|^{d/2-1}}{2^{2/d-1}\Gamma(d/2)} \frac{\{\cosh(\kappa\psi) + \sinh(\kappa\psi)x'\mu\}^{1/\psi}}{P_{1/\psi+d/2-1}^{1-d/2}\{\cosh(\kappa\psi)\}}, \quad x \in S^d, \qquad (16)$$

where $P_\alpha^\beta(z)$ is the associated Legendre function of the first kind of degree α and order β (Gradshteyn and Ryzhik, 1994, Sections 8.7, 8.8). It follows from equation (8.711.1) of Gradshteyn and Ryzhik (1994) and equation (2) of McCullagh (1989) that density (16) reduces to density (15) if $\kappa = d \log\{(1+\rho_1)/(1-\rho_1)\}/2$, $\mu = \eta_1$, and $\psi = -2/d$. Also, the extended model (14) includes the model with density

$$f(x) = \frac{1}{A_{d-1}} \frac{(1-\rho_1^2)^{d+1}}{1+\rho_1^2} \frac{1}{\|x - \rho_1 \eta_1\|^{2d}}, \quad x \in S^d, \tag{17}$$

when $\rho_1 = \rho_2$, $\eta_1 = \eta_2$. It can be seen that this model is another submodel of Jones and Pewsey's (2005) family (16) by putting $\psi = -1/d$ and $\kappa = d \log\{(1+\rho_1)/(1-\rho_1)\}$. Also, notice that, if $d = 2$, density (17) corresponds to Case 2, (4), of circular submodels in Section 2.4. In addition to these two submodels, the multivariate model (14) contains the uniform distribution ($\rho_1 = \rho_2 = 0$), one-point distribution ($\rho_1 = 0$ and $\rho_2 \to 1$, or $\rho_2 = 0$ and $\rho_1 \to 1$), and two-point distribution ($\rho_1 = \rho_2$, $\mu_1 \neq \mu_2$ and $\rho_1 \to 1$).

In a similar manner to Theorem 8, it is easy to prove that model (14) can be derived from Bayesian analysis of the exit distribution.

Theorem 12. *Let $X|\xi$ be distributed as the exit distribution Exit($\rho_1\xi$) with known ρ_1. Suppose that the prior distribution of ξ is Exit($\rho_2\eta_2$). Then the posterior distribution of ξ given $X = \eta_1$ is given by density (14).*

From this result and the random number generator of the exit distribution given by Kato (2009, Section 3.3), it is possible to generate random samples from (14) through the Metropolis-Hastings algorithm.

Appendices

A. Proof of Theorem 1

Write $B_{\tau_1} = (\cos\Theta, \sin\Theta)'$ and $B_{\tau_2} = \rho_2^{-1}(\cos\Xi, \sin\Xi)'$. Let $f_{\Theta|\Xi}$ be the conditional density of Θ given Ξ. Note that if there exists a joint density for (Θ, Ξ), it can be expressed as

$$f_{\Theta|\Xi}(\theta|\xi) = \frac{f_\Theta(\theta) g_{\Xi|\Theta}(\xi|\theta)}{h_\Xi(\xi)}, \quad -\pi \le \theta < \pi, \tag{18}$$

where f_Θ is the marginal density of Θ, $g_{\Xi|\Theta}$ is the conditional density of Ξ given $\Theta = \theta$, and h_Ξ is the marginal density of Ξ. The marginal of Θ is known to be the exit distribution for the unit circle or, equivalently, the wrapped Cauchy distribution $WC(\mu_1, \rho_1)$. (See, for example, McCullagh (1996).) Because of the Markov property of Brownian motion, the conditional distribution of Ξ given $\Theta = \theta$ is also equivalent to an exit distribution, that for the circle with radius ρ_2^{-1} generated by Brownian motion starting at $(\cos\theta, \sin\theta)$. It follows that this conditional distribution is the wrapped Cauchy $WC(\theta, \rho_2)$. Similarly, the marginal distribution of Ξ is shown to be $WC(\mu_1, \rho_1\rho_2)$. Thus, from (18), the conditional distribution of Θ given $\Xi = \xi$ is given by

$$f_{\Theta|\Xi}(\theta|\xi) \quad = \quad \frac{1}{4\pi^2} \frac{1-\rho_1^2}{1+\rho_1^2 - 2\rho_1 \cos(\theta - \mu_1)} \frac{1-\rho_2^2}{1+\rho_2^2 - 2\rho_2 \cos(\xi - \theta)}$$

$$
\Big/ \left\{ \frac{1}{2\pi} \frac{1-\rho_1^2\rho_2^2}{1+\rho_1^2\rho_2^2-2\rho_1\rho_2\cos(\xi-\mu_1)} \right\}
$$

$$
= \frac{1+\rho_1^2\rho_2^2-2\rho_1\rho_2\cos(\xi-\mu_1)}{2\pi(1-\rho_1^2\rho_2^2)} \frac{1-\rho_1^2}{1+\rho_1^2-2\rho_1\cos(\theta-\mu_1)} \frac{1-\rho_2^2}{1+\rho_2^2-2\rho_2\cos(\theta-\xi)}.
$$

By putting $\xi = \mu_2$, we obtain the density (2). □

B. Proof of Theorem 3

Clearly, if $\mu_1 = \mu_2$, $\mu_1 = \mu_2 + \pi$, or $\rho_2 = 0$, the density (2) is symmetric about $\theta = \mu_1$ and $\mu_1 + \pi$. When $\rho_1 = 0$, the density is symmetric about $\theta = \mu_2$ and $\mu_2 + \pi$. If $\rho_1 = \rho_2$, then the density can be expressed as

$$
\begin{aligned}
f(\theta) &\propto \left[\left\{ 1+\rho^2-2\rho\cos(\theta-\mu_1) \right\} \left\{ 1+\rho^2-2\rho\cos(\theta-\mu_2) \right\} \right]^{-1} \\
&= \left[\left\{ 1+\rho^2-2\rho\cos(\theta-\mu_1) \right\} \left\{ 1+\rho^2-2\rho\cos(\theta-2\overline{\mu}+\mu_1) \right\} \right]^{-1},
\end{aligned}
$$

where $\overline{\mu} = \arg(\sum_{j=1}^2 \cos\mu_j + i\sum_{j=1}^2 \sin\mu_j)$. Then it is easy to see that the density is symmetric about $\theta = \overline{\mu}$ and $\overline{\mu} + \pi$.

Next we prove the necessary condition for symmetry. Let $\rho_1 \neq \rho_2$, $\rho_j > 0$ and $\mu_1 \neq \mu_2 + (j-1)\pi$ $(j = 1, 2)$. Suppose that there exists a constant $a\,(\in [-\pi, \pi))$ such that $f(a-\theta) = f(a+\theta)$ for any $-\pi \leq \theta < \pi$. Then it holds that

$$
\frac{1+\rho_1^2-2\rho_1\cos(a-\mu_1+\theta)}{1+\rho_1^2-2\rho_1\cos(a-\mu_1-\theta)} = \frac{1+\rho_2^2-2\rho_2\cos(a-\mu_2-\theta)}{1+\rho_2^2-2\rho_2\cos(a-\mu_2+\theta)}, \quad -\pi \leq \theta < \pi. \tag{19}
$$

If $\theta = \pi/2$, we have

$$
\frac{1+\rho_1^2+2\rho_1\sin(a-\mu_1)}{1+\rho_1^2-2\rho_1\sin(a-\mu_1)} = \frac{1+\rho_2^2-2\rho_2\sin(a-\mu_2)}{1+\rho_2^2+2\rho_2\sin(a-\mu_2)}.
$$

Since $\rho_j > 0$ $(j = 1, 2)$, it follows that

$$
\sin(a-\mu_2) = -\frac{\rho_1}{\rho_2}\frac{1+\rho_2^2}{1+\rho_1^2}\sin(a-\mu_1). \tag{20}
$$

If $\theta = \pi/4$ in (19), then

$$
\sin(a-\mu_2) = -\frac{\rho_1}{\rho_2}\frac{1+\rho_2^2-\sqrt{2}\rho_2\cos(a-\mu_2)}{1+\rho_1^2-\sqrt{2}\rho_1\cos(a-\mu_1)}\sin(a-\mu_1). \tag{21}
$$

Combining (20) and (21), we have

$$
\cos(a-\mu_2) = \frac{\rho_1}{\rho_2}\frac{1+\rho_2^2}{1+\rho_1^2}\cos(a-\mu_1). \tag{22}
$$

Then, from (20) and (22),

$$
\tan(a-\mu_2) = -\tan(a-\mu_1).
$$

Thus we get $a = \overline{\mu}$ or $\overline{\mu} + \pi$. However, substituting $\overline{\mu}$ for a in (20) and using the assumption $\mu_1 \neq \mu_2, \mu_2 + \pi$, we obtain $\rho_1 = \rho_2$. This is contradictory to the initial assumption. □

C. Proof of Theorem 11

For convenience, write $X = B_{\tau_1}$ and $Y = B_{\tau_2}$. Let $f_{X|Y}$ be the conditional density of X given Y. The distribution of interest can be expressed as

$$f_{X|Y}(x|y) = \frac{f_X(x) g_{Y|X}(y|x)}{h_Y(y)}, \tag{23}$$

where f_X is the marginal density of X, h_Y is the marginal density of Y and $g_{Y|X}$ is the conditional density of Y given $X = x$. From the derivation of the model, it follows that the marginal distribution of X is the exit distribution $\text{Exit}_d(\xi_1)$. (See, for example, Durrett (1984, Section 1.10).) Because of the Markov property of Brownian motion, it can be seen that the conditional distribution of Y given $X = x$ is the same as the distribution of the position of a Brownian particle which first hits the sphere with radius ρ_2^{-1} given the starting point at x. Thus this conditional distribution follows an exit distribution on the sphere with radius ρ_2^{-1} whose density is given by

$$g_{Y|X}(y|x) = \frac{1}{A_{1/\rho_2, d-1}} \frac{\|y\|^2 - 1}{\|y - x\|^d}, \quad y \in \rho_2^{-1} S^d,$$

where $A_{r,d-1} = 2\pi^{d/2} r^{d-1} / \Gamma(d/2)$ and $cS^d = \{x \in \mathbb{R}^d \, ; \, \|x\| = c, \ c > 0\}$. Similarly, the density for the marginal distribution of Y is

$$h_Y(y) = \frac{1}{A_{1/\rho_2, d-1}} \frac{\|y\|^2 - \|\xi_1\|^2}{\|y - \xi_1\|^d}, \quad y \in \rho_2^{-1} S^d.$$

Since the three densities in the right-hand side of equation (23) have been obtained, it follows that the conditional density of X given $Y = y$ has the expression

$$f_{X|Y}(x|y) = \frac{1}{A_{d-1}} \frac{\|y - \xi_1\|^d}{\|y\|^2 - \|\xi_1\|^2} \cdot \frac{1 - \|\xi_1\|^2}{\|x - \xi_1\|^d} \cdot \frac{\|y\|^2 - 1}{\|y - x\|^d}. \tag{24}$$

Substituting $y = \rho_2^{-1} \eta_2$ and $\xi_1 = \rho_1 \eta_1$ in (24), we obtain the density (14). □

Acknowledgements

The authors are grateful to the Associate Editor and a referee for suggesting a simpler proof of Theorem 4, the alternative derivation of the proposed model via Brownian motion given in Theorem 7, and the conformal invariance properties discussed in Section 5.2.

References

Abramowitz, M. and Stegun, I.A. (1970). *Handbook of Mathematical Functions*. New York: Dover Press.

Batschelet, E. (1981). *Circular Statistics in Biology*. London: Academic Press.

Durrett, R. (1984). *Brownian Motion and Martingales in Analysis*. Belmont, CA: Wadsworth.

Gatto, R. and Jammalamadaka, S.R. (2007). The generalized von Mises distribution. *Stat. Methodol.* **4** 341–353.

Gradshteyn, I.S. and Ryzhik, I.M. (2007). *Table of Integrals, Series, and Products*, 7th ed. San Diego: Academic Press.

Jones, M.C. (2004). The Möbius distribution on the disc. *Ann. Inst. Statist. Math.* **56** 733–742.

Jones, M.C. and Pewsey, A. (2005). A family of symmetric distributions on the circle. *J. Amer. Statist. Assoc.* **100** 1422–1428.

Kato, S. (2009). A distribution for a pair of unit vectors generated by Brownian motion. *Bernoulli* **15** 898–921.

Kato, S. and Jones, M.C. (2010). A family of distributions with links to, and applications arising from, Möbius transformation. *J. Amer. Statist. Assoc.* **102** 249–262.

Kent, J.T. and Tyler, D.E. (1988). Maximum likelihood estimation for the wrapped Cauchy distribution. *J. Appl. Statist.* **15** 247–254.

Maksimov, V.M. (1967). Necessary and sufficient statistics for a family of shifts of probability distributions on continuous bicompact groups. *Teor. Veroyatn. Primen.* **12** 307–321 (in Russian). English translation: *Theor. Probab. Appl.* **12** 267–280.

Mardia, K.V. (1972). *Statistics of Directional Data*. London: Academic Press.

McCullagh, P. (1989). Some statistical properties of a family of continuous univariate distributions. *J. Amer. Statist. Assoc.* **84** 125–129.

McCullagh, P. (1996). Möbius transformation and Cauchy parameter estimation. *Ann. Statist.* **24** 787–808.

Pewsey, A. (2008). The wrapped stable family of distributions as a flexible model for circular data. *Comput. Statist. Data Anal.* **52** 1516–1523.

Pollard, D. (2002). *A User's Guide to Measure Theoretic Probability*. Cambridge: Cambridge University Press.

Rudin, W. (1987). *Real and Complex Analysis*, 3rd ed. New York: McGraw-Hill.

Uspensky, J.V. (1948). *Theory of Equations*. New York: McGraw-Hill.

Yfantis, E.A. and Borgman, L.E. (1982). An extension of the von Mises distribution. *Commun. Statist. – Theor. Meth.* **11** 1695–1706.