Filomat 32:17 (2018), 5931–5947
https://doi.org/10.2298/FIL1817931M

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Transformation of Circular Random Variables Based on Circular Distribution Functions

Hatami Mojtaba[a], Alamatsaz Mohammad Hossein[a]

[a] *Department of Statistics, University of Isfahan, Isfahan, 8174673441, Iran*

Abstract. In this paper, we propose a new transformation of circular random variables based on circular distribution functions, which we shall call inverse distribution function (idf) transformation. We show that Möbius transformation is a special case of our idf transformation. Very general results are provided for the properties of the proposed family of idf transformations, including their trigonometric moments, maximum entropy, random variate generation, finite mixture and modality properties. In particular, we shall focus our attention on a subfamily of the general family when idf transformation is based on the cardioid circular distribution function. Modality and shape properties are investigated for this subfamily. In addition, we obtain further statistical properties for the resulting distribution by applying the idf transformation to a random variable following a von Mises distribution. In fact, we shall introduce the Cardioid-von Mises (CvM) distribution and estimate its parameters by the maximum likelihood method. Finally, an application of CvM family and its inferential methods are illustrated using a real data set containing times of gun crimes in Pittsburgh, Pennsylvania.

1. Introduction

There are various general methods that can be used to produce circular distributions. One popular way is based on families defined on the real line, i.e., linear distributions such as normal, Cauchy and etc. Examples of such methods are the projection (or offsetting), conditioning, wrapping and (inverse) stereographic projection (see Mardia and Jupp [16] and Jammalamadaka and SenGupta [10] for more details).

Other methods of obtaining circular models are based on circular families. A first general method of this type is perturbation. In this approach, the density of an existing circular density is multiplied by some function chosen to ensure that their product is also a circular density; e.g., cardioid density is obtained by cosine perturbation of the continuous circular uniform density. Extending this idea, Umbach and Jammalamadaka [22] adapted the perturbation approach of Azzalini [5] to the circular context. One of the special cases of this general approach is the sine-skewed family of distributions studied by Abe and Pewsey [1]. Also, Abe and Pewsey [2] using duplication and cosine perturbation proposed models with two diametrically opposed modes. A second method is to apply transformation of argument to some existing density, $f(\theta)$, replacing its argument θ by some function of θ. This method was first applied by

2010 *Mathematics Subject Classification*. Primary 60E05; Secondary 62H11
Keywords. Trigonometric moments; Skewness; Unimodality; Finite mixtures; Cardoid distribution; Cardoid-von Mises distribution.
Received: 19 March 2018; Accepted: 16 July 2018
Communicated by Aleksandar Nastić
Email addresses: m.hatami.v@gmail.com (Hatami Mojtaba), alamatho@sci.ui.ac.ir (Alamatsaz Mohammad Hossein)

Papakonstantinou [17] and Batschelet [6] and was studied further by Abe et al. [4], Pewsey et al.[19] and Jones and Pewsey [11]. General properties of distributions derived using two particular forms of change of argument are given in Abe et al. [3]. A further type of transformation that can be applied to a distribution defined on the unit circle is the so-called Möbius transformation. Kato and Jones [12] built, via a Möbius transformation, a new four-parameter family of circular distributions.

Another method is based on non-negative trigonometric sums (NNTS) proposed by Fernández-Durán [7]. NNTS densities are finite Fourier series constrained to be non-negative. Finally, some new circular distributions can be derived by extensions of classic circular distributions; for instance, Gatto and Jammalamadaka [8] and Kim and SenGupta [14] presented extensions of the von Mises distribution and also Kato and Jones [13] proposed an extension of the wrapped Cauchy.

In this paper, we intend to introduce a new transformation of circular random variables in order to construct new and more flexible circular models. In this method, proper choices of circular distribution functions provide various models which can be symmetric, asymmetric and skew-symmetric. We, then, produce a special class of distributions using cardioid distribution function. The resulting distribution can be used to model both symmetric and asymmetric, unimodal or bimodal data with a very wide range of skewness and kurtosis.

The paper is organized as follows. In Sec. 2, we first define our idf transformation. Then, we shall provide very general results concerning modality and moments of the resulting new models. We shall also present a random variate generation method and investigate maximum entropy and finite mixtures of the proposed families of idf transformation. We also show that Möbius transformation is a special case of idf transformation. In Sec. 3, we consider cardioid distribution function as a special case and introduce a new class of circular distributions and provide very general results for their structural properties. In Sec. 4, applying the new transformation, Cardioid-von Mises (CvM) circular distribution is introduced. We also study properties and moments estimation of its symmetric Cardioid-von Mise (SCvM) subclass. Then, inferences are made concerning parameters of certain members of the family by the maximum likelihood method. A simulation study is also presented to assess the performance of the estimators. Application of the proposed model is considered in Sec. 5 for the times of gun crimes in Pittsburgh, Pennsylvania data.

2. General Method

A circular distribution is usually described in terms of a circular density, which is a function $f(\theta)$ defined on angle θ, satisfying the conditions

1. $f(\theta) \geq 0 \ \ for \ -\infty < \theta < \infty$,
2. $f(\theta + 2k\pi) = f(\theta) \ \ for \ k \in \mathbb{Z}, -\infty < \theta < \infty$,
3. $\int_{-\pi}^{\pi} f(\theta)d\theta = 1$.

Let $F(\theta) = \int_{-\pi}^{\theta} f(w)dw$. A circular distribution function (d.f.) is defined by F restricted on $[-\pi, \pi]$, i.e.,

$$P(-\pi < \Theta \leq \theta) = F(\theta), \ \ -\pi \leq \theta \leq \pi, \tag{1}$$

and

$$F(\theta + 2\pi) - F(\theta) = 1, \ -\infty < \theta < \infty. \tag{2}$$

Eq. (1) is a circular analogue of the usual definition of a distribution function for a random variable (r.v.) observed on the real line (linear r.v.). Eq. (2) is an extra condition imposed to reflect the periodicity of a circular distribution, i.e., probability of obtaining a point on the unit circle within any arc of length 2π radians equals 1. As we observe, the circular d.f. F defined above differs from a linear d.f. in having the following mathematical properties:

$$\lim_{\theta \to -\infty} F(\theta) = -\infty, \ \lim_{\theta \to \infty} F(\theta) = \infty \tag{3}$$

Further, we have

$$\int_{\omega}^{\psi} f(\theta)d\theta = F(\psi) - F(\omega), \quad \omega \leq \psi \leq \omega + 2\pi, \ \ \omega \in \mathbb{R}.$$

By definition, $F(-\pi) = 0$, $F(\pi) = 1$ (Mardia and Jupp [16], Pewsey et al. [20]).

A method for constructing new distributions on the real line is using transformation of r.v.s, such as Box-Cox, Yeo-Johnson, g-and-h, sinh-arcsinh transformations. Any function of linear random variable is a linear r.v. For example, in the easiest case, let X be a continuous linear r.v. with probability density function $f_X(x)$ and, let $Y = u(X)$, where u is a monotone function and u^{-1} is continuous and differentiable. Then, the probability density function of Y is $f_Y(y) = \left|\frac{d}{dy}u^{-1}(y)\right| f_X(u^{-1}(y))$. But any function of a circular random variable does not necessarily lead to a circular r.v.. In the following theorem, we present certain conditions under which this will happen.

Theorem 2.1. *Let $T : \mathbb{R} \to \mathbb{R}$ be a continuous function whose first derivative of T^{-1} is continuous and Ψ be a continuous circular r.v.. Then, $\Theta = T(\Psi)$ has a circular distribution if*

1. *$T : [-\pi, \pi) \to [t_1, t_2)$ such that $t_2 - t_1 = 2\pi$,*
2. *T is monotone on $[-\pi, \pi]$,*
3. *$T^{-1}(\theta + 2k\pi) = T^{-1}(\theta)$ or $T^{-1}(\theta + 2k\pi) = 2k\pi + T^{-1}(\theta)$, for $-\infty < \theta < \infty$ and $k \in \mathbb{Z}$,*

Proof. The density function of the r.v. $\Theta = T(\Psi)$ is

$$f_{\Theta}(\theta) = \left|\frac{d}{d\theta}T^{-1}(\theta)\right| f_{\Psi}(T^{-1}(\theta)).$$

Thus, it clearly follows that $f_{\Theta}(\theta+2k\pi) = f_{\Theta}(\theta)$ for $k \in \mathbb{Z}$. Therefore, $\Theta = T(\Psi)$ has a circular distribution. □

Now, consider the function T such that $T^{-1}(\theta) = 2\pi F(\theta - \xi) - \pi$, where F is the circular d.f. associated with circular density f and $-\pi \leq \xi \leq \pi$. Obviously, the function T satisfies conditions of Theorem 2.1. Thus, $\Theta = T(\Psi)$ is a circular r.v..

Corollary 2.2. *Suppose that the circular r.v. Ψ has density function $g(\psi)$, $\psi \in [-\pi, \pi)$. Then, the resulting circular density function of transformation, $\Theta = T(\Psi)$, defined above is*

$$f^*(\theta) = 2\pi g(2\pi F(\theta - \xi) - \pi) f(\theta - \xi), \ \theta \in [-\pi, \pi). \tag{4}$$

For simplicity, we assume that $\xi = 0$ except when otherwise stated. We shall refer to the method (4) of constructing a circular distribution as the inverse distribution function (idf) transformation and call it a circular Fg distribution. F, in Fg, represents the circular d.f. used in the idf transformation and g represents the base density function of circular r.v. Ψ.

If F in the idf transformation is the continuous circular uniform distribution function, then Θ and Ψ are identically distributed and if Ψ is the circular uniform r.v., then Θ is a circular r.v. with d.f. F.

The idf transformation method can generally be used to obtain new skew-symmetric circular distributions from base symmetric circular distributions. It is interesting to note that our circular Fg distribution may be viewed as the conditional distribution of Wehrly and Johnson [23]. However, here our results and study goes beyond their arguments.

2.1. General properties of circular Fg distributions

In this section, we consider certain basic properties of any Fg distribution with density (4). Specifically, we provide their distribution function, trigonometric moments and a random variate generation method. We also provide a result for finite mixtures and conditions for maximum entropy of the distribution.

Clearly, the distribution function of the density (4) is given by

$$F^*(\Theta) = p(\Theta \leq \theta) = p(\Psi \leq T^{-1}(\theta)) = G(2\pi F(\theta) - \pi), \tag{5}$$

where G denotes the d.f. of g. Trigonometric moments of f^* can be expressed in terms of expected values of g. That is, we have $E\{h(\Theta)\} = E_g\left\{h\left(F^{-1}\left(\frac{\Psi+\pi}{2\pi}\right)\right)\right\}$ where $F^{-1}(\theta) = \inf\{u : F(u) \geq \theta\}$. Thus, $E(h(2\pi F(\Theta) - \pi)) = E_g(h(\Psi))$. In particular, we have $E\{-\cos(2\pi F(\Theta))\} = E_g\{\cos(\Psi)\}$ and $E\{-\sin(2\pi F(\Theta))\} = E_g\{\sin(\Psi)\}$.

Consider the density of a mixture $pf_1^*(\theta) + (1-p)f_2^*(\theta)$ of two Fg distributions with densities f_1^* and f_2^* in which $f_1 = f_2 = f$, where $0 \leq p \leq 1$ is the mixing probability. Then, we have

$$\begin{aligned} pf_1^*(\theta) + (1-p)f_2^*(\theta) &= 2\pi\{pg_1(2\pi F(\theta) - \pi)f(\theta) + (1-p)g_2(2\pi F(\theta) - \pi)f(\theta)\}, \\ &= 2\pi\{pg_1(2\pi F(\theta) - \pi) + (1-p)g_2(2\pi F(\theta) - \pi)\}f(\theta). \end{aligned}$$

Thus, the resulting mixture is also a Fg distribution where g is a mixture of g_1 and g_2.

A basic algorithm for random variate generation of a Fg distribution is immediate using the distribution function (5): generate U from the uniform distribution and set $\Theta = F^{-1}((G^{-1}(U) + \pi)/2\pi) \pmod{2\pi}$.

2.2. Maximum entropy

The concept of entropy arose in Shannon [21] when attempting to create a theoretical model for the transmission of information. The entropy of a circular distribution with density $f(\theta) > 0$ over$[-\pi, \pi)$ is given by $-\int_{-\pi}^{\pi} log(f(\theta))f(\theta)d\theta$. Distributions maximizing the entropy often have important properties. The following theorem gives conditions for recognizing certain families of maximum entropy distributions.

Theorem 2.3. *Let f and g be circular densities and there exist 2π-periodic functions $w_i(\theta)$; $i = 1, ..., n$; and $e_j(\theta)$, $j = 1, ..., m$; and constants c_{1i}; $i = 1, ..., n$; and c_{2j}; $j = 1, ..., m$; such that*

a) $\int_{-\pi}^{\pi} w_i(F^{-1}(\theta))g(\theta - \pi)d\theta = c_{1i}$; $i = 1, ..., n$,
b) $\int_{-\pi}^{\pi} e_j(\theta)g(\theta - \pi)d\theta = c_{2j}$; $j = 1, ..., m$.

Then, the entropy is maximized when the circular density is of the form $f^*(\theta) = 2\pi g(2\pi F(\theta) - \pi)f(\theta)$, where

$$f(\theta) = e^{\sum_{i=1}^{n} \lambda_{1i} w_i(\theta)} \quad \text{and} \quad g(\theta) = e^{c_0 + \sum_{j=1}^{m} \lambda_{2j} e_j(\theta)}, \tag{6}$$

provided that there exist $c_0, \lambda_{11}, \lambda_{12}, ..., \lambda_{1n}$ and $\lambda_{21}, \lambda_{22}, ..., \lambda_{2m}$ such that a) and b) hold.

Proof. Theorem 3.1 in Mardia [15] states that a circular distribution which maximizes the entropy subject to $\int_{-\pi}^{\pi} w_i(\theta)f(\theta)d\theta = c_i$, $i = 1, ..., n$, is of the form $f(\theta) = e^{c_0 + \sum_{i=1}^{n} \lambda_i w_i(\theta)}$. Thus, the circular density

$$f^*(\theta) = e^{c_0 + \sum_{i=1}^{n} \lambda_{1i} w_i(\theta) + \sum_{j=1}^{m} \lambda_{2j} e_j(F(\theta))}$$

maximizes the entropy subject to a) and b). □

2.3. Möbius transformation as a special case of the idf transformation

Möbius transformation maps the unit circle onto itself. The general form of a Möbius transformation from circle $|z| = 1$ to circle $|z'| = 1$ is given by $z' = a\frac{z+b}{1+z\bar{b}}$, where $|a| = 1$, $|b| < 1$ and $\bar{b}$ is the conjugate of b. That is, the Möbius transformation from, say, $\Psi \to \Theta$ is given by

$$e^{i\Theta} = e^{i\mu}\frac{e^{i\Psi} + re^{i\nu}}{re^{i\Psi - i\nu} + 1} \quad or \quad \Theta = T(\Psi) = \nu + \mu + 2\arctan\left[\frac{1-r}{1+r}\tan\left(\frac{\Psi - \nu}{2}\right)\right],$$

where $0 \leq \mu, \nu < 2\pi$ and $0 \leq r < 1$. Since,

$$T^{-1}(\Theta) = \nu + 2\arctan\left[\frac{1+r}{1-r}\tan\left(\frac{\Theta - \gamma}{2}\right)\right], \quad \gamma = \nu + \mu,$$

the function T satisfies conditions of Theorem 2.1 and so the corresponding density of $\Theta = T(\Psi)$ is given by

$$f_\Theta(\theta) = f_\Psi(T^{-1}(\theta))\frac{1-r^2}{1+r^2-2r\cos(\theta-\gamma)}.$$

This is exactly equivalent to the idf transformation when we consider $T^{-1}(\theta) = 2\pi F(\theta-\gamma)+\nu-\pi$ and F is the circular distribution function of the wrapped Cauchy distribution. Thus, Möbius transformation is a special case of the idf transformation.

2.4. Modality property

Since there are no restrictions in choosing F and g in Corollary 2.2, we can not make a general comment on the modality of a Fg distribution. However, in the next theorem, we shall discuss the modality property of Fg distributions under certain conditions.

Theorem 2.4. *Let f and g be symmetric and unimodal circular densities at $\theta = 0$ and their first derivatives exist and are continuous. Then, $f^*(\theta) = 2\pi g(2\pi F(\theta)-\pi)f(\theta)$ is also symmetric and unimodal at $\theta = 0$, where F is the d.f. of f.*

Proof. f^* is symmetric, because

$$\begin{aligned} f^*(-\theta) &= 2\pi g(2\pi F(-\theta)-\pi)f(-\theta) \\ &= 2\pi g(2\pi(1-F(\theta))-\pi)f(\theta) = 2\pi g(\pi-2\pi F(\theta))f(\theta) = f^*(\theta). \end{aligned}$$

Differentiating the density f^* with respect to θ, one obtains

$$f^{*'}(\theta) = 2\pi\left\{2\pi f^2(\theta)g'(2\pi F(\theta)-\pi)+f'(\theta)g(2\pi F(\theta)-\pi)\right\}. \tag{7}$$

Since f and g are symmetric and unimodal at $\theta = 0$, they have antimodes at $\theta = -\pi$. Thus, we have $f'(0) = f'(-\pi) = 0$ and $g'(0) = g'(-\pi) = 0$. Hence, $f^{*'}(\theta) = 0$ at $\theta = 0$ and $\theta = -\pi$. Since $F(0) = 1/2$, we have $F(\theta) < 1/2$, for $-\pi < \theta < 0$, or $2\pi F(\theta)-\pi < 0$. Hence, $g'(2\pi F(\theta)-\pi) > 0$ and $f'(\theta) > 0$ for $-\pi < \theta < 0$. Consequently, it follows that $f^{*'}(\theta) > 0$. Similarly, we obtain $f^{*'}(\theta) < 0$ for $0 < \theta < \pi$. Hence, $f^*(\theta)$ has a mode at $\theta = 0$ and an antimode at $\theta = -\pi$. As required. □

3. An Alternative Special Case

Different choices of F and g provide wide classes of Fg models with various structural properties. To achieve mathematical tractability of Fg densities, we use another closed-form circular d.f. F for the idf transformation. That is, we consider the cardioid distribution function, i.e., $F(\theta) = (\pi+\theta+\lambda\sin(\theta))/2\pi$, $\theta \in [-\pi,\pi)$ and $\lambda \in [0,1]$ as F in the idf transformation. The cardioid density function is $f(\theta) = \frac{1}{2\pi}(1+\lambda\cos(\theta))$, where $\theta \in [-\pi,\pi)$. Consequently, we produce a class of distributions by the idf transformation whose densities are given by

$$f^*(\theta) = (1+\lambda\cos(\theta))g(\theta+\lambda\sin(\theta)),\ \theta \in [-\pi,\pi). \tag{8}$$

Clearly, choosing $\lambda = 0$ leads to the original distribution g. In the following, we consider that $g(\theta)$ is a symmetric and unimodal circular density so that for a function g_0 it can be written as $g(\theta) = g_0(\cos(\theta-\tau))$, $\theta \in [-\pi,\pi)$, where $-\pi \le \tau \le \pi$ is a location parameter. Notice that this is not a major restriction. Indeed, this is the case in many known circular distributions such as cardioid, power-of-cosine, wrapped Cauchy, von Mises and Jones-Pewsey circular distributions. Thus, we obtain

$$f^*(\theta) = (1+\lambda\cos(\theta))g_0(\cos(\theta+\lambda\sin(\theta)-\tau)). \tag{9}$$

Note that, here, $-\pi < \tau \le \pi$ is not a location parameter of f^*. Detailed results are presented below for the modality and shape properties of the circular densities (9).

3.1. Modality

Now assume that the first derivative of $g(\theta)$, $g'(\theta)$, exists and is continuous so that $g'(\tau) = 0$. Then, $g(\theta)$ has a mode at $\theta = \tau$ and an antimode at $\theta = \tau - \pi$. Clearly, the function

$$g_\lambda(\theta) = g(\theta + \lambda \sin(\theta)) = g_0(\cos(\theta + \lambda \sin(\theta) - \tau)) \tag{10}$$

is symmetric about τ if $\lambda = 0$. Further, trivially the function

$$H_\lambda(\theta) = \theta + \lambda \sin(\theta) : [-\pi, \pi] \to [-\pi, \pi] \tag{11}$$

is monotonically increasing for different values of $\lambda \in [0, 1]$ (see, Fig. 1). Thus, $H_\lambda^{-1}(\theta)$ is defined uniquely and we have the following results.

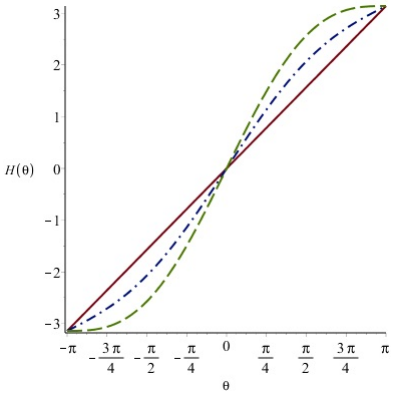

Figure 1: Function $H_\lambda(\theta)$ for $\lambda = 0$ (solid curve); $\lambda = 0.5$ (dash−dotted curve); $\lambda = 1$ (dashed curve).

Lemma 3.1. *The function (10) is unimodal with a mode at $\theta = H_\lambda^{-1}(\tau)$ and an antimode at $\theta = H_\lambda^{-1}(\tau - \pi)$. The value of the function at the mode equals $g(0)$.*

Proof. The derivative of $g(H_\lambda(\theta))$ with respect to θ is

$$g'_\lambda(\theta) = -\sin(H_\lambda(\theta) - \tau)(1 + \lambda \cos(\theta)) {g'}_0(\cos(H_\lambda(\theta) - \tau)).$$

Since $g(\theta)$ is unimodal, it follows that we must have ${g'}_0(\cos(\theta)) > 0, \forall\theta$. Thus, $g'_\lambda(\theta) = 0$ if either $\sin(H_\lambda(\theta) - \tau) = 0$ or $1 + \lambda \cos(\theta) = 0$. Since $\sin(H_\lambda(\theta) - \tau)$ is 0 at $\theta = H_\lambda^{-1}(\tau - \pi)$ and $\theta = H_\lambda^{-1}(\tau)$, it follows that $g'(H_\lambda(\theta))$ is non-zero for all $\theta \neq H_\lambda^{-1}(\tau - \pi)$ and $\theta \neq H_\lambda^{-1}(\tau)$. Therefore, as $g_\lambda(H_\lambda^{-1}(\tau - \pi)) < g_\lambda(H_\lambda^{-1}(\tau))$, $g_\lambda(\theta)$ is unimodal with a mode at $M_\lambda = H_\lambda^{-1}(\tau)$ and an antimode at $M_\lambda^* = H_\lambda^{-1}(\tau - \pi)$. The value of the function at the mode is $g_\lambda(M_\lambda) = g(0)$. □

Theorem 3.2. *For the density (9) we have:*

a) If $\tau = 0$, then f^* is symmetric and unimodal at $\theta = 0$,
b) If $0 < |\tau| < \pi$, then f^* is either unimodal or bimodal,
c) If $\tau = \pi$, then f^* is symmetric and unimodal with a mode and an antimode at either $\theta = -\pi$ and $\theta = 0$ or f^* is symmetric and bimodal with two diametrically opposed modes.

Proof. **a)** When $\tau = 0$, g is also symmetric and unimodal. Thus, according to Theorem 2.4, $f^*(\theta)$ is symmetric and unimodal at $\theta = 0$.

b) First, we suppose $0 < \tau < \pi$. Differentiating the density (9) with respect to θ, one obtains

$$f^{*'}(\theta) = -\lambda \sin(\theta) g_0(\cos(H_\lambda(\theta) - \tau)) - \sin(H_\lambda(\theta) - \tau)(1 + \lambda \cos(\theta))^2 {g'}_0(\cos(H_\lambda(\theta) - \tau)). \tag{12}$$

The number of solutions of $f^{*'}(\theta) = 0$ equals to the points of intersection of the functions

$$\gamma(\theta) = -\frac{g'_0(\cos(H_\lambda(\theta) - \tau))}{g_0(\cos(H_\lambda(\theta) - \tau))}, \tag{13}$$

and

$$k(\theta) = \frac{\lambda \sin(\theta)}{(1 + \lambda\cos(\theta))^2 \sin(H_\lambda(\theta) - \tau)}.$$

Since $g_0(\cos(\theta))$ is symmetric and unimodal with mode at $\theta = 0$ and antimode at $\theta = -\pi$, we have $g'_0(\cos(\theta)) > 0$. Since $g_0(y) > 0$ and $g'_0(y) > 0$ for $-1 \le y \le 1$, therefore, $g''_0(y)$ is always negative, always positive or always zero for $-1 \le y \le 1$. Thus, $g''_0(\cos(\theta)), \forall\theta$, is always negative, always positive or always zero. When $g''_0(\cos(\theta)) > 0 (< 0)$, $g'_0(\cos(\theta))$ is a unimodal function with mode at $\theta = 0(-\pi)$ and antimode at $\theta = -\pi(0)$. When $g''_0(\cos(\theta)) = 0$, $g'_0(\cos(\theta))$ is a constant function. Thus, by lemma 3.1, $g'_0(\cos(H_\lambda(\theta) - \tau))$ has a mode at $\theta_1^* = H_\lambda^{-1}(\tau - \pi)$ or $\theta_2^* = H_\lambda^{-1}(\tau)$. Consequently, function $\gamma(\theta)$ has a mode and an antimode at either θ_1^* or θ_2^* and it is negative. $k(\theta)$ has two vertical asymptotes at θ_1^* and θ_2^*, i.e.,

$$\lim_{\theta\to\theta_2^{*+}} k(\theta) = \lim_{\theta\to\theta_1^{*+}} k(\theta) = +\infty,$$
$$\lim_{\theta\to\theta_2^{*-}} k(\theta) = \lim_{\theta\to\theta_1^{*-}} k(\theta) = -\infty.$$

Also, $k(\theta) < 0$ for $\theta \in (-\pi, \theta_1^*) \cup (0, \theta_2^*)$ and $k(\theta) > 0$ for $\theta \in (\theta_1^*, 0) \cup (\theta_2^*, \pi)$. Since $\gamma(\theta) < 0$, it is sufficient to find the points of intersection $k(\theta)$ and $\gamma(\theta)$ in the interval $(-\pi, \theta_2^*)$. To do this, we need to find extremum points of $k(\theta)$ in $(-\pi, \theta_2^*)$.

The number of roots of $k'(\theta) = 0$ equals to the number of solutions of

$$-\frac{\lambda\cos(\theta)^2 - \cos(\theta) - 2\lambda}{\sin(\theta)(1 + \lambda\cos(\theta))^2} = \frac{\cos(H_\lambda(\theta) - \tau)}{\sin(H_\lambda(\theta) - \tau)}.$$

But, $\cot(H_\lambda(\theta) - \tau)$ is zero at $z_1 = H^{-1}(\tau - \pi/2)$ and $z_2 = H^{-1}(\tau + \pi/2)$ and has two vertical asymptotes at θ_1^* and θ_2^*, i.e.,

$$\lim_{\theta\to\theta_2^{*-}} \cot(H_\lambda(\theta) - \tau) = \lim_{\theta\to\theta_1^{*+}} \cot(H_\lambda(\theta) - \tau) = +\infty,$$
$$\lim_{\theta\to\theta_2^{*+}} \cot(H_\lambda(\theta) - \tau) = \lim_{\theta\to\theta_1^{*-}} \cot(H_\lambda(\theta) - \tau) = -\infty.$$

On the other hand, function

$$\varphi(\theta) = -\frac{\lambda\cos(\theta)^2 - \cos(\theta) - 2\lambda}{\sin(\theta)(1 + \lambda\cos(\theta))^2},$$

has two real roots at $\theta_1 = -\pi + \arccos\left(\frac{-1+\sqrt{8\lambda^2+1}}{2\lambda}\right)$ and $\theta_2 = -\theta_1 \in [\pi/2, \pi]$. Also, $\varphi(\theta)$ has three vertical asymptotes at $-\pi$, 0 and π, i.e.,

$$\lim_{\theta\to 0^-} \varphi(\theta) = \lim_{\theta\to\pi^-} \varphi(\theta) = -\infty,$$
$$\lim_{\theta\to 0^+} \varphi(\theta) = \lim_{\theta\to-\pi^+} \varphi(\theta) = +\infty.$$

Since $\theta_1 < z_1$, $\varphi(\theta)$ and $\cot(H_\lambda(\theta) - \tau)$ do not intersect in $[\theta_1^*, \theta_2^*]$. Hence, $k'(\theta) < 0$ and thus $k(\theta)$ strictly decreases in $[\theta_1^*, \theta_2^*]$. Therefore, $f^{*'}(\theta) = 0$ has only one root in $[\theta_1^*, \theta_2^*]$. We observe that $\varphi(\theta)$ is either strictly decreasing or has 2 extremum points in the interval $(\theta_1, 0)$. When $\varphi(\theta)$ is strictly decreasing in $(\theta_1, 0)$, $k'(\theta) < 0$. Therefore, $f^{*'}(\theta) = 0$ has only one root in (θ_1, θ_1^*) and two roots in $[\theta_1, \theta_2^*]$. Consequently, $f^*(\theta)$ is unimodal.

Now, let us assume $\varphi(\theta)$ has a minimum and a maximum in the interval $(\theta_1, 0)$. Since, $\cot(H_\lambda(\theta) - \tau)$ is strictly decreasing, $\varphi(\theta)$ and $\cot(H_\lambda(\theta) - \tau)$ have two points of intersection and therefore there are two solutions for $k'(\theta) = 0$ in (θ_1, θ_1^*). Thus, $k(\theta)$ has a minimum and a maximum in (θ_1, θ_1^*). Since $k(-\pi) = 0$, $k(\theta)$ and $\gamma(\theta)$ have three points of intersection in (θ_1, θ_1^*), $f^{*'}(\theta)$ has four roots in $[\theta_1, \theta_2^*]$. Consequently, $f^*(\theta)$ is bimodal.
Hence, $f^*(\theta)$ is either unimodal or bimodal for $0 < \tau < \pi$ and since $f^*(-\theta; \lambda, \tau) = f^*(\theta; \lambda, -\tau)$, $f^*(\theta)$ is either unimodal or bimodal for $-\pi < \tau < 0$.

c) For $\tau = \pi$, $f^*(\theta)$ is symmetric about $\theta = 0$ and thus, $\theta = 0$ is a mode or an antimode. From (13), we also have that $\gamma(\theta)$ is symmetric about $\theta = 0$ and unimodal at $\theta = -\pi$ or $\theta = 0$. Further, at $\tau = \pi$, $k(\theta)$ is also symmetric about $\theta = 0$ and $k'(\theta)$ is positive in $(-\pi, 0)$, i.e., $k(\theta)$ is strictly increasing in $(-\pi, 0)$. Thus, $\gamma(\theta)$ and $k(\theta)$ have maximum one point of intersection in $(-\pi, 0)$. When $\gamma(\theta)$ and $k(\theta)$ do not intersect, $f^*(\theta)$ is unimodal. When $\gamma(\theta)$ and $k(\theta)$ have one point of intersection, $f^*(\theta)$ is bimodal with diametrically opposed modes. □

When $\tau = 0$, the density (8) is symmetric and unimodal at $\theta = 0$. When $\tau = \pi$, $f^*(\theta) = (1 + \lambda\cos(\theta))g_0(-\cos(\theta + \lambda\sin(\theta)))$ is symmetric about zero and it is unimodal at $\theta = -\pi$ or $\theta = 0$ and bimodal with two diametrically opposed modes. While, for example, if $\tau = \pm\pi/2$, we have $\cos(\theta + \lambda\sin(\theta) \mp \pi/2) = \pm\sin(\theta + \lambda\sin(\theta))$ and, thus, the density function $f^*(\theta)$ is asymmetric. Hence, for $\lambda \neq 0$ in cases $\tau \neq \pi$ and $\tau \neq 0$ the ensuing distributions are asymmetric. It is important to note that if $\lambda = 0$, ξ and τ are non-identifiable.

3.2. Shape properties

In addition to the assumptions of Sec. 2.4, here we further assume that the second derivative of $g(\theta)$, $g''(\theta)$, exists and is continuous. The following theorem contains useful conditions under which the peakedness of the density (8) can be controlled.

Theorem 3.3. *For $\lambda_1 < \lambda_2$ in the density (9), we have:*

a) $f^*_{\lambda_1}(0) < f^*_{\lambda_2}(0)$ *and* $f^*_{\lambda_1}(-\pi) > f^*_{\lambda_2}(-\pi)$, *if* $\tau = 0$,
b) $f^{*''}_{\lambda_1}(0) > f^{*''}_{\lambda_2}(0)$ *and* $f^{*''}_{\lambda_1}(-\pi) > f^{*''}_{\lambda_2}(-\pi)$, *if* $\tau = 0$,
c) $f^*_{\lambda_1}(M_{\lambda_1}) < f^*_{\lambda_2}(M_{\lambda_2})$, *if* $|\tau| \in (0, \pi/2]$,

where M_{λ_i}'s are the modes (the major mode in the bimodal case) corresponding to density functions $f^*_{\lambda_i}$, $i = 1, 2$.

Proof. **a)** The assertion follows because

$$f^*_{\lambda_1}(0) = (1 + \lambda_1)g(0) < (1 + \lambda_2)g(0) = f^*_{\lambda_2}(0)$$

and

$$f^*_{\lambda_1}(-\pi) = (1 - \lambda_1)g(-\pi) > (1 - \lambda_2)g(-\pi) = f^*_{\lambda_2}(-\pi).$$

b) Since for $\tau = 0$, $g(\theta)$ is symmetric and unimodal at $\theta = 0$, we have $g'(0) = 0$ and $g''(0) < 0$. Now,

$$f^{*''}_\lambda(\theta) = (1 + \lambda\cos(\theta))^3 g''(H_\lambda(\theta)) - 3\lambda\sin(\theta)(1 + \lambda\cos(\theta))g'(H_\lambda(\theta)) - \lambda\cos(\theta)g(H_\lambda(\theta)).$$

Thus, we have

$$f^{*''}_{\lambda_1}(0) = g''(0)(1 + \lambda_1)^3 - \lambda_1 g(0) > g''(0)(1 + \lambda_2)^3 - \lambda_2 g(0) = f^{*''}_{\lambda_2}(0).$$

Similarly, it follows that $f^{*''}_{\lambda_1}(-\pi) > f^{*''}_{\lambda_2}(-\pi)$.

c) First, we suppose $\tau \in (0, \pi/2]$. Since the function $g_{\lambda_i}(\theta)$ has a mode at $\theta^*_{\lambda_i} = H^{-1}_{\lambda_i}(\tau)$, where H_{λ_i} is defined in (11) and $\theta^*_{\lambda_i} \in (0, \tau)$, $i = 1, 2$, for $\tau \in (0, \pi/2]$ and $1 + \lambda_i\cos(\theta)$ are decreasing on $(0, \pi]$, we have the (major) modes $M_{\lambda_i} \in (0, \theta^*_{\lambda_i})$, $i = 1, 2$. We have $\theta^*_{\lambda_2} < \theta^*_{\lambda_1}$. Now, there are two following situations:

(i) $\theta^*_{\lambda_2} < M_{\lambda_1} < \theta^*_{\lambda_1}$; in this case since $\tau \in (0, \pi/2]$ we have

$$\begin{aligned} f^*_{\lambda_1}(M_{\lambda_1}) &= (1 + \lambda_1 \cos(M_{\lambda_1}))g_{\lambda_1}(M_{\lambda_1}) < (1 + \lambda_2 \cos(\theta^*_{\lambda_2}))g_{\lambda_1}(M_{\lambda_1}) \\ &< (1 + \lambda_2 \cos(\theta^*_{\lambda_2}))g_{\lambda_1}(\theta^*_{\lambda_1}) = (1 + \lambda_2 \cos(\theta^*_{\lambda_2}))g_{\lambda_2}(\theta^*_{\lambda_2}) = f^*_{\lambda_2}(\theta^*_{\lambda_2}) \\ &< f^*_{\lambda_2}(M_{\lambda_2}). \end{aligned}$$

The second equality follows by lemma 3.1 because the values of the functions g_{λ_i}, $i = 1, 2$, at the modes are the same, i.e., $g_{\lambda_1}(\theta^*_{\lambda_1}) = g_{\lambda_2}(\theta^*_{\lambda_2}) = g(\tau)$.

(ii) $M_{\lambda_1} < \theta^*_{\lambda_2} < \theta^*_{\lambda_1}$; in this case

$$\begin{aligned} f^*_{\lambda_1}(M_{\lambda_1}) &= (1 + \lambda_1 \cos(M_{\lambda_1}))g_{\lambda_1}(M_{\lambda_1}) < (1 + \lambda_2 \cos(M_{\lambda_1}))g_{\lambda_1}(M_{\lambda_1}) \\ &< (1 + \lambda_2 \cos(M_{\lambda_1}))g_{\lambda_2}(M_{\lambda_1}) = f^*_{\lambda_2}(M_{\lambda_1}) < f^*_{\lambda_2}(M_{\lambda_2}). \end{aligned}$$

Note that, since $H_{\lambda_i}(M_{\lambda_i}) \in (0, \tau)$; $i = 1, 2$; for $\tau \in (0, \pi/2]$ and $H_{\lambda_1}(M_{\lambda_1}) < H_{\lambda_2}(M_{\lambda_1})$, we have $g_{\lambda_1}(M_{\lambda_1}) = g(H_{\lambda_1}(M_{\lambda_1})) < g(H_{\lambda_2}(M_{\lambda_1})) = g_{\lambda_2}(M_{\lambda_1})$.

Thus, $f^*_{\lambda_1}(M_{\lambda_1}) < f^*_{\lambda_2}(M_{\lambda_2})$ for $\tau \in (0, \pi/2]$ and since $f^*(-\theta; \lambda, \tau) = f^*(\theta; \lambda, -\tau)$, we have $f^*_{\lambda_1}(M_{\lambda_1}) < f^*_{\lambda_2}(M_{\lambda_2})$ for $\tau \in [-\pi/2, 0)$. □

Theorem 3.3 clearly implies that in the unimodal case λ is a shape parameter which controls the peakedness properties of the density f^*.

Corollary 3.4. *For $\tau = 0$, $f^*_{\lambda_2}$ is more sharply peaked than $f^*_{\lambda_1}$ and also $f^*_{\lambda_1}$ has heavier tails than $f^*_{\lambda_2}$.*

Proof. The curvature of $f^*_\lambda(\theta)$ at θ equals $\frac{f^{*''}_\lambda(\theta)}{(1+\{f^{*'}_\lambda(\theta)\}^2)^{3/2}}$. Since $f^{*'}_\lambda(0) = 0$, the curvature of $f^*_\lambda(0)$ is equal to $f^{*''}_\lambda(0)$. Thus, $f^*_{\lambda_1}(0)$ is more sharply peaked than $f^*_{\lambda_2}(0)$. Since $f^*_{\lambda_1}(-\pi) > f^*_{\lambda_2}(-\pi)$, $f^*_{\lambda_1}$ has heavier tails than $f^*_{\lambda_2}$. □

To illustrate the shape flexibility of the class of cardioid-g, we consider von Mises, wrapped Cauchy and cardioid distributions as the g distribution and portray cardioid-von Mises, cardioid-wrapped Cauchy and cardioid-cardioid densities in Fig. 2. The left-hand panels of Fig. 2 show the effect of the change of the values of the parameter λ, with the other parameters kept constant when distributions are symmetric and unimodal. Note how the peakedness of the distributions increase with the increase of the values of the parameter λ. Also, in the bimodal case, values of the density at major modes increase with the increase of the parameter λ. In comparing the cases $\tau = 0$, $\tau = \pi/2$ and $\tau = 2$, it is observed how change of values of parameter τ effects the skewness of the distribution. In the next section, we discuss some properties of the cardioid-von Mises case.

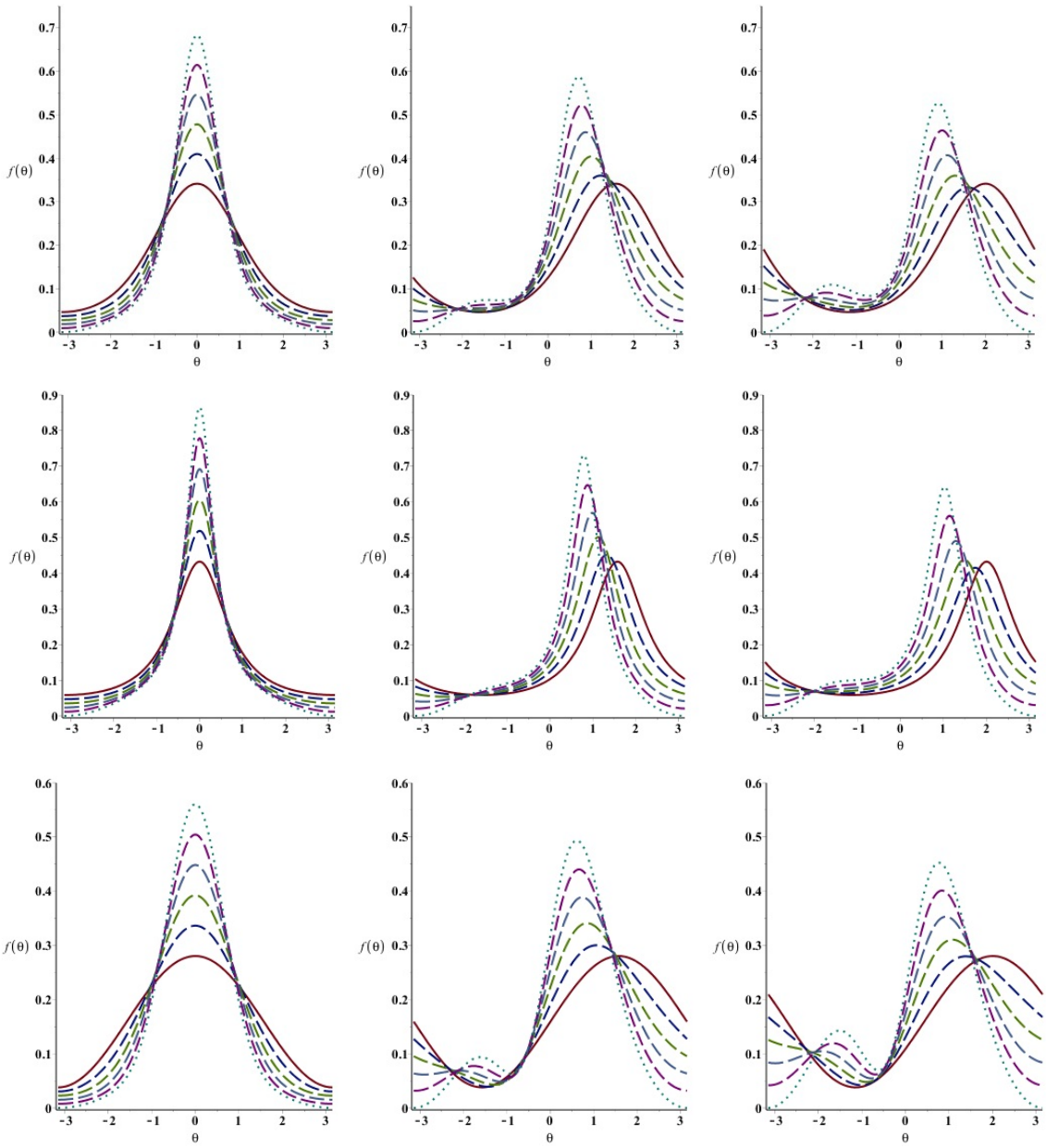

Figure 2: Densities of cardioid-von Mises (CvM) for $\kappa = 1$ (top row), cardioid-wrapped Cauchy (CwC) for $\rho = 0.46$ (middle row) and cardioid-cardioid (CC) for $\rho = 0.38$ (bottom row) distributions for various parameter values. In the first column $\tau = 0$, in the second column $\tau = \pi/2$ and in the third column $\tau = 2$. In each panel, $\lambda = 0$ (solid); $\lambda = 0.2(0.2)0.8$ (dashed); $\lambda = 1$ (dotted).

4. Cardioid-von Mises Distribution and its Special Cases

In this section, we consider the von Mises distribution for the density g, which is a unimodal symmetric circular distribution and plays a central role in the analysis of circular data. The von Mises distribution has density

$$f_{vM}(\theta) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \xi)}, \tag{14}$$

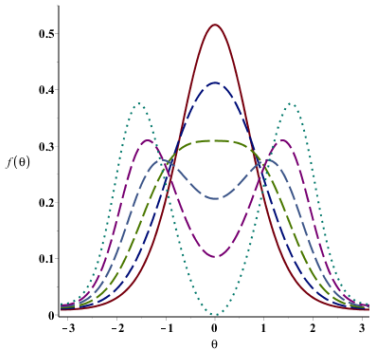

Figure 3: The CvM densities for $\xi = \tau = \pi$, $\kappa = 2$ and $\lambda = 0$ (solid); $\lambda = 0.2(0.2)0.8$ (dashed); $\lambda = 1$ (dotted).

where $\kappa \geq 0$ is a concentration parameter and $I_r(\kappa)$ is the modified Bessel function of the first kind of order r, defined by

$$I_r(\kappa) = \frac{1}{2\pi}\int_0^{2\pi} \cos(r\theta)e^{\kappa\cos(\theta)}d\theta, \quad r = 0, \pm1, \pm2, \ldots .$$

Substituting density (14) for g in (9), the density of the Cardioid-von Mises (CvM) distribution is given by

$$f^*_{CvM}(\theta) = \frac{1 + \lambda\cos(\theta - \xi)}{2\pi I_0(\kappa)} e^{\kappa\cos(\theta-\xi+\lambda\sin(\theta-\xi)-\tau)}, \quad -\pi \leq \theta < \pi, \tag{15}$$

where $0 \leq \lambda \leq 1$ and $-\pi < \tau \leq \pi$.

As Figures 3 and 2 show, the CvM distribution becomes symmetric and asymmetric for certain values of the parameters τ: for $\tau = 0$ it is symmetric and unimodal, for $\tau = \pm\pi/2$ it is asymmetric and unimodal or bimodal and for $\tau = \pi$ it is symmetric and unimodal or bimodal with two diametrically opposed modes. We shall refer to these submodels as symmetric cardioid-von Mises (SCvM), asymmetric cardioid-von Mises (ACvM) and bipolar cardioid-von Mises (BCvM) distributions, respectively. Using these submodels we can investigate the effect of the τ parameter on the model fitting in Sec. 5. The flexibility of such distributions is illustrated in the first row of Fig. 2 and Fig. 3. In the unimodal case, the top row in Fig. 2 shows that CvM distribution has more peakedness than vM distribution and Fig. 3 shows that CvM distribution has more flatness than vM distribution. In Fig. 4, contour plots of the circular skewness and circular kurtosis (see Sec. 4.1 for a general definition) are portrayed as functions of κ and λ, for the CvM family.

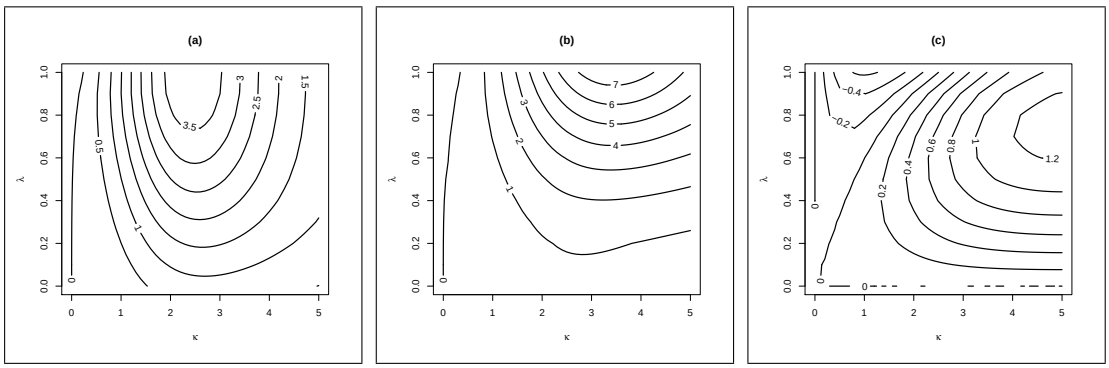

Figure 4: Contour plots of (a) the circular kurtosis ($\tau = 0$) and (b) the circular kurtosis ($|\tau| = \pi/2$) and (c) the circular skewness ($|\tau| = \pi/2$) , as functions of κ and λ, for CvM family.

According to Theorem 3.2, the CvM distribution becomes bimodal for certain values of the parameters κ, λ and τ. But, they cannot be found algebraically. A numerical solution is obtained by solving for the

number of roots of the derivative of the density CvM. This work is well done with the **rootSolve** package in R software. Fig. 5 shows a plot of the boundary region of λ and κ values for $\tau = \pi$, $|\tau| = 1$ and $|\tau| = 2$. As seen in Fig. 5, as the parameter $|\tau|$ decreases from $\tau = \pi$, regions of the parameter space of λ and κ wherein the density CvM is unimodal increase. Finally, when $\tau = 0$, CvM is unimodal for all the parameter space.

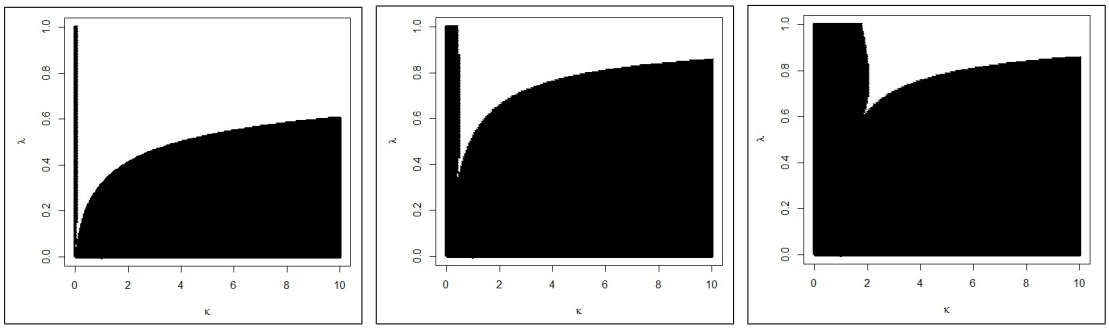

Figure 5: Plot of bimodal region for CvM distribution for $\tau = \pi$, $|\tau| = 2$ and $|\tau| = 1$ from left to right, for some values of λ and κ. (Black and white areas are unimodal and bimodal regions, respectively).

Kato and Jones [12] used the Möbius transformation of a r.v. following the von Mises distribution. They built a new four-parameter family of circular distribution. As explained in Sec. 2.3, this new four-parameter family is the same as the wrapped Cauchy-von Mises distribution, i.e., when F is the wrapped Cauchy distribution and g is a von Mises density. We shall compare this distribution with CvM distribution in Sec. 5 by fitting both to the same real data set.

4.1. Other properties of the SCvM distribution

The characteristic function of a circular r.v. Θ is $\phi_p = E\left(e^{ip\Theta}\right) = \alpha_p + i\beta_p$, $p = 0, \pm1, \pm2, \ldots$ where $\alpha_p = E(\cos(p\Theta))$ and $\beta_p = E(\sin(p\Theta))$ are referred to as the pth cosine and sine moments, respectively. The pth trigonometric moments about the mean direction μ, i.e., $\mu = \mu_1 = \arg\{\alpha_1 + i\beta_1\}$, are defined by $\bar{\alpha}_p = E(\cos p(\Theta - \mu))$ and $\bar{\beta}_p = E(\sin p(\Theta - \mu))$. It follows that an alternative representation of ϕ_p is $\phi_p = (\bar{\alpha}_p + i\bar{\beta}_p)e^{i\mu p}$. Since the density of SCvM is symmetric about ξ, it follows that $\mu = \xi$ and $\bar{\beta}_p = 0$, and thus $\phi_p = \bar{\alpha}_p e^{i\mu p}$.

The pth central cosine moment is $\bar{\alpha}_p = C_p(\kappa, \lambda) + \frac{\lambda}{2}\left[C_{p-1}(\kappa, \lambda) + C_{p+1}(\kappa, \lambda)\right]$, where

$$C_p(\kappa, \lambda) = \int_{-\pi}^{\pi} \frac{\cos(p\theta)}{2\pi I_0(\kappa)} e^{\kappa\cos(\theta + \lambda\sin(\theta))} d\theta, \quad p = 0, 1, \ldots .$$

Consequently, $\phi_p = (C_p(\kappa, \lambda) + \frac{\lambda}{2}\left[C_{p-1}(\kappa, \lambda) + C_{p+1}(\kappa, \lambda)\right])e^{i\mu p}$, and the pth cosine and sine moments follow as

$$\alpha_p = \cos(p\mu)(C_p(\kappa, \lambda) + \frac{\lambda}{2}\left[C_{p-1}(\kappa, \lambda) + C_{p+1}(\kappa, \lambda)\right]),$$

$$\beta_p = \sin(p\mu)(C_p(\kappa, \lambda) + \frac{\lambda}{2}\left[C_{p-1}(\kappa, \lambda) + C_{p+1}(\kappa, \lambda)\right]).$$

The pth mean resultant length of a circular distribution is defined to be $\rho_p = \sqrt{\alpha_p^2 + \beta_p^2}$. Because of the symmetric property of SCvM's distribution, its pth mean resultant length is identical to the pth central cosine moment, i.e., $\rho_p = |\bar{\alpha}_p|$ and also $\mu_p = p\mu$. The circular variance and circular standard deviation are then given by

$$v = 1 - \rho = \frac{1}{2}[2 - \lambda(1 + C_2(\kappa, \lambda)) + (\lambda^2 - 2)C_1(\kappa, \lambda)]$$

and

$$\sigma = (-2\ln(1 - v))^{1/2} = (-2\ln 2(\lambda(1 + C_2(\kappa, \lambda)) + (2 - \lambda^2)C_1(\kappa, \lambda)))^{1/2}.$$

Using these results, the circular skewness and kurtosis are given by $s = \frac{\bar{\beta}_2}{(1-\rho)^{3/2}} = 0$ and $k = \frac{\bar{\alpha}_2 - \rho^4}{(1-\rho)^2}$.

4.2. Maximum Likelihood Estimation

In this section, we consider the maximum likelihood estimation for the vector parameter $\eta = (\xi, \kappa, \lambda, \tau)$ of the density (15). Let $\theta_1, ..., \theta_n$ be a random sample of size n from the distribution CvM. Then, the log-likelihood function for η is expressed as

$$\begin{aligned} l(\eta) = & - n \log(2\pi I_0(\kappa)) \\ & + \sum_{i=1}^{n} \log(1 + \lambda \cos(\theta_i - \xi)) + \kappa \sum_{i=1}^{n} \cos(\theta_i - \xi + \lambda \sin(\theta_i - \xi) - \tau). \end{aligned} \tag{16}$$

Taking partial derivatives from the log-likelihood function with respect to ξ, κ, λ and τ, respectively, and equating the resulting expressions to zero yield likelihood equations. The score equations and observed information matrices are provided in the Appendix. To solve these equations, it is usually more convenient to use nonlinear optimization methods to numerically maximize $l(\eta)$. Function `optim` in program R provides the non-linear optimization routine for solving such problems. We make use of **L-BFGS-B** optimization method in the following illustrative examples in Sec. 5. In the case of grouped data, suppose that there are m class intervals $(\theta_0, \theta_1), ..., (\theta_{m-1}, \theta_m)$, where $\theta_0 = -\pi$ and $\theta_m = \pi$. Denoting the number of data values in the jth class interval by n_j and thus a total of $n = n_1 + n_2 + ... + n_m$ observations, the log-likelihood function is given by

$$\begin{aligned} l(\eta) = & - n \log(2\pi I_0(\kappa)) \\ & + \sum_{j=1}^{m} \int_{\theta_{j-1}}^{\theta_j} \log(1 + \lambda \cos(\theta - \xi)) + \kappa \cos(\theta - \xi + \lambda \sin(\theta - \xi) - \tau) d\theta. \end{aligned}$$

In general, no closed-form expressions exist for the maximum likelihood estimates (MLEs) and so numerical methods must be used to identify them. The maximization of this function can also be achieved using the optimization options available in R referred to above.

4.3. Monte Carlo comparison

In this section, we study the performance and accuracy of the MLEs of the parameters of the CvM by conducting various simulations for different sample sizes and different parameter values. To generate data from CvM distribution, we can use the method mentioned in Sec. 2.1. In our Monte Carlo experiment, for each combination of $(\xi, \kappa, \lambda, \tau)$, with $n = 50, 100, 500$, we have simulated 3000 samples of size n from the density (15). CvM distribution with parameters $(\xi, \kappa, \lambda, \tau) = (0, 1, 0.9, 0.8), (0, 1, 0.5, 0.3), (0, 1, 0.9, 1.5)$ is symmetric, skew and bimodal, respectively. In our simulation study, mean bias and mean square error of the MLE of the parameters, defined below, are computed and discussed.

1. Mean bias (Bias) of the MLE of the parameter of interest ϵ (e.g., ξ, κ, λ or τ) is defined as:

$$Bias_{\epsilon}(n) = \frac{1}{N} \sum_{i=1}^{N} (\hat{\epsilon}_i - \epsilon).$$

2. Mean squared error (MSE) of the MLE of the parameter of interest ϵ is defined as:

$$MSE_{\epsilon}(n) = \frac{1}{N} \sum_{i=1}^{N} (\hat{\epsilon}_i - \epsilon)^2,$$

where $\hat{\epsilon}_i$ is the MLE of ϵ based on a sample of size n.

Table 1 illustrates Bias, and MSE values of $(\xi, \kappa, \lambda, \tau)$ for the different sample sizes. It can be concluded that as the sample size n increases, the Bias and MSE decay towards zero.

Table 1: Mean Bias (Bias) and Mean squared error (MSE) in parenthesis for the ML estimators calculated using 3000 samples of size n simulated from CvM distribution.

ξ	κ	λ	τ	n	$\hat{\xi}$	$\hat{\kappa}$	$\hat{\lambda}$	$\hat{\tau}$
0	1	0.9	0.8	50	-0.0155(0.037)	0.0461(0.089)	0.0093(0.009)	0.0124(0.097)
				100	-0.0016(0.015)	0.0269(0.042)	0.0035(0.005)	0.0079(0.041)
				500	-0.0002(0.002)	0.0076(0.008)	0.0009(0.0009)	-0.0002(0.008)
0	1	0.5	0.3	50	0.0038(0.167)	0.1204(0.136)	0.0360(0.033)	0.0056(0.081)
				100	0.0003(0.097)	0.0581(0.065)	0.0199(0.017)	0.0032(0.049)
				500	-0.0001(0.017)	0.0100(0.012)	0.0035(0.003)	0.0007(0.010)
0	1	0.9	1.5	50	-0.0316(0.051)	0.0662(0.133)	-0.0324(0.014)	0.0960(0.203)
				100	-0.0083(0.038)	0.0373(0.055)	-0.0002(0.014)	0.0241(0.129)
				500	-0.0001(0.006)	0.0008(0.010)	-0.00005(0.002)	-0.0006(0.020)
0	1	0.7	2	50	-0.0208(0.166)	0.1157(0.147)	0.0221(0.037)	0.0168(0.283)
				100	0.0078(0.089)	0.0527(0.063)	0.0101(0.019)	-0.0212(0.185)
				500	0.0023(0.010)	0.0083(0.011)	0.0012(0.003)	-0.0028(0.021)

5. Application

We consider a grouped data set of size $n = 15831$, which consists of times of gun crimes committed in Pittsburgh, Pennsylvania, recorded over the period 1987-1998 (Gill and Hangartner [9]). These data are recorded at the hourly level on the 24-h clock.

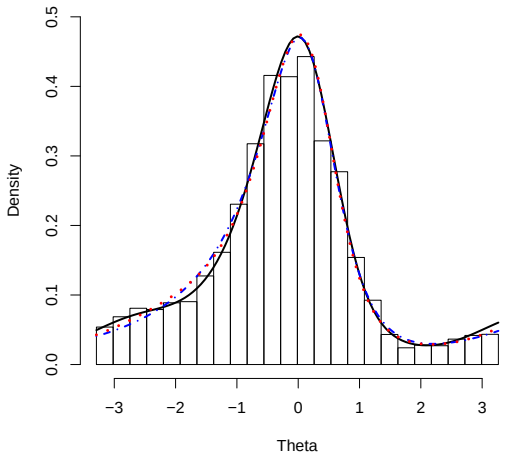

Figure 6: Histogram of the gun crimes data represented in radians, together with the densities of the maximum likelihood fits for the cardioid-von Mises (solid curve), sine-skewed Jones-Pewsey (dotted curve) and asymmetric extended Jones-Pewsey (dash−dotted curve) distributions.

In Fig. 6, a histogram of the data is illustrated. Similar to Kato and Jones [13], the data were converted from 24 hours to angles in $[-\pi, \pi)$; for clarity, $-\pi$ corresponds to midday, 0 to midnight, etc. The data are asymmetric and peaked around 22 to 1 hours. The test of Pewsey [18], with a p-value of 0.0000, rejects the underlying distribution as being symmetric emphatically. In Table 2, we present the maximum likelihood estimates, the values of the maximized log-likelihood (MLL), Akaike information criterion (AIC) and

Bayesian information criterion (BIC) obtained by fitting the cardioid-von Mises (CvM), asymmetric cardioid-von Mises (ACvM), sine-skewed Jones-Pewsey (SSJP) model of Abe and Pewsey [1], the asymmetric extended Jones-Pewsey (AEJP) distribution of Abe et al. [3], Full model Kato-Jones (FKJ) of Kato and Jones [13] and four-parameter Kato-Jones (KJ2010) model of Kato and Jones [12]. Because the fitted models are very similar in the plot, we show just three fitted models on the histogram. The fitted skew-symmetric densities are superimposed on the histogram of the data set in Fig. 6.

Table 2: Maximum likelihood estimates, the values of the maximized log-likelihood (MLL), Akaike information criterion (AIC) and Bayesian information criterion (BIC) for the fits to the gun crime data of the cardioid-von Mises (CvM), asymmetric cardioid-von Mises (ACvM), Full model Kato-Jones (FKJ), sine-skewed Jones-Pewsey (SSJP), Kato-Jones (KJ2010) and asymmetric extended Jones-Pewsey (AEJP)models.

Model					MLE	AIC	BIC
	$\hat{\xi}$	$\hat{\kappa}$	$\hat{\lambda}$	$\hat{\tau}$			
CvM	1.33	1.56	0.54	-2.13	−44676.94	89361.87	89392.55
	$\hat{\xi}$	$\hat{\kappa}$	$\hat{\lambda}$	τ			
ACvM	0.92	1.24	0.42	$(-\pi/2)$	−44797.96	89601.92	89624.93
	$\hat{\xi}$	$\hat{\kappa}$	$\hat{\psi}$	$\hat{\lambda}$			
SSJP	0.23	1.07	-0.81	-0.65	−44731.86	89471.71	89502.39
	$\hat{\xi}$	$\hat{\kappa}$	$\hat{\psi}$	$\hat{\nu}$			
AEJP	-2.04	1.40	-0.17	-0.56	−44720.72	89449.45	89480.12
	$\hat{\xi}$	$\hat{\gamma}$	$\hat{\alpha}_2$	$\hat{\beta}_2$			
FKJ	0.38	0.55	0.23	0.16	−44741.00	89490.00	89520.68
	$\hat{\mu}$	$\hat{\kappa}$	$\hat{\nu}$	$\hat{r}$			
KJ2010	0.89	1.40	1.90	0.29	−44730.24	89468.47	89499.15

From the results in Table 2, it can be seen that among all these three four-parameter distributions; the cardioid-von Mises distribution has the highest MLL and lowest AIC and BIC values. Thus, the cardioid-von Mises distribution provides a better fit to the data than the other three four-parameter skew models. Considering the MLL values for the cardioid-von Mises and asymmetric cardioid-von Mises distributions, the test statistic for the usual likelihood-ratio test is calculated to be 2(- 44676.9+44797.96) = 242.12 . Comparing this value with the quantiles of the χ^2_1 distribution, the p-value of the test is 0.000. Thus, the cardioid-von Mises distribution provides a significant improvement comparing to the asymmetric cardioid-von Mises distribution. Also, this reflects the effect of the τ parameter in the model. Despite these findings, all models are rejected by chi-squared goodness-of-fit tests because of the very large sample size.

Nominally, 95% confidence intervals for the parameters ξ, κ, λ and τ of the CvM distribution are calculated using standard errors computed from the observed information matrix evaluated at the ML solution and the asymptotic normality of ML estimators. They are (1.28, 1.37), (1.51, 1.61), (0.52, 0.56) and (-2.18,-2.08), respectively. But those calculated using profile log-likelihood function and standard chi-squared theory are (1.29, 1.37), (1.51, 1.61), (0.52, 0.56) and (-2.18,-2.08). Clearly, both sets of intervals are very similar and the fact that none of the intervals for τ contain the value $\tau = \pi/2$ confirms the inadequacy of the asymmetric cardioid-von Mises as a model for the data.

6. Conclusions

We proposed a new transformation of circular r.v.s based on circular distribution functions in order to generate new circular distributions, called *Fg* models. We first arrived at some general results for our new models and then showed that the Möbius transformation is a special case of this transformation when *F* is the wrapped Cauchy distribution. By employing the cardioid distribution as *F*, we introduced the Cardioid-*g* class of circular distributions. Some general structural properties of our model such as modality, skewness and shape were studied. As explained in the Sec. 3, our main motivation is to provide to densities that are

mathematically tractable. We have also focused our attention primarily on the Cardioid-von Mises (CvM) family and three of its submodels. As Fig. 2 and the results in Sec. 4 show, the resulting CvM family is capable of modeling distributional forms ranging from symmetric to skew cases. The examples presented illustrate the potential of Cardioid-von Mises distributions as models for real circular data.

Acknowledgments

The first author is grateful to the Graduate Office of the University of Isfahan for their support. The authors would also like to thank anonymous reviewers whose valuable comments and suggestions led to improvement of the manuscript.

References

[1] T. Abe, A. Pewsey, Sine-skewed circular distributions, Statistical Papers 52(3) (2011a) 683–707.
[2] T. Abe, A. Pewsey, Symmetric circular models through duplication and cosine perturbation, Computational Statistics and Data Analysis 55(12) (2011b) 3271–3282.
[3] T. Abe, A. Pewsey, K. Shimizu, Extending circular distributions through transformation of argument, Annals of the Institute of Statistical Mathematics 65(5) (2013) 833–858.
[4] T. Abe, K. Shimizu, A. Pewsey, On Papakonstantinous extension of the cardioid distribution, Statistics and Probability Letters 79 (2009) 2138–2147.
[5] A. Azzalini, A class of distributions which includes the normal ones, Scandinavian Journal of Statistics 12 (1985) 171–178.
[6] E. Batschelet, Circular Statistics in Biology, Academic Press, London, 1981.
[7] J. J. Fernández-Durán, Circular distributions based on nonnegative trigonometric sums, Biometrics 60(2) (2004) 499–503.
[8] R. Gatto, S. R. Jammalamadaka, The generalized von Mises distribution, Statistical Methodology 4(3) (2007) 341–353.
[9] J. Gill, D. Hangartner, Circular data in political science and how to handle it, Political Analysis 18(3) (2010) 316–336.
[10] S. R. Jammalamadaka, A. SenGupta, Topics in Circular Statistics, World Scientific, Singapore, 2001.
[11] M. C. Jones, A. Pewsey, Inverse Batschelet distributions for circular data. Biometrics 68 (2012) 183–193.
[12] S. Kato, M. C. Jones, A family of distributions on the circle with links to, and applications arising from, Möbius transformation, Journal of the American Statistical Association 105(489) (2010) 249–262.
[13] S. Kato, M. C. Jones, A tractable and interpretable four-parameter family of unimodal distributions on the circle, Biometrika 102(1) (2014) 181–190.
[14] S. Kim, A. SenGupta, A three-parameter generalized von Mises distribution, Statistical Papers 54(3) (2013) 685–693.
[15] K. V. Mardia, Characterizations of directional distributions, In A Modern Course on Statistical Distributions in Scientific Work, Springer, Dordrecht, 1975.
[16] K. V. Mardia, P. E. Jupp, Directional Statistics, John Wiley, Chichester, 1999.
[17] V. Papakonstantinou, Beiträge zur zirkulären Statistik, Ph.D. thesis, University of Zurich, Switzerland, 1979.
[18] A. Pewsey, Testing circular symmetry, Canadian Journal of Statistics 30(4) (2002) 591–600.
[19] A. Pewsey, K. Shimizu, de la Cruz, R, On an extension of the von Mises distribution due to Batschelet, Journal of Applied Statistics 38 (2011) 1073–1085.
[20] A. Pewsey, M. Neuhäuser, G. D. Ruxton, Circular statistics in R, Oxford University Press, 2013.
[21] C. E. Shannon, A mathematical theory of communication. ACM SIGMOBILE Mobile Computing and Communications Review 5(1) (2001) 3–55.
[22] D. Umbach, S. R. Jammalamadaka, Building asymmetry into circular distributions, Statistics and Probability Letters 79(5) (2009) 659–663.
[23] T. Wehrly, R. A. Johnson, Bivariate models for dependence of angular observations and a related Markov process, Biometrika 67-1 (1980) 255–256.

Appendix

Let $Q(\theta_i, \xi, \lambda, \tau) = \theta_i - \xi + \lambda \sin(\theta_i - \xi) - \tau$. The score equations are obtained by differentiating (16) as below:

$$\frac{\partial l}{\partial \xi} = \sum_{i=1}^{n} \frac{\lambda \sin(\theta_i - \xi)}{1 + \lambda \cos(\theta_i - \xi)} + \kappa \sum_{i=1}^{n} (1 + \lambda \cos(\theta_i - \xi)) \sin(Q(\theta_i, \xi, \lambda, \tau)) = 0,$$

$$\frac{\partial l}{\partial \lambda} = \sum_{i=1}^{n} \frac{\cos(\theta_i - \xi)}{1 + \lambda \cos(\theta_i - \xi)} - \kappa \sum_{i=1}^{n} \sin(\theta_i - \xi) \sin(Q(\theta_i, \xi, \lambda, \tau)) = 0,$$

$$\frac{\partial l}{\partial \kappa} = -n \frac{I_1(\kappa)}{I_0(\kappa)} + \sum_{i=1}^{n} \cos(Q(\theta_i, \xi, \lambda, \tau)) = 0,$$

$$\frac{\partial l}{\partial \tau} = \kappa \sum_{i=1}^{n} \sin(Q(\theta_i, \xi, \lambda, \tau)) = 0,$$

The maximum likelihood (ML) estimates satisfy the score equations. The elements of the observed information matrix are the negative of the second partial derivatives of the log-likelihood of (16). Denoting them by $\ell_{\xi\xi}, \ell_{\xi\lambda}, ..., \ell_{\tau\tau}$, one obtains:

$$\ell_{\xi\xi} = \sum_{i=1}^{n} \frac{-\lambda(\cos(\theta_i - \xi) + \lambda)}{(1 + \lambda \cos(\theta_i - \xi))^2} + \kappa \sum_{i=1}^{n} \Big(\lambda \sin(\theta_i - \xi) \sin(Q(\theta_i, \xi, \lambda, \tau)) - (1 + \lambda \cos(\theta_i - \xi))^2 \cos(Q(\theta_i, \xi, \lambda, \tau)) \Big),$$

$$\ell_{\xi\lambda} = \sum_{i=1}^{n} \frac{\sin(\theta_i - \xi)}{(1 + \lambda \cos(\theta_i - \xi))^2} + \kappa \sum_{i=1}^{n} (\cos(\theta_i - \xi) \sin(Q(\theta_i, \xi, \lambda, \tau)) + \sin(\theta_i - \xi)(1 + \lambda \cos(\theta_i - \xi)) \cos(Q(\theta_i, \xi, \lambda, \tau))),$$

$$\ell_{\xi\kappa} = \sum_{i=1}^{n} (1 + \lambda \cos(\theta_i - \xi)) \sin(Q(\theta_i, \xi, \lambda, \tau)),$$

$$\ell_{\xi\tau} = -\kappa \sum_{i=1}^{n} (1 + \lambda \cos(\theta_i - \xi)) \cos(Q(\theta_i, \xi, \lambda, \tau)),$$

$$\ell_{\lambda\lambda} = -\sum_{i=1}^{n} \frac{\cos(\theta_i - \xi)^2}{(1 + \lambda \cos(\theta_i - \xi))^2} - \kappa \sum_{i=1}^{n} \sin(\theta_i - \xi)^2 \cos(Q(\theta_i, \xi, \lambda, \tau)),$$

$$\ell_{\lambda\kappa} = -\sum_{i=1}^{n} \sin(\theta_i - \xi) \sin(Q(\theta_i, \xi, \lambda, \tau)),$$

$$\ell_{\lambda\tau} = \kappa \sum_{i=1}^{n} \sin(\theta_i - \xi) \cos(Q(\theta_i, \xi, \lambda, \tau)),$$

$$\ell_{\kappa\kappa} = n \left(\frac{I_1(\kappa)^2}{I_0(\kappa)^2} - \frac{\kappa I_0(\kappa) - I_1(\kappa)}{\kappa I_0(\kappa)} \right),$$

$$\ell_{\kappa\tau} = \sum_{i=1}^{n} \sin(Q(\theta_i, \xi, \lambda, \tau)),$$

$$\ell_{\tau\tau} = -\kappa \sum_{i=1}^{n} \cos(Q(\theta_i, \xi, \lambda, \tau)).$$